



European
University
Institute

ROBERT
SCHUMAN
CENTRE FOR
ADVANCED
STUDIES

WORKING PAPERS

RSCAS 2018/03
Robert Schuman Centre for Advanced Studies
Global Governance Programme-292

Last Mile for Tuna (to a Safe Harbor)
What is the TBT Agreement all About?

Petros C. Mavroidis

European University Institute

Robert Schuman Centre for Advanced Studies

Global Governance Programme

Last Mile for Tuna (to a Safe Harbor)

What is the TBT Agreement all About?

Petros C. Mavroidis

EUI Working Paper **RSCAS** 2018/03

This text may be downloaded only for personal research purposes. Additional reproduction for other purposes, whether in hard copies or electronically, requires the consent of the author(s), editor(s). If cited or quoted, reference should be made to the full name of the author(s), editor(s), the title, the working paper, or other series, the year and the publisher.

ISSN 1028-3625

© Petros C. Mavroidis, 2018

Printed in Italy, January 2018

European University Institute

Badia Fiesolana

I – 50014 San Domenico di Fiesole (FI)

Italy

www.eui.eu/RSCAS/Publications/

www.eui.eu

cadmus.eui.eu

Robert Schuman Centre for Advanced Studies

The Robert Schuman Centre for Advanced Studies (RSCAS), created in 1992 and directed by Professor Brigid Laffan, aims to develop inter-disciplinary and comparative research and to promote work on the major issues facing the process of integration and European society.

The Centre is home to a large post-doctoral programme and hosts major research programmes and projects, and a range of working groups and *ad hoc* initiatives. The research agenda is organised around a set of core themes and is continuously evolving, reflecting the changing agenda of European integration and the expanding membership of the European Union.

Details of the research of the Centre can be found on:

<http://www.eui.eu/RSCAS/Research/>

Research publications take the form of Working Papers, Policy Papers, Policy Briefs, Distinguished Lectures, Research Project Reports and Books.

Most of these are also available on the RSCAS website:

<http://www.eui.eu/RSCAS/Publications/>

The EUI and the RSCAS are not responsible for the opinion expressed by the author(s).

The Global Governance Programme at the EUI

The Global Governance Programme is one of the flagship programmes of the Robert Schuman Centre for Advanced Studies at the European University Institute (EUI). It aims to: build a community of outstanding professors and scholars, produce high quality research and, engage with the world of practice through policy dialogue. At the Global Governance Programme, established and early career scholars research, write on and discuss, within and beyond academia, issues of global governance, focussing on four broad and interdisciplinary areas: European, Transnational and Global Governance; Global Economics; Europe in the World; and Cultural Pluralism.

The Programme also aims to contribute to the fostering of present and future generations of policy and decision makers through its unique executive training programme, the Academy of Global Governance, where theory and “real world” experience meet. At the Academy, executives, policy makers, diplomats, officials, private sector professionals and academics, have the opportunity to meet, share views and debate with leading academics, top-level officials, heads of international organisations and senior executives, on topical issues relating to governance.

For more information: <http://globalgovernanceprogramme.eui.eu>

Abstract

The WTO Agreement on TBT (Technical Barriers to Trade) aims at taming NTBs (nontariff barriers), the main instrument segmenting markets nowadays. Some of the terms used to flesh out the commitments undertaken are borrowed from the GATT, and some originate in the modern regulatory reality as expressed through SDOs (standard-development organizations). It does not share a copy-cat function with the GATT, though. Alas, the WTO Appellate Body, by understanding words as ‘invariances’, e.g., interpreting them out of context (without asking what is the purpose for the TBT?), has not only exported its GATT case law, but also misapplied it into the realm of TBT, and ended up with significant errors. In the most recent report of the tuna-saga, an ongoing dispute between Mexico and the United States, a panel corrected some of the pre-existing misunderstandings regarding the scope and function of TBT. Problems though, persist. Importantly, we cannot rely on instincts of isolated adjudicators to get it right. Only a coherent test for consistency with the TBT can serve this function. This is what we still miss. In what follows, we explain why the current approach is erroneous, and advance an alternative understanding, which could help implement the TBT in a manner faithful to its negotiating intent, and objective function.

Keywords

WTO; TBT; revealed preferences.

JEL Classification: K40

1. Tuna Has Been Swimming in the GATT/WTO Waters Since 1989*

The tuna-saga, a series of disputes between Mexico and the United States regarding commercialization of tuna in the US market, has been going on for almost thirty years now. It has earned its place among the legendary Softwood lumber-, Bananas-, Hormones-, Airbus/Boeing-, and the DISC/FSC disputes. Unlike many of them though, we are yet to see the end of it. Or, so it seems. In part at least, this is due to the fact that the WTO courts have not managed to explain in clear terms what the United States should do. In turn, this is due to their convoluted understanding of the TBT Agreement.

Following repeated condemnations of its measures by a GATT panel already in 1991, the United States eventually allowed all tuna, no matter how it had been harvested, to be sold in its market. It reserved nevertheless the label “dolphin-safe” to tuna harvested in particular manner, and following specific verification procedures. Mexico complained and prevailed yet again. The United States modified some aspects of its measures, without however, touching upon its quintessential elements. The amount of proof required to show that tuna could legitimately have access to the “dolphin-safe” label would depend on the likelihood that dolphin life had been accidentally taken when fishing for tuna (that is, the US measures were ‘calibrated’ to the likelihood of accidentally taking the life of dolphins when harvesting for tuna).

The most recent reincarnation of the US measure was finally found to be WTO-consistent by a ‘compliance panel’, that is, a WTO ‘court’ examining the consistency of US implementing measures with an adverse decision by a WTO panel or the Appellate Body. This report has already done a lot of good, but its fate is uncertain as it is currently under appeal. Irrespective of the eventual outcome though, there is still considerable way to go, since this report changed the outcome of the dispute without changing the test of consistency with the TBT that the Appellate Body had already elaborated. This is what we attempt to do in this paper.

Our argument runs like this. The TBT Agreement is asking different questions than the GATT. Its function is not to insure against erosion of tariff concessions. It is that as well. Its main function though, is to ensure that NTBs will not be erected when they do not genuinely serve a legitimate regulatory objective. In this vein, the agreement imposes a three-prong test, whereby members will intervene only when necessary to advance an objective, will adopt the least restrictive measures when doing so, and will apply adopted measures in non-discriminatory manner. Unilateral measures will be adopted, only if existing international standards, which are presumed to be necessary (e.g., the least restrictive option to advance an objective), cannot be appropriately used.

The architecture thus, of the TBT Agreement inherently carries a sequence, in the sense that members should first see to what extent an international standard can be appropriately used to serve their preference. Recourse to unilateral measures is appropriate only when response to this question is negative.

There is another sequence though, embedded in the TBT Agreement, which is the focus of this paper, and which the Appellate Body has totally overlooked. The key difference between the GATT and the TBT is that the latter requires the adoption not simply of non-discriminatory measures, but crucially of necessary measures (a subset of all possible non-discriminatory measures) that should be applied in even-handed manner. Necessity is presumed when international standards have been privileged, but not when, in the absence of appropriate international standards, recourse to unilateral measures is made.

* For helpful comments on previous draft and numerous discussions on this issue, I thank Tom Bollyky, Cary Coglianesi, Rob Howse, Damien J. Neven, Neeraj RS, André Sapir, and Alan O. Sykes. Henrik Horn af Rantzen deserves a special mention in light of numerous ongoing discussions and repeated comments on previous drafts. Remaining errors are his own.

The protectionism-test (since non-discrimination is legalese for absence of protectionism) is thus embedded in the non-discrimination provision in the GATT-context, but in the necessity-requirement in the TBT. Non-discrimination in TBT is a follow-up question, whereby the issue is whether A and B have received the same treatment assuming that they have both adopted necessary measures.

The TBT is, in other, words akin to Articles XX(b) or XX(d) of GATT, which also reflect the necessity-requirement, and the Appellate Body would be well-advised to follow the sequential analysis that it has adopted there. Alas, the WTO Appellate Body has completely ignored it. Unless sequence is acknowledged and embedded in the test of consistency with the TBT, the risk of erroneous outcomes cannot be discarded.

To sustain our main claim, we go through the case law on TBT. Talking about erroneous outcomes, the United States has been on the receiving end twice. Its legislation reserving the label “dolphin-safe” only to tuna that had demonstrably been fished in a manner not endangering dolphin life was judged WTO-inconsistent for unfathomable, as we explain, reasons. Its labelling scheme obliging producers to indicate place of birth, slaughtering etc. for bovine meat, was also found to be WTO-inconsistent, even though the challenged measure was not judged more trade-restrictive than necessary, and, in principle, was applied across goods irrespective of their origin.¹

The errors by the Appellate Body in interpreting the TBT are not confined to the understanding of the obligations imposed on unilateral measures, such as those employed in the tuna disputes. The Appellate Body has, unnecessarily so, created havoc regarding the policy relevance of international standards, one of the pillars of the TBT. Peru had a narrow escape in EC-Sardines, when the Appellate Body forgot for a moment its “textualist” self and adopted an interpretation of the term “except”, featured prominently in the body of Article 2.4 of TBT, which is hard to reconcile with the intrinsic meaning of this term.

In short, case law under the TBT leaves a lot to be desired. In this paper, we use the tuna disputes as background to first explain what is wrong with the current test for consistency as developed by the Appellate Body, why we ended up with so many erroneous interpretations, and then advance our own preferred test for consistency with the TBT Agreement. Section 2 explains the test for consistency with the TBT Agreement as developed in successive Appellate Body reports. In this Section, we also examine what changes with the recent ruling of the compliance panel. Section 3 is divided in two parts. We first discuss what, in our view, is still wrong with the test for consistency developed in case law, the recent improvements notwithstanding, and then we ask why this has been the case. It is in Section 4 that we advance our suggested approach, while Section 5 recaps the main conclusions.

2. Background

The choice of the tuna disputes as background for our critical evaluation of the case law does not suffer from sample bias. Panels and the Appellate Body have not seen in any of the reports issued sequence across the obligations embedded in the TBT.

¹ We will not be focusing on this dispute, since the legal reasoning of the Appellate Body here echoes that followed in the tuna disputes anyway. Mavroidis and Saggi (2014) discuss the US-COOL dispute dealing with the labeling scheme for bovine meat in detail. Marceau and Trachtman (2003) provide a very good overview of the early TBT disputes, whereas Howse and Levy (2012) discuss the later TBT disputes that essentially gave rise to the standard of review adopted by panels and the Appellate Body as we now know it.

2.1 The Importance of the Issue

The test of consistency with the TBT Agreement, when it comes to unilateral acts, has been developed in three reports, namely, US-Clove Cigarettes, US-COOL (Canada),² and US-Tuna II (Mexico). The approach developed though, has an impact on a much wider class of cases:

- 52 disputes, where at least one claim coming under the ambit of the TBT Agreement, have been formally raised since the advent of the WTO.³ This suggests a rate of more than 2 cases yearly, if the current trend continues, approximately that is, the annual workload of the International Court of Justice;
- In March 2016, the 500th⁴ specific trade concern (STC) was raised. STCs are a hybrid between transparency and dispute adjudication.⁵ One would expect that contract completion through the case law regarding the consistency with the TBT Agreement, will affect the discussions in the realm of STCs;
- And of course, the Agreement on Sanitary and Phyto-Sanitary Measures (SPS), which endorses a similar test for consistency and its own STCs as well will also be heavily influenced by the test established under TBT case law.

One can thus, easily appreciate that a wide realm of WTO practice will be influenced, actually, is being influenced by the current test for consistency.

2.2 The Appellate Body Understanding of the TBT Agreement

The Appellate Body has held that:

- The obligations regarding consistency of technical regulations (standards, and conformity assessment),⁶ namely, non-discrimination, and necessity are independent from each other;
- Non-discrimination has two legs, since complainant must show that two goods are like, and imported goods are afforded less favourable treatment than domestic goods
 - The term “like goods” should be understood in the same way as in its GATT case law, namely, that it is consumers that decide on likeness
 - Conversely, the term “less favourable treatment” should be understood in GATT-unlike manner. Contrary to its GATT case-law, regulatory distinctions that lead to disparate trade outcomes are not discriminatory if they are not origin-based. Note though, that the (disparate) trade outcome must be exclusively due to the (legitimate) regulatory distinction;

² There were two distinct reports against the US in this dispute, DS383 US-COOL (Canada), and DS384 US-COOL (Mexico), since both Canada and Mexico complained against the rule of origin system practiced by the United States. The two reports are identical in substance.

³ Our cut-off date is December 31, 2017. All information regarding disputes formally raised before the WTO can be accessed at www.wto.org

⁴ https://www.wto.org/english/news_e/news16_e/tbt_11mar16_e.htm

⁵ Horn et al. (2013) discuss the origins, function and practice of STCs.

⁶ The TBT Agreement covers four instruments: international standards (e.g., standards adopted by the ISO, International Standardization Organization), technical regulations, standards (instruments regulating the production process of goods compliance with which is necessary condition for market access with respect to the former, but not to the latter), and conformity assessment, that is the procedure allowing the importing state to verify whether imports conform to the content of international standards, technical regulations, or standards. International standards are presumed to meet the necessity-criterion (Article 2.5 of TBT). There is no case so far where a panel had to review a claim that an international standard had not met the necessity-criterion. For a detailed discussion of this issue, see Mavroidis (2016), vol. 2, pp. 389 et seq.

- Necessity has the same content in GATT as in TBT, that is, it requires from WTO members to reduce cost-shifting by adopting the least trade restrictive measures in order to pursue legitimate regulatory objectives.⁷

2.2.1 No Sequence

There is no need to delve too much on this point. Suffice to repeat that neither the Appellate Body nor panels have raised this point in any of the reports discussing claims under the TBT.

2.2.2 Necessity

In US-Tuna II (Mexico), the Appellate Body found that the US measure was necessary to reach the legitimate objective pursued, but discriminatory. The United States had in place a measure, whereby vessels fishing for tuna observe enhanced verification and reporting requirements (certifying no incidental taking of dolphin life) when fishing in areas with high concentration of dolphins than when fishing in areas with low concentration of dolphins. The measure was thus, calibrated to the likelihood that dolphins could be accidentally killed (high likelihood would lead to enhanced reporting requirements, whereas low likelihood would lead to ‘lighter’ requirements).

The Appellate Body found that the US reporting requirements were effectively calibrated to the magnitude of risk of encountering dolphins when fishing for tuna. It was thus, necessary to achieve its stated objective. In doing that, it employed the classic GATT-analysis regarding the “necessity”-requirement:

- In light of the objective pursued, is the measure the least restrictive option (that is, the option that has the least negative impact on international trade)?
- If yes, is it reasonably available to the regulating state?
- If no (if it imposes undue hardship on regulator), is the next in line least restrictive measure reasonably available to it?

2.2.3 Non-discrimination

In US-Tuna II (Mexico), the Appellate Body asked whether consumers would treat dolphin-safe and – unsafe tuna as like products. Having responded affirmatively to this question, the Appellate Body went on to ask whether the regulatory distinction operated by the United States afforded less favourable treatment to Mexican tuna. It asked whether the (disparate) trade impact on Mexican producers was exclusively due to the legitimate regulatory distinction operated in US law.

It went ahead to find that this measure was discriminatory because most Mexican vessels were fishing in the area with high concentration of dolphins and, hence, had to observe the more onerous reporting requirements. The US measure was thus, found to be discriminatory, but necessary.

2.3 Other Issues

There are two more issues that the Appellate Body case law had to grapple with, and, as we show later, failed to rise to the occasion: the allocation of the burden of proof when a WTO member has deviated from a standard, and the understanding of the term ‘standard’. We take them in turn.

⁷ Mavroidis (2013) describes in detail the findings of the three reports.

2.3.1 Allocation of Burden of Proof (International Standards)

In EC-Sardines, the Appellate Body held that, in case of deviation from an international standard, the allocation of the burden of production of proof should not shift to the member deviating from the standard (and, obviously, possessing the information why it had done so).⁸ The Appellate Body justified its approach stating that the complainant would not be disadvantaged, since the TBT Agreement contained many transparency obligations and, as a result, all WTO members are well-acquainted with measures coming under its ambit.

2.3.2 Standards

In US–Tuna II (Mexico), traders were not banned from accessing the US market if their tuna did not qualify as “dolphin-safe”. The US position was thus, that the measure adopted was a standard, and not a technical regulation. Mexico protested, arguing that the label conferred an advantage to those traders that could have access to it, and that, by restricting the use of the “dolphin-safe” label to only those fishers who used a specific harvesting technique, the United States had enacted a technical regulation, not a standard.

The panel agreed with Mexico. What mattered was that traders did not have unconditional access to the use of the label (§§ 7.100 et seq. especially 7.120 and 7.131). A separate opinion issued by one member of the panel stressed that, what was relevant for the qualification of the measure as a technical regulation or standard, was ultimately whether imports of tuna could (or could not) be lawfully traded in the US market without the label “dolphin-safe.” Evidence showed that Mexican traders that did not conform to the criteria established could sell their products in the US market anyway, even though sales had taken a hit. They would have to market them in this case as “tuna,” not “dolphin-safe tuna.” Consequently, the measure at hand, in this view, was a standard (§§ 7.146 et seq.).

The Appellate Body in §§172–199 endorsed the majority view, holding that the US measure was indeed a technical regulation. In its view, what mattered was that goods had to observe the statutory requirements (and had no discretion at all to this effect), otherwise, they could not legitimately be considered “dolphin-safe.”

2.4 The Facts before the Second Compliance Panel Report

Following the condemnation of its policies, the United States adopted implementing legislation in order to address the concerns of the Appellate Body. Its first effort was thwarted. In US-Tuna II (Mexico) (Article 21.5-Second Recourse), the second implementing effort by the United States, the US measure, the consistency of which with the WTO rules was at stake,⁹ was as follows:

- Tuna caught through two fishing methods, namely, purse seine nets with setting on dolphins and driftnet, can never be eligible for the “dolphin-safe” label;
- Tuna caught using other methods of fishing is, in principle, eligible for the “dolphin-safe” label, if it is proved that no dolphins were killed or were seriously injured through the gear deployment or fishing method used (these are the so-called “eligibility” criteria, in the sense that tuna caught with

⁸ Horn and Weiler (2004) explain the repercussions of this approach to the (continuing) relevance of international standards in WTO law.

⁹ A side remark is in order here. It is questionable whether the now well-embedded practice of having two compliance panels in the same dispute is consistent with the letter and the spirit of the WTO Dispute Settlement Understanding (DSU). Losing defendants in WTO litigation have only one reasonable period of time within which they can bring their measures into compliance. The second compliance panel by definition examines measures that were adopted outside the reasonable period of time, that is, new measures, that should normally be adjudicated before a new panel. As said though, all this is by now water under the bridge.

any but the two aforementioned methods is, in principle, eligible to access the “dolphin-safe” label);

- To prove that tuna can legitimately carry the “dolphin-safe” label, interested producers must observe certification, tracking and verification, as well as determination requirements:
 - For tuna caught in the ETP (Eastern Tropical Pacific), certification requirements are heavier: both the captain of the ship as well as an independent observer must attest to the fact dolphins were killed or injured);
 - For tuna caught outside the ETP, certification requirements are “lighter”: the captain of the boat, having received training to this effect, can confirm whether dolphins were killed or injured, and there is no need to also have an observer on board);
 - The differentiation between ETP et non-ETP tuna is predicated on the fact that it is in ETP that high concentration of dolphins is observed, and most of the setting on of dolphins takes place;
 - The US authorities can determine that an observer must be on board and certify that dolphin-life has not been endangered even for non-ETP fisheries, if the fisher(s) at hand have a record of systematic accidental killing or endangerment of dolphin-life;
 - Finally, the new, modified measure added heightened tracking and verification requirements (processors and importers of tuna must collect and keep information for two years so as to enable the US authorities to track non-“dolphin-safe” tuna back to the timing of its harvesting). Not only fishers harvesting tuna in the ETP, but also fishers outside the ETP have to observe this requirement as well.

The US measure is thus, calibrated to an “exogenous” event, namely, the likelihood to endanger dolphin life when fishing for tuna in areas with high and low concentration of dolphins. It is thus, markedly unlike the situation in EU-Seal Products, where the challenged measure had been “calibrated” to whatever the domestic political economy could take.¹⁰

2.4.1 What Changed with this Report?

The panel had in front of it one important new element only: the new tracking and verification requirements that the US authorities were now imposing on fisheries outside the ETP. Calibration of the measure to the likelihood of endangering dolphin life continued to be the key element of the US measure. The objectives of the US measure continued to be the same, namely, protection of consumers, as well as protection of dolphin life, as the panel itself explicitly acknowledged (§7.49).

The regulatory changes between the first and the second compliance procedure should thus, not be exaggerated.

To decide whether the US measure was discriminatory, the panel focused its attention on whether it afforded less favourable treatment to Mexican tuna. It decided that this had not been the case (§§7.529-717). It first echoed the legal standard developed by the Appellate Body, according to which no less favourable treatment can result if disparate effects for imported goods are exclusively attributed to a legitimate regulatory distinction. Since panels and the Appellate Body have construed the TBT *à la* GATT, that is an instrument protecting competitive conditions, this panel did not look into actual trade effects. It held that, because the variation in the regulatory requirements had been calibrated to the risk, no less favourable treatment could have ever resulted for Mexican tuna.

¹⁰ Conconi and Voon (2016) discuss this issue in detail.

2.4.2 What Stayed the Same?

This panel did not touch on the understanding of the term “like products”, as well of the “necessity”-requirements. It is consumers that will decide on whether two goods are like. Furthermore, WTO members must adopt the least restrictive measure reasonably available to them in pursuance of their unilaterally set legitimate objective.

2.4.3 End Result: United States Wins (For Now)

Against this background, the panel found that the US measure was not discriminatory, and exonerated the United States from responsibility. Whereas the outcome makes sense though, there is no guarantee that it will be reproduced in future disputes.¹¹ The reason why this is the case is the fact that the underlying test of consistency has not been modified. We explain in what now follows.

3. What is Wrong with the Current Understanding of TBT and Why?

The Appellate Body has committed, in our view, four errors when addressing claims under the TBT:

- It interpreted the two obligations regarding technical regulations and/or standards (namely, non-discrimination, and necessity) as if they were independent from each other;
- Second, and because of this error, it also adopted erroneous understanding of the term “like products”;
- Third, it reduced the relevance of international standards, against the expressed negotiating will; and
- Finally, the distinction between: technical regulations” on the one hand, and “standards” on the other, became a line in the sand (although statutory language had set it in stone).

3.1 What Is Wrong?

The most basic error in the current test for consistency is the construction of the obligations embedded in Article 2.1 (non-discrimination), and 2.2 (necessity) of TBT as two parallel, independent obligations. We start from this.

3.1.1 Lack of Sequence Can Lead to Absurd Results

Assume that the tuna dispute had been adjudicated under the GATT, and not under the TBT. Article III.4 of GATT would have been the relevant provision to entertain the dispute.

The panel would have to ask if “dolphin-safe”, and “dolphin-unsafe” tuna were like products. Since consumers define likeness, and, as we know from consistent case-law, the response to this question should be in the affirmative.

The next question would be whether the United States, by allowing (some) US tuna to carry the “dolphin-safe” label, and disallowing Mexican tuna to do the same would be according the latter less favourable treatment than the former. GATT case-law has gone full circle here. In *Korea-Various Measures*, the Appellate Body had held that modification of conditions of competition to the detriment

¹¹ Recall that it was the introduction of flexibility (in the sense of allowing for trading of shrimps harvested through fishing methods other than turtle excluding devices) was the lynchpin for the Appellate Body to accept that the US has brought its measure into compliance with its obligations in the *US-Shrimp* litigation. And it is introduction of flexibility that did the trick in the tuna-dispute as well.

of imported goods equalled less favourable treatment (§149).¹² Later, in Dominican Republic-Import and Sale of Cigarettes, the Appellate Body held that detrimental effects would not amount to less favourable treatment if the reason for it was a regulatory distinction unrelated to the origin of goods traded (§96). In EC-Seal Products, the Appellate Body returned to its Korea understanding, and thus, all disparate effects would amount to less favourable treatment (§5.109-110).

If the Appellate Body followed its jurisprudence under the Dominican Republic case, then it would have concluded that no violation of Article III.4 of GATT had occurred. If it followed its jurisprudence under either of the other two disputes, the opposite would have been the case. Defendants would then have to invoke Article XX of GATT to justify their violation. This is what the United States, in our case, would do.

Assume that the United States chose to invoke Article XX(b) of GATT.¹³ The measure would be considered necessary anyway (since it was considered necessary by the original Appellate Body and the two compliance panels). The last remaining question would be whether it had been applied in even-handed manner.

Recall US-Shrimp. Once persuaded that the measure (banning imports of shrimps that had not been harvested with turtle excluding devices, TEDs) was relating to the protection of an exhaustible natural resource (sea turtles), the Appellate Body asked whether the measure had been applied in non-discriminatory (even-handed) manner. It held that it was not the case. The reasoning nevertheless matters. It concluded as much, only because the US objective could have been reached through other fishing methods as efficient as use of TEDs. Implicitly thus, it accepted the legitimacy of the US distinction between shrimps harvested with TEDs (as well as with other equally efficient methods), and shrimp harvested in a manner that endangered the life of sea turtles. It did not ask consumers whether they would treat the two sorts of shrimps as like goods. When in the realm of Article XX of GATT, there is no more room for market-likeness.

Think now, of ETP and non-ETP as two WTO members. In the former, the risk of killing dolphin is high; in the latter, it is not. These are not two countries where the same conditions prevail. Imposing thus, differentiated regulatory requirements would be a perfectly legitimate measure in light of the regulatory objective pursued.

This is not of course, what the Appellate Body did in all its reports in the tuna-saga. By reverting to ask consumers' opinions on likeness, and then ask whether the detrimental impact for Mexican tuna was exclusively due to a legitimate regulatory distinction, the Appellate Body misapplied its case law under the GATT, the basic thrust of which it had borrowed for its analysis. It held that the measure was discriminatory, although properly calibrated to the risk. The second compliance panel did not modify the test. It simply, modified the outcome holding that the measure was non-discriminatory because it had now been properly calibrated to the risk (because of the new tracking and verification requirements).

If the United States would have prevailed under Article XX(b) of GATT, as we have shown here, should it then be in the receiving end under the TBT? This is a question the Appellate Body should have asked, but did not ask.

¹² In Mavroidis (2016, pp. 385 et seq.), we held that this is a deplorable outcome. The Korean measure had led to fewer outlets selling imported beef because of the legal import quota that had been in place (since traders would rather sell the cheaper imported, rather than the domestic more expensive beef). Furthermore, all of the quota for imported beef had been routinely sold in the Korean market. If at all, the number of outlets should have been treated as a services-, not a goods issue.

¹³ Following the US-Shrimp jurisprudence, it could of course, invoke Article XX(g) of GATT, since living organisms as well are considered "exhaustible natural resources", and thus have an even easier road to victory. It does not matter though, for the reasons we explain here.

The US measure was found twice to violate the TBT, even though it would have sailed through the GATT. Worse, under the TBT, the burden of proof stays with Mexico, whereas in the GATT-context it shifts to the United States only after Mexico has established a violation of Article III of GATT.

The original, as well as the compliance Appellate Body, effectively asked the question whether the measure is calibrated to the risk twice, once under the necessity- and once under the non-discrimination requirements. This as a matter of pure logic cannot be. The necessity- and the non-discrimination requirements have differentiated scope.

Necessity refers to the policy tools to serve a given objective. In our case, the question is whether it is necessary to condition use of the “dolphin-safe” label to the likelihood that dolphin life has been endangered. Yes, responds the Appellate Body.

Non-discrimination would then respond to the question whether the United States has been treating goods (tuna) in the high- and the low risk areas the same, irrespective of their origin. Yes, is the response again, as a matter of fact.

The Appellate Body did state a number of times that the TBT combines elements of Articles III and XX of GATT. Consequently, by its own standard, the Appellate Body should have found that the US measure was overall TBT-consistent since it satisfies, as we have seen, the requirements of Article XX(b), which are substantially more stringent than those in Article III of GATT, since they provide a narrow exception to basic, a fundamental obligation to not discriminate. The opposite would be absurd. Absurd happened, though.

3.1.2 Consumers’ Preferences Are Immaterial

In the TBT-context, consumers’ perceptions for likeness are immaterial. To start with, if government and consumers saw eye to eye, if they shared preferences, there would be no need to enact technical regulations or standards. Governments intervene either because they have private information (this is very often the case for example, in the realm of environmental legislation), or simply because it has different preferences from (some) consumers (as in our case). Governments have of course, the right to do that under domestic constitutions, and the WTO (TBT) has not questioned this right. After all, the WTO is a negative integration-contract, where preferences are set unilaterally and, to the extent that by embedding them into regulations they shift costs to trading partners, must observe non-discrimination (and in the case of TBT, the necessity-principle, as well).

What matters thus, is not the consumer definition of likeness. What matters is the government definition of likeness. Likeness in TBT is not market-, it is policy-likeness. It is for the United States to decide on what is “dolphin-safe” tuna. The TBT contains, of course, important safeguards against abusive interventions.

First, governments must think whether interventions are necessary at all. They should go ahead, only in presence of evidence that non-intervention would be costly. Second, measures must be necessary to achieve a legitimate regulatory objective. The agreement contains an indicative list to this effect, which informs users of what the framers had in mind, and help adjudicators to avoid false negatives. Third, and assuming its measure is necessary, that is, assuming the “dolphin-safe” label represents the least restrictive means to reach its ends (protection of consumers and life of dolphins), the measure must be applied in non-discriminatory manner.

Thus, the only remaining question for the Appellate Body, having satisfied itself on the necessity of the US measure, should be whether Mexican tuna that meets the US definition has been treated in the same manner as US tuna.

Why is this case? Well, by accepting the necessity to distinguish between “dolphin-safe” and “dolphin-unsafe” tuna, the Appellate Body has accepted the legitimacy of distinguishing between two classes of like products in order to serve the overall legitimate objective, namely, the protection of

consumers and the protection of dolphin life. The only remaining question should be whether goods falling under each category are treated in even-handed manner, irrespective of their origin.

Going back to market-likeness is putting into question is tantamount to asking the same question twice, once from the perspective of government intervention that is deemed necessary, and once by totally neglecting the necessity for intervention.

Finally, think of a case where an international standard has been used. Assume for example, that an ISO¹⁴ standard exists for “dolphin-safe” tuna. Assume further, that the United States refuses to apply it in non-discriminatory manner, the claims by Mexicans that their tuna meets the standard notwithstanding. Would the Appellate Body in this case ask what the US consumers’ view on likeness is, or would it ask if Mexican tuna conforms to the ISO standard? Obviously, the latter would be the case. International standards are presumed necessary (Article 2.5) but must be applied in non-discriminatory manner. Technical regulations and/or standards are not presumed necessary and must be applied in non-discriminatory manner. The latter requirement is identical across international standards on the one hand, and technical regulations and/or standards on the other. To the extent thus, that technical regulations and/or standards have been found to comply with the necessity-requirement, it is their content, and not consumers’ preferences that should provide the benchmark for likeness.

3.1.3 Detriment Exclusively Caused by Legitimate Distinction

For some unfathomable reason the Appellate Body, in its report on US-Tuna II (Mexico), held that a measure producing detrimental effects will still not considered to accord less favourable treatment to imported goods, if detrimental impact is due exclusively to a legitimate regulatory distinction (§297). This is a very demanding test, and quite unwarranted if the purpose of this provision, as the Appellate Body itself has explicitly admitted in numerous reports, is to ensure that domestic and imported goods will be treated in even-handed manner. We explain.

If we were to unravel the test as applied, we would first need to be clear about its constituent elements:

- If the challenged measure modifies the conditions of competition
- To the detriment of imported goods
- And the detrimental impact results exclusively from a legitimate regulatory distinction
- Then, the measure does not afford less favourable treatment to imported goods.

Modification of conditions of competition should not be an issue. After all, WTO members have signed to a negative integration contract, and retained the right to adopt any policies they wish to the extent that they apply them in even-handed manner.

Some detrimental impact for imported goods (e.g., the even-handedness requirement must not have been adhered to) must result for imported goods, otherwise less favourable treatment has not been accorded.¹⁵ The first question is how should one measure the relative detrimental impact. Imported goods must be hit harder than domestic goods, that much is clear. But by how much?

The simple answer is that we do not know. The Appellate Body has discarded the relevance of trade effects. So, the best evaluation we can have here is some sort of sophisticated qualitative guess as to how trade trends will evolve because of the measure adopted (and challenged).

¹⁴ ISO stands for International Organization for Standardization. It is the only SDO explicitly mentioned in the TBT.

¹⁵ Modification of conditions of competition that does not result in detrimental impact for imported goods can still be the object of a non-violation complaint. The study of this eventuality though, escapes the purview of this paper.

But then comes the even bigger thorn. Assuming the Appellate Body guesstimates that the hit will be harder on imported goods, it must then ask whether the (contemplated) impact results exclusively from a legitimate regulatory distinction. This is quite a hurdle because of the context of the TBT.

The TBT does not request the adoption of first-best, but of necessary policies. Negative external effects are minimized, even though there is no absolute guarantee that they are eliminated altogether, when first-best policies are adopted.¹⁶ When necessary policies are adopted, then we are by definition in presence of some burden on international trade, even though this burden is reduced if, for example, the counterfactual had been the adoption of non-discriminatory (but not necessary) measures. The presence of negative external effects might influence the trading of imported goods in one way or another. How then can the detrimental impact be exclusively due to the legitimate regulatory distinction?

Furthermore, other (than regulation) factors might influence the outcome. Is it for example, crystal clear that Mexican producers suffered damage exclusively from the “dolphin-safe” label? What if some consumers would anyway buy “dolphin-safe” tuna? Or what if fish lovers turn to halibut because there has been an oversupply of halibut in a given year in the US market, and, as a result, its price is now much lower than that of tuna?

The Appellate Body routinely turns a blind eye to any sophisticated analysis when various factors affect one outcome, and this is irrespective whether we are dealing with a dispute in the TBT-, the Antidumping- or the Subsidies context. Indeed, the recent panel report on US-Tuna II (Mexico) (Article 21.5-Mexico) found that the US measure met the exclusivity-requirement because it was calibrated to the likelihood of endangering life of dolphins (§7.717). This finding makes sense when it comes to discussing the necessity-requirement, not the question whether the detrimental impact was due exclusively to the legitimate regulatory distinction.

3.1.4 Allocation of Burden of Proof is Wrong

In EC-Sardines, Peru had argued before the panel that, since it was the European Union that had deviated from an appropriate international standard, it was for it to explain why. The panel agreed (§§7.48 et seq.). The Appellate Body overturned this finding, arguing that there was nothing exceptional about a deviation from an international standard, and adding that anyway Peru, because of the elaborate transparency-related obligations embedded in the TBT, should be well aware of the rationale for the measures adopted by the European Union (§§285 et seq.).¹⁷

This simply cannot be. First, there is a clear legislative mandate stating that international standards must be followed, except when they cannot appropriately serve as basis. The term “exception” has been relied in all reports dealing with claims Article XX of GATT as the basis for allocating the burden of proof to the party invoking one of the grounds mentioned therein.

Second, and more substantively so, what if the European Union (or any WTO member to this effect) had withheld information? What if it had behaved in non-transparent manner? Indeed, in cases of voluntary deviation from an international standard, similar behaviour would be incentive-compatible. How can then, the transparency-obligations be of help to complainants in similar circumstances? Could Peru possess the rationale for EU’s decision to deviate from the international standard, had the latter decided not to communicate it to the TBT Committee?

Or, what if the notification is inadequate? What if for example, the notifying member failed to explain in detail the rationale for a measure that deviates from an international standard? Coglianes

¹⁶ Bhagwati and Ramaswami (1963).

¹⁷ The Appellate Body has repeated this allocation of burden of proof in subsequent cases, and most recently in its report on Indonesia-Import Licensing Regimes (United States) in §§5.49 et seq.

(2009) usefully distinguishes between “fishbowl”- and “reasoned” transparency. The first covers cases where a measure is simply described. The latter extends to explanations of the rationale for the notified measure. According to the TBT (Article 2.9.1 and 2.9.2) a member must submit a brief only description of the objective and rationale of the measure, and additional information will be provided only upon request (Article 2.9.3). How then, can complainants protect their interests when they have partial only knowledge of the objectives and rationale of the challenged measure? Did the Appellate Body implicitly add a requirement of due diligence, which would require from WTO members to ask questions regarding the rationale for the measure (as opposed to enjoying discretion to this effect, as per the current architecture of the TBT)? Recall, that it is the same Appellate Body, which has consistently held that the policy rationale matters when it comes to evaluation whether the non-discrimination obligation has been adhered to.

One might further ask, from a policy perspective, whether this finding eviscerates the incentive to adopt international standards. Indeed, harmonization is one of the key objectives of the TBT, and relegating the policy relevance of international standards might act as disincentive to pursue harmonizing efforts (and thus, defeat a key objective of the TBT).¹⁸

To complete the discussion, we should probably mention that Peru managed to win its argument. Nevertheless, it is quite remarkable that Peru did not add anything before the Appellate Body to what it had already argued before the panel. That is, Peru won by providing the same quantum of evidence in a context where the burden of production of proof rested with the other party, and in a context where it had to assume it itself. This in and of itself is quite telling of the functionality of the allocation of the burden of production of proof operated by the Appellate Body.

3.1.5 No Room for Standards Anymore?

Standards, as opposed to technical regulations, do not condition market access for a relevant product market upon the satisfaction of statutory requirements. They can and do have a market effect, but they do not exhibit the binary function of technical regulations: either products meet the statutory requirements and are granted market access, or they do not and stay out of the market as a result.

Were in the case before us, the United States to accept only “dolphin-safe” tuna in its market, then yes, we would have been in presence of a technical regulation. By allowing “dolphin-safe” and “dolphin-unsafe” tuna into its market, the United States thought (legitimately so) that it had enacted a standard.

The Appellate Body disagreed. It held that, because the United States was not allowing all traders to use the “dolphin-safe” label, it was effectively (de facto) imposing a technical regulation on Mexican fishers.

It seems that the Appellate Body confused two separate issues: binding language regarding access to the tuna market, and binding language regarding access to the “dolphin-safe” tuna market. The former is a technical regulation, and the latter is a standard.

Consistency with standards-requirements must be safeguarded, of course. Standards are denied their *raison d'être*, if recourse to them is open even for products that fall short of meeting the established statutory requirements. Consumers will not be protected at all if, for example, products can carry the “dolphin-safe” label regardless of whether the tuna sold has been fished in a dolphin-friendly manner. Recall, that all panels and the Appellate Body acknowledged that the objective pursued by the

¹⁸ On this score, see the excellent analysis in Horn and Weiler (2004). The authors underscore that by lowering the burden of persuasion of Peru, the relevance of international standards has been strengthened. This sounds plausible. The question though, remains whether panels and the Appellate Body will consistently be satisfied with similar burden of persuasion in future case law as well. This remains to be seen. What is clear for now, is that the burden of production of proof does not switch in case of deviation from international standard.

United States was to inform consumers about the manner in which traded tuna had been fished. Opening up the use to the “dolphin-safe” label to all tuna marketed, regardless of the harvesting method, would have led to misinforming consumers. The United States would have never achieved its stated (legitimate) objective.

3.1.5 Necessary but Discriminatory

Our analysis suggests that there is room for necessary measures to be found discriminatory in case measures conforming to the substantive requirements that the regulating state has enacted do not profit from the treatment reserved in the challenged measure. Conformity assessment thus, holds the key to similar findings.

Alas, this is not what happened neither in US-COOL, nor in US-Tuna II (Mexico). Take the latter case, which is the focus of this paper. Mexican tuna was denied access to the label simply because it did not conform to the US measure judged necessary, and for no other necessary. And yet the Appellate Body in its original report found the US measure to be discriminatory because consumers treated dolphin-safe and -unsafe tuna as like goods, and the disparate impact on Mexican exports had not been exclusively due to the US regulatory distinction that the Appellate Body itself had judged necessary.

This is what must have confused the United States in the first place, and could not design the implementing measure necessary to bring its measures into compliance with the TBT. The point is this. Unless non-discrimination is understood as even-handed application of necessary measures, we risk seeing dozens of compliance panels in the future as well. Conformity assessment (to which, the Appellate Body itself opened the door in the GATT-context as well, by requesting from the United States to accept shrimps harvested with technology equally efficient to TEDs) holds the key in this context.

3.2 *Why Is the Appellate Body Wrong?*

The Appellate Body has committed the cardinal mistake that no interpreter should ever commit: it did not ask the question what is the rationale explaining the advent of the TBT Agreement? It transposed, in haphazard manner as we have noted, its interpretation of terms that the GATT and TBT share. The same terms nevertheless, can have a different meaning in two different contexts. One does not have to delve into Wittgenstein to reject the view that words are ‘invariances’, e.g., context-independent.

3.2.1 GATT-Think

GATT-think is genial in its simplicity.¹⁹ ‘Protection’ is reduced to one instrument, namely, tariffs. Disciplines on all other instruments affecting trade (including domestic instruments such as those now covered by the TBT Agreement) aim at ensuring that the value of concessions will be eroded. This insurance policy was necessary, otherwise trading partners might have lost the incentive to continue to negotiating the level of tariffs downwards.

Baldwin (1970), in his classic account, defended this integration process. Tariffs and import quotas were blurring the size of the bite of ‘snags’ that were lying behind the border. They had to be tamed first. The ‘snags’ were of course, the various domestic policies. Because of policy substitution, domestic instruments could play the role of tariffs. The rational short- to mid-term solution was to make sure that snags do not bite imported goods more than their domestic counterparts with which they were in competition until they could be properly evaluated. Non-discrimination that they had to

¹⁹ Irwin et al. (2008) provide the full negotiating record on this score.

observe, would thus, reduce the impact of the bite. The consequence was that costs would be shifted in even-handed manner to domestic and foreign products.

Unnecessary costs nonetheless, could be shifted anyway, since trading nations do not have to ‘efficiently’ address distortions, or even address distortions at all. Any regulation, to the extent non-discriminatory, would pass the test of consistency with the GATT, even if it resulted in unnecessary costs for international trade, even if the instrument used to achieve an objective was totally inefficient.²⁰

Furthermore, the GATT non-discrimination provision (Article III) does not contemplate a scenario where an intervention serves a legitimate objective. It contemplates any intervention, irrespective of its rationale. Actually, at the heart of Article III is a common understanding that protectionist policies, that is, plain vanilla policies aiming to favour the domestic producer, should not substitute for customs duties (which do the same). Indeed, even in the WTO-era, some of the disputes adjudicated under Article III concerned cases where tax differentials did not serve a legitimate objective (Japan-Alcoholic Beverages II is a very appropriate illustration).

Under the circumstances, it is only normal that panels established a test where interventions would be judged first by checking consumers’ reactions (market-likeness), and then, assuming government and consumers did not share preferences in this respect, by asking the question whether less favourable treatment had been afforded to imported goods as a result. This approach does not bode with the TBT-think for the reasons expressed in what now follows.

3.2.2 TBT-Think

In a world where NTBs essentially fragment markets, the content of domestic regulatory policies matter. Commitments to simply help avoid tariff concession erosion are of no help. Terms like harmonization, (mutual) recognition, a combination of the two, entered the vocabulary of the world trade lexicon. The TBT Agreement was the ‘vehicle’ that helped introduce this terminology into the GATT first, and the WTO later.²¹

Why the TBT? The drastic reduction of tariffs entailed that domestic lobbies had to look elsewhere for protection. The (almost) irrelevance of tariffs, as well as the enhanced disciplining of subsidies, entailed that regulatory subsidies emerged as the obvious candidates for those seeking protection. By adopting legislation that copied domestic production patterns, trading nations were shifting costs to foreign producers.²²

In principle, one could imagine that imaginative judges could have construed an elaborate non-discrimination test that could eventually have led to a TBT-like outcome. Still, some things would have been missing. The TBT integration discipline imposes a hierarchy: WTO member must use international standards, whenever appropriate. If such do not exist, then they must adopt domestic measures that are the least restrictive option on international trade and apply them in non-discriminatory manner. They should aim to express their domestic measures in terms of performance-rather than process requirements (and thus, invite gains from innovation from the trade of foreign goods that might be different but are equally effective when it comes to protecting the unilaterally set objective). Finally, they should also strive to recognize foreign regulations as equivalent to their own (underlining thus, the quest for gains from innovation).

²⁰ One might retort, why would trading nations shoot themselves on the foot? Political economy, as well as the presence of prospective remedies in the GATT/WTO legal order, might help explain why this can be the case. Grossman et al. (2013)

²¹ Brien (1981), and Winham (1986) discuss the negotiating record of the first TBT Agreement in separate papers. Sykes (1995) discusses the rationale for the second TBT Agreement, which is currently in force.

²² Bütte and Mattli (2011) mention some hilarious examples. Compare Sykes (1995).

Harmonize when necessary, recognize when possible, and anyway adopt necessary measures: this is the TBT recipe.²³ Necessity emerges as the linchpin of the whole edifice. International standards are presumed necessary. When no international standards exist (or when they cannot be appropriately used), WTO members should first think of the necessity to intervene, and if the response to this question is affirmative, they should adopt necessary measures (Article 2.2).

The TBT does not go as far as to impose common policies. It is not an instrument mandating positive integration. It is an instrument that seeks to minimize shifting costs to trading partners. This was a concern only with respect to border instruments (tariffs; import quotas) in the GATT-think. This is the concern for domestic instruments in the TBT world.

Non-discrimination thus, in the TBT world, concerns only the application of necessary measures (either international standards, or domestic technical regulations, standards, and/or conformity assessment).

4. Suggested Approach

With this in mind, here is our proposal regarding the understanding of the TBT. No legislative change is required. In fact, we believe that the approach advocated here is the only faithful to the current drafting of the TBT.

4.1 Sequence Matters

Assuming no international standard exists, or can be appropriately used, WTO panels (and the AB) should first ask whether a measure is necessary, in the sense, that it is the least restrictive option reasonably available to the regulating WTO member in order to reach its unilaterally defined objective:

- If the response is no, the WTO courts would not need to look any further. Violation of the necessity-requirement amounts to violation of the TBT;
- If the response is yes, then the next question should be whether the necessary measure has been applied in non-discriminatory manner. In this part of their analysis, WTO courts should inquire into whether the necessary measure is applied in even-handed manner across WTO members.

The key is that the discussion on non-discrimination, to which we return in what follows, should be sequential to the discussion on necessity. These are not two independent obligations, but one coherent whole. The Appellate Body would simply have to see the TBT as its current understanding of Article XX of GATT, that we have discussed above. There is only one twist: the burden of production of proof stays with the complainant, and not the defendant. The counterbalancing factor is, of course, that regulators must adopt not only non-discriminatory, but necessary measures.

One final point regarding the allocation of the burden of proof is probably warranted here. Practice lends support to the argument advanced by Horn and Mavroidis (2009), that the Appellate Body does not operate as umpire to a tennis game. Complainant does not need to pass the ball above the net before the burden shifts to the defendant. Instead, the Appellate Body asks both parties to submit evidence, and will judge on preponderance of evidence. Still, the first shot should originate in the complainant.

²³ Necessary means not the absolutely least restrictive measure, but the least restrictive measure available to the regulating WTO member. This is a case law contribution, that the Appellate Body introduced first in its report on US-Gambling. The rationale, which we applaud, is that, otherwise, regulating states might have found it impossible to pursue legitimate objectives, simply because pursuing them could be quite costly. The “hardship test” that the Appellate Body has introduced aims at allowing members to use the next in line necessary measure, which they can employ without incurring a disproportionate cost.

4.2 Non-discrimination

4.2.1 Policy Likeness

Market-likeness is irrelevant.²⁴ If at all, it is part of the rather ‘esoteric’ contemplation whether intervention is necessary (the first leg of the necessity-requirement as per Article 2.2 of TBT). If WTO members approved of the manner in which their own consumers were behaving in the market, and/or if they thought that intervention was not necessary, then they would not have enacted a technical regulation or a standard in the first place.²⁵ It is policy-likeness that we care about in TBT, since only measures pursuing a legitimate objective come under its aegis.

4.2.2 Less Favourable Treatment

The Appellate Body should simply ask whether the measure has been applied in even-handed manner. The exclusivity-requirement should be dropped because it is both unwarranted in light of the overall TBT-context as well as almost impossible to prove. There is no need to add a quixotic quest to serve the purpose of the TBT.

4.3 International Standards

The recommendation here is quite simple: assuming deviation from an international standard that could have been appropriately used, the burden of proof to justify the adopted measures should switch to the party possessing information to this effect, that is, the deviating party.

4.4 Standards

In US-Tuna II (Mexico), the Appellate Body reduced the scope for standards to redundancy. Indeed, if access to a standard is uninhibited, and anyone can profit from say a “dolphin-safe” label, irrespective whether it has complied or not with the statutory requirements, why adopt a standard in the first place.

Our recommendation is that standards should revert to the intended meaning. Standards, like technical regulations, are adopted whenever government- and private preferences differ. The wider category is market-likeness, and the narrower segment is the government delineation. The difference between the two instruments is confined to the regulatory intensity. When regulatory choices are important, technical regulations will be adopted. When issues of lesser interest are at stake, it is time for standards.

In our example, allowing for sales of “dolphin-unsafe” tuna is proof in and of itself that the “dolphin-safe” label is a standard. It would have been a technical regulation only if “dolphin-unsafe” tuna could never access the US market.

5. Conclusions

In this paper, we analysed the TBT as “completed” through case law. The outcome is as disappointing intellectually, as is unfaithful to the intentions of the framers. It is not the law that is wanting. It is case law that has failed. This is good news, since all that needs to be done is undo the current case law. It is easy to fix, since there is no need to enter a cumbersome renegotiation of the TBT across 164 heterogeneous partners at a point in time when, as the recent experience in Buenos Aires has amply

²⁴ Neven (2001), and Neven and Trachtman (2013) discuss this point in detail.

²⁵ Mavroidis (2013), and Mavroidis and Saggi (2014) have expressed similar thoughts on this score.

demonstrated, they find it difficult to agree even on the drafting of a simple Ministerial Decision with little substantive content.

The Appellate Body is of course, an agent, and not a principal as Article 3.2 of DSU²⁶ makes it abundantly clear. It must observe the policy space conferred to the WTO but its framers. Its interpretative discretion is what this term suggests: interpretative as opposed to law making function. To be sure, with the exception of its treatment of international standards, the Appellate Body has not affected the policy space committed. It has on the other hand, turned the test for consistency to its head.

Why has this been the case? Largely, because the Appellate Body has consistently failed to understand that the Vienna Convention on the Law of Treaties (VCLT), its instrument for interpreting not only the TBT but all of the WTO, is a rule and not rules of interpretation. Jimenez de Arrechaga (1978) put it so well, when he stated that the reason why Article 31.1 of VCLT was expressed in one sentence was precisely because the guidance to interpreters was to avoid favouring one of the elements mentioned over the other.

The Appellate Body has managed to do by and large the opposite. To use Trachtman's (2007) inimitable expression, Oxford English Dictionary has emerged as its most frequently used source of inspiration. In similar vein, Horn and Weiler (2004) when criticizing the attitude of the Appellate Body towards international standards in its EC-Sardines jurisprudence, mention (p. 140):

True to their belief in a textualist method of interpretation, out come the dictionaries! The Panel comes armed with Webster. The AB fields its favorite Oxford Shorter. And we let the learned wordsmiths whose dictionary definitions are the most extreme example of understanding language independently of context, and with no reference to object and purpose (i.e. the exact opposite approach to meaning of words which a legal interpreter of international texts should adopt), decide for the WTO the relationship between international standard setting and national administrative procedures.

Dictionaries however, per construction, privilege interpretations of the widest possible use. This is why they appeal to wide categories of users, and not to specialists who care about particular uses. Even the most sophisticated dictionaries will refer to a few, frequently encountered, contexts. Trade agreements are not high up on that list.

One might further ask, why bother since, dictionaries or not, the latest report seems to take care of the problem. The problem has not been solved, and this is a salient feature of this paper. The problem will not be solved unless if we can legitimately expect predictable outcomes in the future. This will be the case only when a coherent test for consistency has been developed. We quote from Haas (2016):

One factor ... increasing the odds that world order will survive is that it not require talented statesmen, the supply of which is likely to be insufficient ... Individuals of mediocre or poor skills will enter into positions of responsibility ...

Respecting proportions of course, since Haas refers to the world order and here we deal with the case law of the Appellate Body, his point is highly relevant to our discussion. The recent compliance panel did not change anything in terms of test of consistency with TBT, and still reached a different position. Most likely, a reasonable panellist managed to drive this point through. Similar results though, can only be guaranteed if the test itself becomes more rational. This is what we tried to propose in this paper.

The rationale for our approach is straightforward. The WTO Appellate Body should always ask one question first: why was an agreement signed? Having provided the response to this question, it will find it much easier to interpret the used terms in line with the objective function of the instrument. It

²⁶ DSU stands for Dispute Settlement Understanding, the WTO agreement that regulates dispute settlement before the WTO.

will thus, be respecting the mandate of the VCLT, as well as its institutional role, that of the agent. It did not do it in the context of the TBT so far. It is high time for change.

References

- Baldwin, Robert E. 1970. *Non-tariff Distortions in International Trade*, Brookings Institution: Washington, DC.
- Bhagwati, Jagdish, and V.K. Ramaswami. 1963. Domestic Distortions, Tariffs, and the Theory of Optimum Subsidy, *Journal of Political Economy*, 71: 44–50.
- Brien, Laurel A. 1981. The Tokyo Round Trade Agreements, Technical Barriers to Trade, vol. 4, The US Department of Commerce: Washington DC.
- Büthe Tim, and Walter Mattli. 2011. *The New Global Rulers*, Princeton University Press: Princeton, New Jersey.
- Coglianesi, Cary. 2009. The Transparency President? The Obama Administration and Open Government, *Governance*, 22: 529-544.
- Conconi, Paola, and Tania Voon. 2016. EC-Seal Products: The Tension Between Public Morals and International Trade Agreements, *World Trade Review*, 15: 211-234.
- Grossman, Gene M., Henrik Horn, and Petros C. Mavroidis. 2013. Domestic Instruments pp. 205-345 in Henrik Horn and Petros C. Mavroidis (eds.), *Legal and Economic Principles of World Trade Law*, ALI, The American Law Institute Reporters' Studies on WTO Law, Cambridge University Press: New York City.
- Haas, Richard N. 2016. *A World in Disarray*, Penguin Books: New York City, New York.
- Horn, Henrik, and Petros C. Mavroidis. 2009. Burden of Proof in Environmental Disputes in the WTO: Legal Aspects, *European Energy and Environmental Law Review*, 18: 112-140.
- Horn, Henrik, Petros C. Mavroidis, and Erik N. Wijkström. 2013. In the Shadow of the DSU: Addressing Specific Trade Concerns in the WTO SPS and TBT Committees, *The Journal of World Trade*, 47: 729-760.
- Horn, Henrik, and Joseph H.H. Weiler. 2004. European Communities—Measures Affecting Asbestos and Asbestos-Containing Products, *World Trade Review*, 3: 129-151.
- Howse, Robert, and Philip I. Levy. 2013. The TBT Panels: US-Cloves, US-Tuna, US-COOL, *The World Trade Review*, 12: 327-376.
- Irwin, Douglas A., Petros C. Mavroidis, and Alan O. Sykes. 2008. *The Genesis of the GATT*, (with), Cambridge University Press: Cambridge, Massachusetts.
- Jimenez de Arréchaga, Eduardo. 1978. *International Law in the Past Third of the Century*, *Receuil des Cours*, 159, Brill Academic Publishers: The Hague, Netherlands.
- Marceau, Gabrielle, and Joel Trachtman. 2002. The TBT, the SPS and the GATT: A Map of the WTO Law of Domestic Regulation of Goods, *Journal of World Trade*, 36: 811–82
- Mavroidis, Petros C. 2016. *The Regulation of International Trade*, MIT Press: Cambridge, Massachusetts.
- Mavroidis, Petros C. 2013. Driftin' Too Far from Shore, Why the Test for Compliance with the TBT Agreement Developed by the WTO Appellate Body is Wrong, and what Should the Appellate Body Have Done Instead, *The World Trade Review*, 12: 509-531.

- Mavroidis, Petros C., and Kamal Saggi. 2014. What is not so Cool about US-COOL Regulations? A Critical Analysis of the Appellate Body's Ruling on US-COOL, *World Trade Review*, 13: 1-22.
- Neven, Damien J. 2001. How should Protection be Measured under Art III. of the GATT, *European Journal of Political Economy*, 17-421-444.
- Neven, Damien J., and Joel Trachtman. 2013. Philippines –Taxes on Distilled Spirits. Like Products and Market Definition, *World Trade Review*, 12: 297-323.
- Sykes, Alan O. 1995. *Product Standards for Internationally Integrated Goods Markets*, Brookings Institution: Washington DC.
- Trachtman, Joel. 2007. The WTO Cathedral, *Stanford Journal of International Law*, 43: 127-171.
- Winham, Gilbert R. 1986. *International Trade and the Tokyo Round Negotiations*, Princeton University Press: Princeton, New Jersey.

Author contacts:

Petros C. Mavroidis

Columbia Law School and University of Neuchâtel

Edwin B. Parker Professor of Law at Columbia Law School

Jerome Greene Hall, Room 734

435 West 116th Street

New York, NY 10027

Email: pmavro@law.columbia.edu

Email: Petros.Mavroidis@unine.ch