Michigan Journal of International Law

Volume 18 | Issue 1

1996

Reformulated Gasoline Under Reformulated WTO Dispute Settlement Procedures: Pulling Pandora Out of a Chapeau?

Jeffrey Waincymer
Deakin University School of Law

COMMENTARY

REFORMULATED GASOLINE UNDER REFORMULATED WTO DISPUTE SETTLEMENT PROCEDURES: PULLING PANDORA OUT OF A CHAPEAU?

*Jeffrey Waincymer**

INTRODUCTION

A central question in any field of international law is whether supranational legal systems can promote harmonious intergovernmental relationships and provide suitable mechanisms for the peaceful and equitable settlement of disputes. In the field of international economic law, a key mechanism that is utilized in support of such goals is the dispute settlement system of the World Trade Organization (WTO).[1] The

* Professor of Law, Deakin University School of Law (Burwood, Australia). B.Comm, LL.B, Melbourne University (1976); LL.M, Monash University (Melbourne) (1981).

1. Final Act Embodying the Results of the Uruguay Round of Multilateral Trade Negotiations, Apr. 15, 1994, 33 I.L.M. 1140 (1994) [hereinafter Final Act] which includes the

WTO became operative on January 1, 1995. The WTO and its various Uruguay Round agreements have replaced the General Agreement on Tariffs and Trade (GATT),[2] but have carried on its basic dispute settlement methodology, albeit with significant modifications.

Formal dispute settlement within both the GATT and WTO systems has relied primarily on an internal panel process and ultimate consideration and adoption of the outcome of that process by the Members as a whole. During the 1970s and 1980s, a number of problems arose with the GATT dispute settlement procedures which in turn led to the negotiation of a separate Dispute Settlement Understanding (DSU)[3] as part of the Uruguay Round package. This article concentrates attention on one of the most important developments in that reform, namely the introduction of an appellate process. The primary motivation behind the creation of an appellate process was to ensure that there was a proper mechanism for reviewing the findings of panels. This was seen as particularly desirable once the Uruguay Round negotiators had decided on automatic adoption of panel reports. Previously, panel reports were adopted only by consensus; this had led to problems when losing parties at times blocked adoption of reports.

My analysis of the new appellate process is undertaken through an examination of the very first appeal against a panel decision.[4] The case concerned the United States' domestic regime regulating the pollution effects from gasoline usage.[5] Brazil and Venezuela were the complaining

---

Agreement Establishing the World Trade Organization, opened for signature Apr. 15 1994, 33 I.L.M. 1144 [hereinafter WTO Agreement]. The texts are also to be found in GATT Secretariat, THE RESULTS OF THE URUGUAY ROUND OF MULTILATERAL TRADE NEGOTIATIONS (1994).

2. General Agreement on Tariffs and Trade, Oct. 30, 1947, 55 U.N.T.S. 187 [hereinafter GATT 1947]. The WTO and its various Uruguay Round Agreements have replaced the GATT, but maintained its legal norms by incorporating the current version of GATT, designated as GATT 1994, as a constituent element of the WTO Agreement.

3. A comprehensive Dispute Settlement Understanding, entitled Understanding on Rules and Procedures Governing the Settlement of Disputes, was included as Annex 2 of the WTO Agreement, 33 I.L.M. 1144.

4. For a review of the dispute settlement processes of the GATT see JOHN H. JACKSON, THE WORLD TRADING SYSTEM: LAW AND POLICY OF INTERNATIONAL ECONOMIC RELATIONS (1989); ROBERT HUDEC, THE GATT LEGAL SYSTEM AND WORLD TRADE DIPLOMACY (2nd ed. 1990); ROBERT HUDEC, ENFORCING INTERNATIONAL TRADE LAW: THE EVOLUTION OF THE MODERN GATT LEGAL SYSTEM (1993); William Davey, *Dispute Settlement in GATT*, 11 FORDHAM INT'L L.J. 51 (1987); Ernst-Ulrich Petersmann, *The Dispute Settlement System of the WTO and the Evolution of the GATT Dispute Settlement System Since 1948*, 31 COMMON MKT. L. REV. 1157 (1994).

5. The Panel report and Appellate Body report were each entitled *United States—Standards for Reformulated and Conventional Gasoline*. The Panel report is contained in WTO Document WT/DS2/R (Jan. 29, 1996) [hereinafter Panel report]. The Appellate Body decision was WTO Document WT/DS2/AB/R (Apr. 29, 1996), adopted by the Dispute Settlement Body on May 19, 1996 [hereinafter Appellate Body report].

parties, while the European Community and Norway were interested third parties who also argued against the validity of the U.S. provisions. While the article builds its discussion around that case, it aims to provide an analysis of some of the wider issues that have and will arise with an appellate process. The predominant aim is to see how the Appellate Body approached its function and to identify and evaluate the further evolution of the WTO legal system through that process.

Part I of the article begins by outlining existing GATT/WTO provisions concerning trade-related environmental measures which were relevant to the *Reformulated Gasoline* case.[6] Part II then outlines the facts in the dispute and gives a brief introduction to the decisions at the Panel and Appellate Body stages. Part III deals with the present and potential implications for the appellate process in terms of the substance of the dispute, the methodology and procedure adopted, and the wider issues that the case brings to attention. This Part also addresses some of the theoretical and practical issues that affect the question of the role of law and legal adjudication within the WTO.

Questions to be considered include the substance of exemption provisions, general approaches to interpretation, the relevance of traditional international law adjudicatory techniques, the status to be given to past panel reports and negotiating history, the appropriate grounds for appeal, the scope of measures to be analyzed, the requisite burden and onus of proof, and the general approaches to legal reasoning, articulation of principles and adjudicatory methodology. Because the DSU says little about many of these issues, there were a number of choices open to the Appellate Body members. These choices would, in turn, be likely to have a significant effect on the outcome of the dispute and, in addition, on the dispute settlement system overall.[7]

Ambiguities and interpretational difficulties also arise when considering the core WTO trade principles and the way they allow exemptions for environmental measures. In any international agreement, it is common to find less than clear and comprehensive guidance on such questions, particularly one which was originally drafted when environmental issues

---

6. Part I is not intended as a comprehensive analysis of the entire interface between the WTO agreements and environmental measures. For such an analysis, see ERNST-ULRICH PETERSMANN, INTERNATIONAL AND EUROPEAN TRADE AND ENVIRONMENTAL LAW AFTER THE URUGUAY ROUND (Kluwer Law International 1995); Steve Charnovitz, *Improving Environmental and Trade Governance*, 7 INT'L ENVTL. AFF. 59 (1995); Jennifer Schultz, *Environmental Reform of the GATT/WTO International Trading System*, 18 WORLD COMPETITION 77 (Dec. 1994).

7. The *Reformulated Gasoline* case could also be examined in the context of the current debate about the proper interface between trade and environmental policy and the role and response of the WTO to that debate. That debate is not addressed in this article.

were not given such a political priority. It is also to be expected that negotiators wishing to maintain effective sovereignty in these broad areas of domestic regulation would frame the exempting provisions in broad language. As a result, there are a number of ambiguities or at least interpretational difficulties raised by these provisions. It is very easy for any seemingly narrow question of treaty interpretation to implicate major methodological and conceptual questions that have profound implications for the operation of the system as a whole. The type of questions that may flow from this particular dispute include the following: Should WTO panels and the Appellate Body apply a literal approach to interpretation? Should dictionary definitions be utilized? Is it appropriate to draw presumptions based on the use of different qualifiers in different sections? Should general principles of interpretation in international law apply? If not, why not? If dispute settlement in GATT/WTO is thought to need a more conservative and politically sensitive approach, how would this be best reflected in Appellate Body reasoning?

For anyone with even the vaguest notion of GATT's historically cautious approach to legalism, most such questions would be unwelcome additions to any formal negotiating agenda. But to shy away from such a debate does not obviate the need for a final determination of these issues. One thesis of this article is that, unlike substantive areas of negotiation, once an adjudicatory process is adopted, issues of adjudicatory methodology are inevitably determined either by the empowering authorities articulating the choices and seeking a consensus approach, or by leaving it to individual adjudicators to determine what that consensus would most likely be or to employ their own personal legal philosophy. These issues are dealt with in more detail in the context of the actual decision.

One caveat should be expressed at the outset, which in effect is a justification for the very decision to write such an article. To some supporters of the GATT/WTO, the dispute settlement process has evolved in an understated, conservative, and consequently healthy manner and has benefited from the absence of too fine a scrutiny or too critical a perspective. Thus, a detailed analysis of the first Appellate Body decision may, in the eyes of some, unduly expose the law-making functions of adjudicators, undermine some aspects of the logic of the decision and ultimately adversely affect the respect for the process and the system as a whole.

There are a number of reasons why this should be less of a concern today than ever before. Furthermore, there are reasons why it is now important for supporters of the system to outline the inevitable difficulties of a binding adjudicatory model and help foster an acceptance of these elements of any advanced and dynamic legal system. The first reason is

that the older, "quieter" style could not be expected to last, simply because of the higher profile of the WTO and its dispute settlement function. There are now a critical number of novel disputes in high-profile areas such as environmental protection. New Uruguay Round agreements in areas such as services, investment, and intellectual property, and the debate about possible new areas for coverage such as labor, competition, and corruption, have also raised consciousness of the WTO among a wider range of interest groups. The automatic adoption of reports and the introduction of the new process of appeals will add to this scrutiny.

All these factors ensure that WTO practice in general, and dispute processes in particular, will come under increasing scrutiny from supporters and critics alike. While it is inevitable that some concerns will be raised about interpretations adopted from time to time, respect for the system will depend in part on the respect for the methodology and logic adopted. Responsible and constructive evaluation will hopefully enhance both the practices of, and attitudes toward, the dispute settlement system as a whole.

A final justification for such a study is that as the Members consider amending existing agreements or introducing new areas into the body of WTO law, there are many cautionary and drafting lessons to be learned from an analysis of cases such as *Reformulated Gasoline*.

I. Environmental Exceptions to Basic Liberal Trade Principles

The first important point to make in the context of an analysis of the *Reformulated Gasoline* case is that the GATT/WTO system does not seek to prevent environmental protection provisions per se. Instead, the various agreements seek to limit distortionary and unreasonable measures which limit trade. Thus, an environmental measure that is nondiscriminatory as between countries and as between domestic and foreign producers is perfectly acceptable under the system. Furthermore, even when a measure does not satisfy these requirements, it may still be defined as acceptable under the exempting provisions which were considered in *Reformulated Gasoline*.

Article XX of GATT 1994 contains the key exemptions to these nondiscrimination principles. Article XX(b) allows an exemption for measures that are "necessary to protect human, animal, or plant life or health." Article XX(d) allows for measures that are "necessary to secure compliance with laws or regulations which are not inconsistent with the provisions of this Agreement . . . ." Article XX(g) exempts measures "relating to the conservation of exhaustible natural resources if such

measures are made effective in conjunction with restrictions on domestic production or consumption." In each case, measures which seek Article XX exemption must not be "applied in a manner which would constitute a means of arbitrary or unjustifiable discrimination between countries where the same conditions prevail, or a disguised restriction on international trade . . . ." This has been described as the "chapeau" of Article XX. The Appellate Body decision concentrated on Article XX(g) and the chapeau.[8]

Article XX shows that the GATT Contracting Parties were in agreement that the core GATT principles should not act as a blanket prohibition on general governmental policymaking in key areas such as health, security and the environment. On the other hand, there was also concern that these areas should not become an umbrella for blanket exemptions from GATT principles. The drafting and interpretational difficulties arise from competing policy objectives and the determination of appropriate principles to differentiate between acceptable and unacceptable measures.

II. THE *REFORMULATED GASOLINE* CASE

Reformulated Gasoline dealt with the validity of a U.S. regulation which aimed to control pollution from the combustion of gasoline that was either manufactured in, or imported into, the United States.[9] Stated simply, the case centered around the rules for establishing historical pollution baselines that would operate as allowable ceiling levels. It is necessary, however, to explain the regulations in more detail, as this article seeks to argue that an important implicit element of the case was the proper ambit of measures to consider under the GATT provisions.

The regulation in dispute was enacted under the Clean Air Act of 1990 (hereinafter CAA).[10] The CAA created two gasoline programs: one which aimed to generally contain pollution from gasoline production to levels not exceeding 1990 levels and the other, which aimed to specifically reduce pollution in designated major population centers.

---

8. The Agreement on Technical Barriers to Trade also can deal with environmental protection, as can the Agreement on Agriculture, the Agreement on the Application of Sanitary and Phytosanitary Measures, the Agreement on Trade-Related Aspects of Intellectual Property Rights and the Agreement on Subsidies and Countervailing Measures. While these are important elements of the WTO interface between trade and the environment, they are outside the scope of this article, which is only concerned with those provisions which were of concern to the Panel or Appellate Body.

9. Regulation of Fuels and Fuel Additives-Standards for Reformulated and Conventional Gasoline, 40 C.F.R. § 80-20.

10. Air Pollution Control Act (Clean Air Act) Amendments of 1990, 42 U.S.C. §§ 7401-7671q.

Where the latter program was concerned, only "reformulated" gasoline was allowed to be sold.

The Environmental Protection Agency (EPA) was the agency which administered these programs. It set the requirements for reformulated gasoline. These contained both compositional and performance specifications. Where conventional gasoline was concerned, all elements had to remain in as clean a state as at 1990 levels when assessed on an annual average basis. The Panel report indicates that the policy aim of the conventional gasoline rules was to stop the dumping of the substances that were not allowed to be contained in reformulated gasoline.

The EPA rules provided that from January 1, 1995 to January 1, 1998, domestic refiners, blenders, and importers could use a "Simple Model" to certify compliance with certain fixed specifications. It also required certain elements to be below 1990 levels on an average annual basis.[11] The "Gasoline Rule" provided for baselines to be established for these historical comparison purposes. These were to be determined either on an individual basis or on a statutory basis, as determined by the EPA, to reflect average U.S. gasoline quality in 1990. The rules provided that domestic refiners who were in operation for at least six months in 1990 must determine an individual baseline by one of three methods, all of which related to some aspects of their own historical data. The primary method involved the use of their own quality data and volume records for 1990 gasoline. Blenders and importers were required to use the same method if they had this data, but if not, the two fall-back methods were not available to them and they instead had to use the statutory baseline.

The contentious aspects of this regime were that importers were not allowed to use the actual individual baselines of the foreign refiners from whom they imported, and were also denied resort to the fall-back methods. An importer that was also a foreign refiner could, in theory, use individual baselines if at least 75% of the 1990 foreign-produced gasoline was imported into the United States in that year; in practice, virtually none could have complied with this threshold.

At the time of the formulation of the rules, claims of discrimination were raised. Venezuela had filed an earlier complaint through the GATT in 1994 but agreed to drop the case in return for a U.S. promise to amend the regulation.[12] In May 1994, the EPA had in fact recommended that the

---

11. A second "Complex Model," to be applied from January 1, 1998, was not at issue in the WTO dispute.

12. *See* Aubry D. Smith, *Executive-Branch Rulemaking and Dispute Settlement in the World Trade Organization: A Proposal to Increase Public Participation*, 94 MICH. L. REV. 1267, 1269–70 (1996). Smith argues that there is a need for greater domestic scrutiny of executive branch actions which seek to deal with international obligations and concerns. Steve

individual baselines of foreign refiners be used in a limited manner for reformulated gasoline only. Fully equal treatment for foreign refiners was never advocated, on the stated policy basis that there would be general problems with verifying data and particular avoidance problems given the difficulty of identifying and distinguishing between different batches of gasoline. A negotiated solution was organized between Venezuela, the EPA, the Energy Department, and the U.S. Trade Representative's office.[13] However, when Congress became aware of this deal, it blocked the proposed rule change by denying the funding necessary for its implementation.

In due course, the United States received new requests for consultations under the principles of the new WTO Agreement, namely, Article XXII:1 of GATT 1994, Article 14.1 of the Agreement on Technical Barriers to Trade (TBT Agreement), and Article 4 of the DSU. A panel (hereinafter "the Panel") was established and heard the parties, and its final decision was adopted by the Dispute Settlement Body (DSB) on January 29, 1996.[14] The United States filed a notice of appeal on February 21, 1996. Written submissions were filed and an oral hearing was then held. The Appellate Body asked further specific questions of the parties, some of which were answered orally and some of which were answered in writing. The parties were also invited to provide final written statements of their respective positions. The Appellate Body completed its report on April 29, 1996 and it was adopted by the DSB on May 19, 1996.

When the case was heard before the Panel, the complainants argued that the U.S. regulations were contrary to the Most Favoured Nation (MFN) requirement of Article I and the National Treatment requirement of Article III of GATT 1994. They argued further that the measures were in breach of Article 2 of the TBT and were not exempted by Article XX of GATT 1994. In the alternative, they argued that the rules nullified and

---

Charnovitz suggests that the settlement also involved an agreed quantitative restriction on imports from Venezuela which, if true, raises serious problems about settlements per se under the WTO dispute settlement system and raises the question of whether settlements should be fully notified and vetted for WTO legality. Steve Charnovitz, *Free Trade, Green Trade: Defogging the Debate*, 27 CORNELL INT'L L.J. 459, 522 (1994). The DSU now requires that settlement solutions shall be consistent with the Agreements and be reported to the Dispute Settlement Body [hereinafter DSB], which administers the DSU, and the relevant councils and committees. Marrakesh Agreement Establishing the World Trade Organization, Apr. 15, 1994, Annex 2, art. 3, para. 5–6, *The Results of the Uruguay Round of Multilateral Trade Negotiations: The Legal Texts* (1994).

13. Smith, *supra* note 12, at 1268.

14. Panel report, *supra* note 5.

impaired the benefits accruing to them under the Agreements.[15] The U.S. defended each of these claims and relied specifically on the exemption provisions of Article XX (b),(d), and (g).

The Panel's first major conclusion was that the aspects of the baseline rules dealing with imports were not consistent with Article III:4. In coming to that conclusion, the Panel held that imported and domestic gasoline were "like products" for the purposes of the General Agreement and as such, because imported gasoline was denied the same favorable rules as domestic gasoline, the rules were inconsistent with Article III:4. The Panel also considered that the measures could not be justified under Articles XX(b), (d), or (g). They were not "necessary" in the context of Articles XX(b) and (d) and were not, in the context of Article XX(g), "related to" conservation of exhaustible natural resources. The Panel followed earlier panel reports in considering that a measure must be "primarily aimed at" such a goal to satisfy the test.

The Appellate Body disagreed with some of the Panel's findings but still agreed that the measures were inconsistent with GATT 1994. In its view, the measures which offended against Article III:4 were protected by Article XX(g), as the rules were "related to" conservation, but were in violation of the chapeau of Article XX as they constituted "unjustifiable discrimination" and a "disguised restriction on international trade."

Detailed analysis of these decisions will be undertaken in Part III under a number of broad themes that have general implications for the dispute settlement system as a whole.

III. An Evaluation of the Appellate Process

A. *The Scope of the Appeal*

The first general issue raised by the respective approaches of the Panel and the Appellate Body is that of the proper approach of panels to broad ranging arguments and the resultant effect that various approaches would have on the scope of an appeal. The preliminary issues of coverage arise from the fact that the Panel did not rule on each question raised

---

15. Under Article XXIII:1(b) of GATT 1994 and Article 26 of the DSU, there is a ground for dispute resolution that involves non-violation complaints. Previous GATT approaches to such claims have utilized a *reasonable expectation* test to determine if an otherwise legal governmental measure still can ground a complaint within the GATT system. While this may at first sight appear novel, it is related to the international law principle of *rebus sic stantibus*. Some commentators have, however, called into question whether GATT Contracting Parties ever intended this category of disputes to be empowered. *See* Pierre Pescatore, *The GATT Dispute Settlement Mechanism—Its Present Situation and Its Prospects*, 27 J. World Trade 5 (Feb. 1993).

before it, and the decision of the U.S. to appeal only some of the Panel's findings. Because the Panel's analysis of Articles III:4 and Articles XX(b), (d), and (g) led to a conclusion that the measures should be modified, it declined to deal with other legal issues. These outstanding issues included whether the measures were valid under Articles I and III:1; whether the measures met certain other requirements of Article XX, including the chapeau; whether the measures were inconsistent with the TBT Agreement; and whether they constituted non-violation, nullification, or impairment under Article XXIII.

The decision by the Panel adjudicators to limit their findings to the minimum level necessary to produce a result is a common approach in domestic courts and is at first sight understandable in an international organization that has built up a dispute settlement function cautiously. However, this approach led to difficulties for the complainants at the appellate level, and may portend practical and theoretical problems in the context of WTO dispute settlement.

The Appellate Body had to deal with the preliminary question of whether the complainants could reopen debate on anything other than the narrow grounds of appeal lodged by the United States. The United States had limited its appeal to the interpretation by the Panel of Article XX and in particular, Article XX(g). The complainants had sought to defend the Panel findings on these provisions, and to argue in the alternative that if the Appellate Body upheld the U.S. arguments under Article XX(g), then the other aspects of Article XX and the TBT Agreement should be considered. The U.S. opposed this request on the basis that the complainants themselves had not appealed against the Panel decision. The complainants nevertheless argued that it would be within the scope of the Appellate Body's authority to address these other aspects of the decision. The United States, in turn, argued that this approach would be unfair to both the United States in the instant case, and to third parties who would from time to time wish to rely on this limiting principle when deciding whether to join as participants in any dispute. It was also suggested that widening the ambit of the appeal would further encourage a disregard for the Working Procedures which had been adopted in respect to the appellate process,[16] and which specified time limits and manner of appeals.

The Appellate Body effectively did not even need to rule on this issue as it took the complainants' request to be conditional on the Appellate Body not finding in their favor under Article XX, but it made

16. Article 17.9 of the DSU provided that working procedures were to be drawn up by the Appellate Body in consultation with the Chairman of the DSB and the Director-General of the WTO. These working procedures were drawn up and are found in Document WT/AB/WP/1 (Feb. 15, 1996) [hereinafter Working Procedures].

a ruling in any event which provides clear guidance to WTO Members. The Appellate Body upheld the U.S. contentions, and concluded that the complainants' request was not contemplated by the Working Procedures. It was not prepared to depart from those procedures in the absence of some compelling reason, such as "fundamental fairness or *force majeure*."[17]

This Appellate Body finding is rendered more interesting by its ultimate decision to find in favor of the complainants on grounds quite distinct from the Panel and on a matter not determined by the Panel—namely, the effect of the chapeau of Article XX. Given that the United States had only appealed against the Panel's actual adverse findings, it might not immediately seem clear how the Appellate Body felt able to make a determination on the chapeau of Article XX, while at the same time denying the claimants' right to address the further grounds that they wished to address. The reason for this apparent contradiction can be traced to the actual grounds for appeal—namely, the Panel finding that the measures were contrary to Article XX(g). The United States had appealed against this holding, which was the basis for the Panel's decision not to consider other arguments. Because the Appellate Body considered that the Panel erred on this question, the decision not to address alternative grounds was also an error of law. The U.S. appeal specifically addressed the other aspects of Article XX but did not refer to the TBT Agreement. Thus, in its own decision, the Appellate Body felt able to consider the Article XX chapeau but not the latter agreement.

This aspect of the decision points to some of the inevitable and complex problems inherent in any appeal process, which were not resolved by the negotiators in the DSU or in any direct manner in the Working Procedures. First, there is the question of whether either the Panel or the Appellate Body should limit its conclusions to only some of the grounds raised by the parties, particularly in cases where this would lead to a clear result. The answer to this question depends upon views about the primary aim of the dispute settlement system. The principal argument in favor of such a practice would be that the primary goal of the panel process is to resolve the dispute between the actual parties and not develop wide-ranging legal rulings. Moreover, this limited review is a common approach at first instance in many domestic adjudicatory systems. The primary argument against this approach would be that it could prolong disputes and lead to uncertainty as to the required measures to be taken by the losing party. For example, in the *Reformulated Gasoline* case, it

17. Appellate Body report, *supra* note 5, at 12.

is conceivable that the United States might amend its rules to account for the Appellate Body's concerns about Article XX, but that Brazil and Venezuela would still maintain that the modified rules remain inconsistent with WTO agreements for other reasons. An additional argument in favor of broad review might be that it is important to get nonbinding guidance and rulings from the Appellate Body on the interrelationship between provisions such as Article XX and the TBT Agreement.

Another concern about the scope of appeal raised by the *Reformulated Gasoline* case relates to the potential for wide-ranging and lengthy appeals. For example, given the refusal by the Appellate Body to allow the complainants to address new issues, parties may now be encouraged to draft appeal notices and cross appeals in the widest possible way to avoid losing their rights. This could in turn encourage more lengthy and costly hearings. On the other hand, different problems could have resulted if the Appellate Body took the contrary approach, as appellants could feel ambushed by these additional, arguments which would also tend to increase the length of cases.

This last set of concerns form one aspect of the wider question of how, if at all, to narrow the ambit of appeals to the most appropriate extent. Policy issues raised with respect to domestic courts are that appeals as of right are an important way to promote efficient decisionmaking, but that at the highest level, where legal expertise is also expected to be at the highest, it is appropriate for judges to exercise control and sift out those arguments that would ultimately be a waste of time. Domestic judges are also accepted as having rights to control the duration and coverage of cases in order to promote efficiency and limit the transaction costs of litigation.

The DSU contains no rule requiring parties to seek leave to appeal on any particular issue or stipulating that appeals are only allowed on matters of broader importance, which are the main streamlining procedures employed in many domestic jurisdictions. Such procedures are not advocated for the WTO, as Members would be expected to feel aggrieved if they were denied a right to vent an issue that they considered crucial. Furthermore, tests that try to limit appeals to some standard of general importance could not be employed easily and without contention.

There may, of course, be some self-regulated restraint by Members; if not, there is no obvious means of avoiding a more widespread approach to appeals. If there is also a growth in disputes themselves and a consistent practice of appeals in most cases, then the practical impact on resources could be significant. It is likely that many cases which cannot settle will be appealed as there are few strategic barriers to appeal. Moreover, in contrast to the domestic level, there are at least as many

strategic incentives: no costs are awarded through the WTO; legal advice (for the major players at least) is available on an opportunity cost rather than actual cost basis, owing to the use of existing bureaucratic resources; and finally, it would often be politically difficult for a government to accept an adverse panel finding without attempting to overturn it at the appellate level, particularly where the WTO outcome will be used to explain to domestic vested interests why there is a need to modify legislation.

B. *Remittance of Decisions to Panels*

A related practical issue is when and on what basis the Appellate Body should address a new ground, and when it should instead remit the matter to be decided at first instance by a panel. Regardless of general policy considerations, there is a further problem in the WTO context as neither the DSU nor the Appellate Body's Working Procedures expressly refer to a remittance power. One could hold it to be an inherent element of Appellate Body power or instead, merely allow a de facto and inadequate remittance power to be exercised by the Appellate Body by holding a matter to be a question of fact and requiring the parties to commence the dispute settlement process afresh. Because the time limits for panel proceedings are now relatively short, the end result would not be too different.

If the procedural hurdle were overcome, and a direct remittance power acknowledged, the next question would be when is it appropriate to use such a power? Domestic appellate courts are invariably given broad remittance discretion. A common factor used by domestic courts in determining when to exercise that discretion is whether there are factual questions involved or whether the dispute involves a pure question of law. In the latter situation, domestic appellate courts often deal with these issues without remittance in order to save further delay and expense.

Applying this test to the *Reformulated Gasoline* case, there is a tenable argument either way as to whether an analysis of the Article XX chapeau involved sufficient factual questions to suggest that remittance to the Panel would have been preferable, if such a power existed. An argument in favor of remittance is that the ultimate decision involved an identification of competing policy instruments and at least a partial analysis of their equity and efficiency, to determine whether the measures constituted "unjustifiable discrimination" and a "disguised restriction on international trade." On the other hand, the decision not to remit could be justified on the basis that in the view of the Appellate Body, the uncontroverted evidence showed that the baseline provisions could not be defended under the chapeau.

Such policy considerations cannot be looked at in isolation from the general factfinding and evidentiary approach that the Appellate Body will take from time to time as compared to a panel. If the approaches to factfinding are largely similar, there is a stronger case for not remitting. And if an appeal will truly turn on an unresolved question of fact, it would not appear to be a valid appeal in the first place, as Article 17 of the DSU expressly limits appeals to questions of law. While appellate courts throughout the world tend to broaden this to include mixed questions of fact and law, they do not take over the essential factfinding mission from the courts of first instance. There is no easy answer in the WTO context as factfinding itself, and the many methodological and evidentiary issues that concern domestic courts, have not been addressed to any significant degree. These issues will come to the fore whenever there is a dispute about highly contentious factual questions, which is more likely under some of the newer commitments under the Uruguay Round agreements, such as in the areas of agriculture and intellectual property.[18] It would be desirable for WTO policymakers and the Appellate Body to give some thought to this problem before a contentious dispute arises, as the lack of resources and agreed principles would too easily leave panels and the Appellate Body open to undesirable criticism.

C. *The Article XX(g) Exemption and its Relationship Test*

After dealing with the preliminary issue of coverage, the Appellate Body turned to the issue of justification under the Article XX(g) exemption of GATT 1994. The first central substantive issue addressed in any detail by the Appellate Body was the requirement that the measures were "relating to the conservation of exhaustible natural resources."

Whenever legislation contains broad connecting phrases without further elaboration of meaning and methodology, it is commonly left to adjudicators to determine key questions such as how strong the linkage must be, what relevance—if any—*purpose* will have in determining the linkage, whether purpose is to be determined objectively or subjectively, and what degree of purpose is required (such as sole, dominant, significant, more than minimal, or any purpose). The scope of the measures to be examined under such tests must also be determined. These issues were only touched on indirectly and to varying degrees by the Panel and the

18. For example, disputes might arise as to the true nature and amount of agricultural support, or the effectiveness of intellectual property protection.

Appellate Body. They are addressed in this section and the following one, dealing with the relationship test and scope of measures, respectively.

While the Panel held that clean air was an exhaustible natural resource and that a measure aimed at reducing the depletion of clean air would fit within this description, it also held that the discriminatory aspects of the baseline establishment rules were not related to conservation of an exhaustible natural resource. In dealing with the linkage test of "relating to," the Panel had quoted with approval from an earlier Panel report in the 1987 *Herring and Salmon* case.[19] In that case, the Panel considered that to satisfy the expression "related to," a measure did not need to be necessary or essential but did need to be "primarily aimed" at the stated purpose.

Utilizing this test in the *Reformulated Gasoline* case, the Panel could not see how the importer baseline rules themselves, which provided less favorable treatment for imports, could be primarily aimed at the conservation of natural resources. The Panel saw no "direct connection" between these rules and the permitted object. The Panel also noted that if the United States had used a different regime which provided national treatment consistent with Article III:4, this would not have hindered the conservation objective. In the Panel's view, this showed a further lack of connection between the offending provisions and what it saw as the required relationship.[20]

The Appellate Body rightly criticized the Panel for using the phrase "no direct connection" without indicating whether it intended this to merely be interchangeable with the phrase "primarily aimed at," or whether it was positing an additional test. By asserting that the lack of a direct connection "accordingly" meant that the "primarily" test was not satisfied, the Panel had presumably seen the phrases as identical, but it nevertheless failed to give any reason why they should be seen as such. This equation could be justified on the basis that a *primary aim* test is a *purposive* test and that provisions which are not directly connected to the permitted purpose do not satisfy such a test. The contrary argument is that an indirect connection can also be intended by a legislature or treaty negotiators. At the very least, there is an ambiguity in the Panel's concept of "directness" that justifies its rejection by the Appellate Body.

The Appellate Body then turned its attention to the appropriate meaning of the expression. It ultimately chose to tread a delicate line

---

19. Canada—Measures Affecting Exports of Unprocessed Herring and Salmon, Mar. 22, 1988, GATT B.I.S.D. (35th Supp.) at 98 (1988); 35S/98, *cited in* Panel report, *supra* note 5, at para. 6.39.

20. Panel report, *supra* note 5, at para. 6.40.

between cautiously following the parties' agreed views on these various legal tests and providing its own critical perspective. At a later stage in its report,[21] the Appellate Body noted that all participants in the dispute accepted the view from the 1987 *Herring and Salmon* report that there is an inherent requirement that a measure be "primarily aimed at" conservation to fall within Article XX(g). Because of this consensus, the Appellate Body felt it did not need to examine the point further, but could not resist noting that this phrase "was not itself treaty language and was not designed as a simple litmus test for inclusion or exclusion from Article XX(g)."

There is certainly a danger if adjudicators seek to paraphrase provisions and develop shorthand tests of their applicability. On the other hand, it is inherent in any connecting test that some elaboration of meaning and articulation of relevant factors is required. The Panel in the *Herring and Salmon* case had in fact sought to explain why it needed to further define the test. It formed the view that Article XX(g) did not indicate what the required relationship was. It queried whether *any* relationship would suffice or whether a *particular* relationship was required.[22] The Panel acknowledged the use of a lesser connecting test than "necessary" in Article XX(g), but felt that in light of the context and purpose of the provisions, that while a measure need not be essential, it should be primarily aimed at the stated objective. The *Herring and Salmon* panel felt that the purpose of Article XX was not to widen the scope for trade policy measures but rather to allow for the pursuit of policies aimed at the conservation of exhaustible natural resources.

That panel did not articulate further reasons as to why the "primarily" test best ensured this result, but a strong case can be made to support some form of purposive approach at least. It is unlikely that the drafters intended Article XX to protect measures that *accidentally* related to a specified goal. Thus some purposive or intent-based interpretation of the connecting test should be preferred. Whether that linkage should be as strong as a "primarily" test would require is a separate question. It is desirable to exclude measures that are not really aimed at the particular object, but this could also be accomplished by a lesser standard and is dealt with in part by the chapeau's reference to "disguised" measures. It is also conceivable that a particular provision could have two significant aims, with the non-conservation one being slightly superior. This would fall foul of a "primarily" test. Such a result is more open to challenge

21. Appellate Body report, *supra* note 5, at 18–19.

22. Canada—Measures Affecting Exports of Unprocessed Herring and Salmon, Mar. 22, 1988, GATT B.I.S.D. (35th Supp.) at 113 (1989).

from a policy perspective and the practical application of such a test would also be problematical.

The Appellate Body considered that basic principles of treaty interpretation demand that the "relating to" test imply a different outcome than a "necessary test," simply because the drafters of Article XX used different connecting phrases in different sub-articles. General questions about interpretative methodology are discussed in a later section, but at this stage, attention is given to the way meaning is determined.

The argument that the concept of relationship is different from that of necessity only takes us a very short way towards determining its meaning. The expression "related to," like all legal linkage and causal tests, has its unanswered dimensions. How close must the linkage be? What if a measure has a number of objects, only one of which fits within the GATT terminology? What if a measure technically fits the language but has an ulterior purpose? What if it is indirectly connected to the stated aim? Are questions of equity, efficiency or GATT legality wholly irrelevant? If not, to what extent can they be relevant without the test becoming a synonym for a necessity test?

The Appellate Body did not seek to present answers to these questions. As indicated above, the Appellate Body went on to utilize the paraphrase test of "primarily aimed at," showing a preference at this stage for consensual dispute resolution over judicial creativity. It held that the baseline rules must be looked at in their context to have any meaning. Because they were aimed at permitting scrutiny and monitoring of compliance, they satisfied that test. The baseline rules could not be regarded as being merely incidental to or inadvertently aimed at the conservation of clean air. It effectively used a factual finding of necessity to prove that the lesser relationship test was satisfied, when it held that the baseline establishment rules satisfied Article XX(g) because "without baselines of some kind . . . scrutiny would not be possible . . . and the Gasoline Rule's objective would be substantially frustrated."[23]

Certainly there is little linguistic alternative to seeing the use of "relating to" as opposed to "necessary" as implying a lesser test of justification, even though it is hard to see why exhaustible natural resources should have attracted a lower threshold than human life or health as referred to in Article XX(b). It is respectfully suggested that a "significant aim" test would be preferable from a policy perspective and would not offend acceptable principles of treaty interpretation. It would be justifiable on the basis that the drafters could be presumed to intend

23. Appellate Body report, *supra* note 5, at 19.

to exempt only those provisions that truly were aimed at conservation and that a strict approach to exempting provisions would empower a *de minimus* standard, if not one as intrusive as a "primary aim" test.

The Panel was also criticized for testing what it described as "the less favorable" treatment, rather than the measures themselves, against the terminology of Article XX(g). The Appellate Body saw this as a type of bootstrap argument that would turn the provisions upside down. Certainly, if the Panel's argument was that a conclusion that the regime is discriminatory automatically leads to the result that a requisite conservation purpose cannot be established, then this would be wholly unjustifiable logic. While the Panel was not clear in this aspect of its reasoning, this does not seem to have been its intent. Instead, as indicated above, it indicated that the desired conservation level could have been obtained without discriminatory treatment, and for that reason, the measures could not be said to be primarily aimed at that result.[24]

If one sought to criticize the Panel, one could argue that its comments were not far away from the language of a necessity or at least a reasonableness test, neither of which is expressly included in Article XX(g) and which would not normally be imported into the notion of "related to." It is even tenable to argue from a linguistic perspective, that a sub-optimal or even highly inefficient, inequitable, and misguided measure that truly aims to be an element of a conservation regime, is "related to" the goal of conservation. If one sought to defend the Panel's reasoning, this linguistic argument could be challenged on the basis that Article XX(g) requires a relationship to the *goal* and not simply to a broader statutory regime which itself has that goal. On this basis, one could interpret the Panel's holding as saying that the domestic/import distinctions of the baseline measures had no discernible purpose other than to provide less favorable treatment to importers, which in turn supports the conclusion that they could not be "primarily aimed at" conservation even though they were contained in a broad conservation regime. The difficulty with such a conclusion is that the measures could easily be explained as an attempt at promoting conservation in a less costly way for the administering agency. While both adjudicatory bodies saw this as inefficient and unreasonable, it can still be argued to be a discernible purpose "related to" the conservation goal.

D. *Scope of the Measures to be Examined*

Both the Panel's approach to Article XX(g) and the Appellate Body's conclusion raise a further crucial issue in the application of this and any

24. Panel report, *supra* note 5, at para. 6.40.

other relationship test. Once a determination is made as to the meaning of the test, the next issue is to determine the ambit of facts to be considered. This is probably the most crucial undervalued issue in any regulatory system that uses a causal, purposive or essential character test to distinguish between acceptable and unacceptable practices. This pervades many areas of governmental regulation but is often overlooked by drafters, decisionmakers and analysts.

Where the GATT provisions considered in the *Reformulated Gasoline* case are concerned, the question is whether to apply the relationship test to the entire gasoline rules, to the entire baseline establishment rules, or to some lesser element of those rules, such as those affecting importers alone. No express guidance is provided in the provisions themselves. In international agreements, it is common for vague phraseology to be employed, either because there is no clear consensus on a particular meaning or because the issue was simply never thought of or thought not to be important. That is understandable from a traditional sovereignty and political perspective; however, the irony is that with the advent of automaticity in the dispute settlement process, such vagueness merely empowers adjudicators to make significant contributions to the development of legal norms. While those contributions are theoretically not precedents, it is unlikely that an Appellate Body that is asked to consult all its members with each decision[25] will readily depart from a recently developed line of reasoning.

While both the Panel and the Appellate Body addressed the issue of the breadth of measures selected, neither expressly did so in terms of the potential effect on the application of the connecting test. The Panel had sought to make it clear that in its view, the connecting tests in Articles XX(b), (d), and (g) must only be applied to the measures found to be inconsistent with Article III:4.[26] In the context of Article XX(g), it indicated that the U.S. must, inter alia, show "that the measures for which the exception was being invoked—that is the particular trade measures inconsistent with the General Agreement—were *related to* the conservation of exhaustible natural resources . . . ."[27] The initial issue addressed by the Appellate Body was the meaning of "measures." It noted the view that the question to be determined was whether this term required consideration of the entire Gasoline Rule or instead allowed attention to be limited to the baseline establishment rules. However, the Appellate Body did not see this as a live issue given its view that the Panel held only the

25. *Working Procedures*, *supra* note 16, Part I, para. 4.

26. Panel report, *supra* note 5, at para. 6.22–6.24.

27. Panel report, *supra* note 5, at para. 6.35.

baseline rules to be inconsistent. The Appellate Body agreed that the "measures" to be considered in the context of Article XX were the same measures found to be inconsistent with Article III:4. It presumably thought this disposed of the problem and allowed for it to follow the same approach as the Panel.

Certainly, an exempting provision only makes sense in respect to measures which would otherwise fall afoul of legal requirements. But this does not indicate how to determine the ambit of examination in each case. In the context of an Article III review, are the offending measures merely the discriminatory rules against foreign goods or do they encompass the wider regime that includes the more favorable domestic measures? According to one view, it is only the discriminatory foreign measures that need to be excised. On the other hand, the domestic measures could be made harsher in order to restore nondiscriminatory parity. If both are to be included, is it only in respect to the discriminatory aspects or is the broader regime also encompassed? If so, how much of that regime and on what basis is the determination to be made?

The drafter of any rules will either expressly refer to this question, or the adjudicator will be forced to determine the ambit. This can be done on a technical or purposive basis. A technical basis could merely involve the adjudicator looking at the exact measures as identified by the parties to the dispute. If the parties describe them inappropriately, it could be argued to be their own fault if they lose as a result. The Panel report indicates that the complainants requested the Panel to find that the Gasoline Rule as a whole was contrary to Article III of GATT 1994, was not covered by Article XX, and was also contrary to Article 2 of the TBT Agreement.[28] If this approach was preferred, it could be used to assert that the Panel and the Appellate Body should have looked at a wider range of rules.

A limited technical approach of relying on the parties' own choice might protect adjudicators from accusations of judicial activism but would prevent a just outcome in many instances, and would be likely to adversely affect countries with less experience and resources. It would also encourage broad appeal documents alleging a whole array of offending measures as alternatives to each other, while providing nothing in the way of guidance to other participants in the system. For these reasons, it should be rejected.

An alternative technical approach would be to look at the question from a jurisdictional point of view. Here the Appellate Body could

28. Panel report, *supra* note 5, at para. 3.1.

determine that it is bound by the measures considered at the Panel level, given that its jurisdiction is limited to questions of law, unless it determined that the Panel's choice of ambit of measures was itself an error of law. Even if this approach were adopted, it is by no means clear what measures should have been considered by the Appellate Body in *Reformulated Gasoline* case because of different expressions used by the Panel from time to time in its report, although the better view is that the Panel intended to only examine the narrow discriminatory aspects of the baseline rules. On a number of occasions the Panel had explained that the reason why the Gasoline Rule was contrary to Article III:4 was that under the "baseline establishment methods" the imported gasoline was treated less favorably than domestic gasoline. When applying Article XX, the Panel examined "whether the aspect of the baseline establishment methods found inconsistent with Article III:4 could . . . be justified" under the exempting provisions. Thus, under a technical jurisdictional approach, it would be arguable that attention of the Appellate Body should have been focused on the baseline rules alone or the discriminatory elements of them.

This approach would mean that the highest adjudicatory body in the WTO has a more restricted power of determining the application of provisions than the panels themselves. Furthermore, as indicated above, it would often be possible to argue that a panel's choice of measures was not that which was contemplated by the relevant provisions and hence constituted an error of law. For these reasons, a purposive approach that properly integrates this issue with the general question of the nature and scope of the linkage test is to be preferred. Such a purposive approach would look for what is implied by the object and logic of the legislative scheme. As indicated by the Appellate Body, the general logic of an exempting provision is in respect to that part that would otherwise offend against positive obligations. An overly broad description of measures could be attacked on this principle alone.

From a policy perspective, to require consideration of the entire legal regime would also raise a number of further interpretational difficulties. For example, if part of a legislative regime came within the exemption but part did not, would the latter provisions be protected? If so, quite significant discriminatory measures could be employed as long as they were given a wider protective "blanket." To overcome this problem, could an adjudicator adopt a purposive test on the entire regime to deal with such hybrid provisions? Would any such purposive test look at the sole purpose, dominant purpose, or any significant purpose in order to categorize the entire regime? A sole purpose test would be meaningless when applied to a hybrid provision. Even a dominant purpose test would be

easily overcome by ensuring the legal protective "blanket" is of a sufficient size. Lesser standards would more easily be abused.

Another problem with an approach that looks at the entire regime is that it would further politicize the dispute settlement process, as critics could easily rely on rhetoric that accuses adjudicators of interfering with broad and popular areas of sovereign rights. It is logically and politically preferable to concentrate on those technical aspects of the regime that are found to offend against the positive commitments. Supporters of the WTO would prefer to read a headline saying the "WTO protects Venezuelan refiners," rather than one saying "WTO stops the U.S. from minimizing pollution!"

Criticizing too broad a field of analysis only goes part of the way toward a policy-based solution to the determination of the appropriate measures. Because some GATT exemptions use a "necessary," "justifiable," or "essential" test, either expressly or implicitly, there is also a danger of going too far in the other direction. If the measures to be examined are defined too narrowly, an intellectual evaluation may become impossible, simply because the narrow measures have no meaning in isolation. Many subsidiary parts, when looked at in isolation, will have little discernible purpose other than to support the general regime. Certainly the decisionmaker would be able to look at the context of the measures, but if the ambit was too narrow, it would still be impossible to say that the particular provisions have any discernible object. A second issue is that if such narrow provisions were to be analyzed under Article XX(d), there would then need to be some subsidiary test of whether they were truly ancillary to other provisions that themselves displayed the requisite object.[29] A more difficult and serious problem is that if measures are addressed too narrowly under such tests, given the difficulties of drafting perfect legislation, too many subsidiary parts of broadly acceptable regimes could become subject to pedantic attack.

As with most difficult questions there is a need to strike some balance between the overriding policy aims of the WTO system and the provisions in dispute in a particular case. An overly broad approach would allow widespread abuse and should be rejected. An overly narrow approach can be prevented by establishing a general principle that combines the need for minimum discrete content and a limitation to otherwise offending measures. Under this approach, the measures to be examined must be those that are identified as being in breach, and

29. Provisions such as these are commonly examined by domestic constitutional courts when considering a challenge to a provision's constitutional validity. These are often described as ancillary powers.

identified to as fine a degree of precision as possible that will still allow them to have some discrete object. If there is no discernible object, either inherently or contextually, then the choice is too narrow. If there are non-offending measures, it is too broad.

Determining the proper ambit of relevant questions will not circumscribe how those facts are determined and weighed by different adjudicators. Variations in levels of generality when describing the facts, or variations in shorthand descriptions of them, will raise or lessen the chance of the treaty provisions applying. As indicated above, the Panel in *Reformulated Gasoline* was criticized for concentrating on the discriminatory aspects of the provisions and, in the view of the Appellate Body, for testing its description of the *effect* of the measures, rather than the measures themselves, against the language of Article XX.

The Appellate Body's own approach also shows some inevitable choices and ambiguities. By looking at the baseline rules as a single entity, it was able to conclude with justification that these were aimed at being integral elements of a conservation regime. The first question to be asked is whether the Appellate Body should have looked this broadly, on either jurisdictional or policy grounds. On the test posited above, because reference was made to the entire baseline rules, which contained non-offending and important elements of the conservation regime, this question would be answered in the negative. The next issue is what the result might have been if the Appellate Body was directed to examine the baseline rule for importers only, but to read it in the context of the distinct rule for domestic refiners. In that context, was the rule really aimed at conservation or at discrimination? Alternatively, if the domestic refiner rules had to be included because their more favorable treatment is also discriminatory, what would have been the result if the measures to be examined were described as being limited to the different means employed for the two categories?

While the case against the measures becomes stronger if this is required, it is not automatic. The difference in treatment might still be justified, for example, on the grounds of reliability of data. In addition, the fact that domestic refiners of less than six months' standing also had to employ a statutory baseline would help to minimize the inherent discriminatory nature, or at least any demonstrable and inherent anti-foreign bias of the measures (although this could be explained away if the newer domestics would not have difficulty meeting the minimum standing requirement). But whatever the ultimate conclusion, if this approach was adopted, the case would have had to be made and considered. It would also have required the Appellate Body to be more explicit in terms of its ultimate view about the "primarily aimed at" test as the means of

satisfying the phrase "related to." In turn this would have forced some further analysis of whether the relationship test merely requires some objective connection, or whether it also requires some likely positive effect on conservation to come within Article XX(g). By looking at the broad baseline rules, there was no need to express an opinion on this question as those rules easily satisfied either requirement.

To reiterate the central thesis of this section, these problems arise in any legislative regime that prescribes or proscribes certain purposes, particularly in relation to complex commercial situations. It is difficult to define the parameters of facts to be analyzed, which in turn places the decisionmaker in a difficult and powerful position.[30] To describe the measures in an innocuous fashion as baseline establishment rules makes them clearly related to a regime that requires a baseline. To describe them as the discriminatory aspects, as the Panel did, points to the possibility of holding that they had other objects. While it is not possible or desirable to fully constrict decisionmakers in terms of the way they approach facts, key legal tests should provide more guidance than the adjudicators in the *Reformulated Gasoline* case enjoyed. At the very least, this should be a lesson for drafters of multilateral agreements with broad exempting provisions and for those within the WTO who are considering some further agreement on trade and environment as part of the ongoing working agenda of the organization.

E. *Treaty Interpretation and the Role of International Law*

A further interesting aspect of the *Reformulated Gasoline* decision relates to the process of treaty interpretation, as this obviously has implications way beyond environmental disputes under Article XX(g). As indicated above, the Appellate Body felt that the Panel had wrongly applied a test of "necessity" to Article XX(g).[31] In doing so, the Panel had,

---

30. For example, domestic antitrust rules and anti-avoidance tax provisions give rise to these adjudicatory difficulties.

31. Appellate Body report, *supra* note 5, at 16. It is conceivable that the Panel was merely utilizing the same finding of fact to support both a finding of lack of necessity under Article XX(b) and the lack of an *effective* relationship under Article XX(g). Nonetheless, this could be criticized under the view that the relationship test does look at effectiveness in advancing the policy goal. Regardless of the interpretational question, it is even possible to argue that the Appellate Body's own approach showed regard for considerations of necessity, not as a synonym for "related to," but as a means of showing when the latter phrase is clearly satisfied. This view is based on its argument that the baseline rules satisfy the connection requirement of Article XX(g) because "[w]ithout baselines of some kind . . . scrutiny would not be possible . . . and the *Gasoline Rule*'s objective . . . would be substantially frustrated." Appellate Body report, *supra* note 5, at 19.

in its view, breached the fundamental rule of treaty interpretation as articulated in Article 31 of the Vienna Convention on the Law of Treaties (the "Vienna Convention").[32] Article 31.1 stipulates that: "A treaty shall be interpreted in good faith in accordance with the ordinary meaning to be given to the terms of the treaty in their context and in the light of its object and purpose." This rule had been relied on by all parties at some stage in their argument and had, in the view of the Appellate Body, attained the status of customary or general international law.[33]

When reading the decision, it would be fair to say that the Appellate Body was concerned to make it very clear that the WTO is not a unique international organization that only needs look at its own terminology and practices for guiding principles; rather, it is one of a number of organizations whose existence is determined by treaty and which must, as a result, follow general principles of treaty interpretation in its adjudicatory methodology. It also pointed out that Article 3(2) of the DSU directs the application of such "customary rules of interpretation of public international law." In the Appellate Body's view, this reflected "a measure of recognition that the General Agreement is not to be read in clinical isolation from public international law."[34]

A number of policy issues can be raised in this regard. First, the Appellate Body's reminder should help disputants understand that because the WTO agreements are treaties operating under principles of international law, the corollary is that many disputed actions are breaches of those legal responsibilities and cannot be defended on the grounds of sovereignty, as the adoption of the treaty with its incumbent obligations was itself a manifestation of sovereign autonomy that required the support of international law for its conceptual validity.

Secondly, while some "GATTophiles" might shudder at the reference to general principles of international law and the citation of ICJ and other international law cases,[35] such a development was inevitable even without express reference in the DSU, and has been going on for some time. As disputes become more and more complex, they will tend to raise more and more choices between literal and purposive interpretation and between different methods of resolving ambiguities, whether through an examination of *travaux préparatoire*, internal context, discerned purpose, or case

32. Vienna Convention on the Law of Treaties, May 23, 1969, 1155 U.N.T.S. 331.

33. This finding was necessary because the Vienna Convention can only apply directly to countries which have ratified it. As of yet, the United States has not ratified the treaty.

34. Appellate Body report, *supra* note 5, at 17.

35. For example, the Appellate Body cited *Territorial Dispute (Libya v. Chad)*, 1994 I.C.J. 6 (Feb. 3). Appellate Body report, *supra* note 5, at 17 n.34.

law. Adjudicators always make such choices, whether they articulate them or not. By referring to accepted and binding international rules, the Appellate Body was making appropriate reference to the only jurisprudentially acceptable principles that could guide it in its task. Furthermore, this approach has certainly been taken by panels from time to time.[36]

This then invites an analysis of the accepted principles of international legal adjudication and a consideration of the implications of adopting these as the guiding principles of WTO methodology. As always, there is a wide field of scholarship that addresses these principles but little by way of clearly accepted content. Broad principles, however, have been recognized.[37] While there has been a debate about the binding nature of rules of interpretation, the inclusion of express provisions in the Vienna Convention makes those rules clearly binding on the parties to that convention at least.

As is the case with municipal courts, a short amount of research will uncover conflicting principles. International law scholars identify the following key alternative approaches to treaty interpretation: literal, purposive (described also as contextual or systematic), historical, subsequent practice, and teleological. The literal approach gives primacy to the plain meaning of the words used. The purposive approach looks to the context and object. The historical approach looks for the original intent of the parties. Subsequent practice looks for an evolving consensus on meaning. The teleological approach, also described as the functional approach, combines these in a normative manner to adopt an interpretation that best gives effect to the function of the provision or treaty.

Categorizing these methods further tends to reveal the primary divide being between a strictly text-based approach and some version of a purpose-based approach. Both approaches start from the same principle but divide on methodology. Each accepts that international treaties are consensual limitations on national sovereignty and as such should be interpreted in the context of consensus. One camp finds consent only in the words of the treaty. The other finds it in the identifiable purpose, through either the policy discernible in the treaty, the *travaux préparatoire*, or a combination. For these reasons, the wording of the interpretative rule as found in the Vienna Convention (like most principles

36. *E.g.*, *United States—Countervailing Duties on Non-Rubber Footwear from Brazil*, SCM/94, Oct. 4, 1989.

37. Schwarzenberger describes it as "International Judicial Law." GEORG SCHWARZENBERGER, INTERNATIONAL LAW AS APPLIED BY INTERNATIONAL COURTS AND TRIBUNALS 3, (vol. IV 1986).

of judicial interpretation) is itself sufficiently wide to enable decisionmakers to adopt quite different approaches and conclusions.

A related and crucial question in adjudicatory methodology, and one which underlies a growing number of WTO disputes, is the relative importance of *substance* and *form*. This debate is not usually conducted in general and express terms within GATT/WTO practice, but is best highlighted by certain categories of disputes. Examples include tariff specialization, and standards that purport to attack real problems but do so in an intentionally discriminatory manner. Standards which are imposed in a practically irrelevant way at home could technically satisfy Article III on a literal interpretation, but still operate as a restriction on imports. The practice of tariff specialization, whereby tariff categories are divided to effectively give a preference to particular suppliers, is best exemplified by the "dappled cow" example cited by Curzon[38] where German legislators drafted a tariff category that could only have applied to Swiss cows. A related issue was the subject of a complaint by the EC against Chile in respect of lower sales taxes on local spirits than on whisky.[39] The specialization issue has also come up indirectly in "like product" cases under Article III. Another example was the environmental levy imposed by the Canadian province of Ontario on beer packaged in cans. Critics of this measure, which technically satisfied Article III of GATT, pointed out that most Canadian beer is packaged in bottles while most U.S. beer is packaged in cans and, further, that the measure did not seek to impose any limits on the use of cans for non-beer products.[40]

The conceptual problem is the way to deal with rules which do not offend the core principles of GATT in the literal sense, but which might do so in a practical sense. This will be an increasingly important area of concern and is also inherent in the chapeau of Article XX, which is further discussed below. A formalistic approach to the substance versus form debate tends to go hand in hand with an overly literal approach to treaty interpretation. The general tendency of judges in the Western world has been to prefer a purposive approach. A failure to adopt such an approach

---

38. GERARD CURZON, MULTILATERAL COMMERCIAL DIPLOMACY 60 n.1 (1965).

39. ROBERT E. HUDEC, ENFORCING INTERNATIONAL TRADE LAW: THE EVOLUTION OF THE MODERN GATT LEGAL SYSTEM, 585 (1993).

40. Peter P. Uimonen, *Trade Rules and Environmental Controversies During the Uruguay Round and Beyond*, 18 THE WORLD ECONOMY 71, 76 (1995). The increasing concern with substance over form and avoidance of WTO principles is perhaps one reason why the negotiators of the Agreement on the Application of Sanitary and Phytosanitary Measures (SPS Agreement) took the chapeau from Article XX of the General Agreement and made it a positive obligation under Article 2.3 of the SPS Agreement. The irony from a legal perspective is that changing it from an exception to an obligation will change the burden of proof and in that sense make it harder for a complainant to succeed.

tends to be an invitation to avoid express obligations wherever ambiguities can be found and casuistic arguments constructed. Other contentious aspects of legal interpretation are shown by the many linguistic presumptions of intent used by adjudicators from time to time. For example, the Appellate Body only felt the need to raise the issue of treaty interpretation because it considered that the Panel failed to take account of the distinction made in Article XX between various adjectives and linkage tests chosen, namely, "necessary," "relating to," "in pursuance of," "essential," "for the protection of," and "involving." Because of these different expressions, it did not seem reasonable to conclude that the drafters intended the same degree of connection in each case. Thus, even a literal approach does not ignore context. Different wording often implies different intent. If this intent is not articulated, adjudicators search for implicit distinctions.

In addition to the problem of identifying meaningful distinctions, the international treaty context poses particular questions as to the reliability of presumptions of intent through the adoption of different forms of expression in various GATT articles. Here there is no safe route to adopt for an adjudicator. On the one hand, the Appellate Body can argue justifiably that it would be a usurpation of legislative power to ignore the difference in meaning between terms such as "necessary" in Article XX(b) and "related to" in Article XX(g). On the other hand, an understanding of the typical pressures, compromises, and meanderings of international treaty negotiations makes it unsafe to read too much into slight differences in wording. This is not to say that the words quoted above have only slight differences, but that such differences often arise because suggested clauses are presented by a whole range of negotiators or are taken in block form from other agreements. The final version may be a last-minute "cut and paste" without the time or political will that is typically available at the domestic level for attention to internal linguistic consistency. Where this is the case, differences in wording do not necessarily display any difference in intent.

The problem remains acute because the knowledge that such differences may sometimes be accidental could never justify an adjudicator always ignoring them. Furthermore, any legal system has to be accepted as having a dynamic and evolutionary existence. Thus if the WTO Members are unhappy with any decisions that make too much out of differences of wording, they can seek to ensure that draft agreements are more carefully vetted to ensure that only intended differences remain. In a domestic system, it is usually the job of parliamentary counsel to ensure consistency in language both within documents and between them. Legal drafting also tends to be better in those systems that involve counsel from the outset,

namely, at the policy design stage. In the WTO context, it would be desirable to at least consider giving the Legal Secretariat, or some of the newly constituted experts groups, an expanded reform and drafting assistance function to address some of these issues. This might be a politically contentious course of action, as differences are at times deliberately used to cloud meaning and hide a fundamental lack of consensus. In the view of the author, however, the development of an equitable and efficient body of international economic law through the WTO would, on balance, be enhanced by such a function.

F. *The Relationship Between Article XX and the Positive Commitments of the General Agreement*

The Appellate Body then turned to the question of the relationship between the affirmative commitments in Articles I, III, and XI and the policies embodied in Article XX. Legal drafters can address such issues by indicating a hierarchy. The opening words of Article XX, referring to the fact that nothing in the various agreements shall be construed so as to take away the enumerated policy rights referred to in Article XX, might suggest some hierarchy. To the Appellate Body, however, the relationship can only be given meaning "by a treaty interpreter on a case by case basis, by careful scrutiny of the factual and legal context in a given dispute . . . ."[41] It is certainly correct to say that the application of these provisions in any particular case will depend on the facts. It is perhaps also wise not to have said too much of a broad nature about any of the provisions, including the relationship between Articles III and XX, but it can surely be said that the two articles should be given effect in their terms and that a provision which is inconsistent with Article III, but which satisfies a limb of Article XX, including the chapeau, is clearly an acceptable measure. At the very least, such a statement by the Appellate Body might have helped the WTO's publicity efforts with the environmental lobby.

G. *The "Made Effective" Requirement in Article XX(g)*

After concluding that the measures satisfied the relationship test in Article XX(g), the Appellate Body then turned its attention to the additional requirement that the measures be "made effective in conjunction with restrictions on domestic production or consumption." The difficult expression is "made effective," as there is no clear purpose in using the

41. Appellate Body report, *supra* note 5, at 18.

word "effective" in this context. The U.S. argued that the expression merely requires that a measure cannot only apply to imported goods. The complainants argued that the expression required that a measure be primarily aimed at making domestic restrictions effective. The U.S. view would be clearly correct if the Article merely said "made in conjunction with." The complainants' view would be appropriate if the Article said "made to effect restrictions on domestic production or consumption." Thus the Appellate Body was faced with the all-too-common problem of looking for an ordinary and sensible meaning in the face of inelegant drafting. It held that the phrase refers to governmental measures being brought into effect in conjunction with domestic restrictions. It did not go further textually and require identical treatment, which would have resulted in the measure satisfying Article III:4 in the first place.[42]

The Appellate Body then dealt with the linguistic argument that to make something "effective" means to make it work. It rejected an empirical "effects test" on practicality grounds, but conceded that the predictable effects of a measure may be relevant in some cases. In particular, if there can be no possible positive conservation effect, then it would normally not satisfy the relationship test.[43]

Whatever the merits or otherwise of an "effects test" from the policy perspective, such a test would be more justifiable if it was directly aimed at the measure itself rather than at the domestic implementation alone. Thus, bringing in such a test merely through the reference to domestic implementation in Article XX(g) is less than optimal. While a more general effects test could be supported on the broad grounds that rules should foster only equitable and efficient environmental measures, the Appellate Body was correct to raise the very serious practical problems of utilizing such a test in an adjudicatory context. This warning should be born in mind by policymakers looking at changes to the trade and environment interface.

H. *The Chapeau of Article XX*

The Appellate Body then turned its attention to the chapeau of Article XX. Even if a measure came within a specific subsection of Article XX, it would not be acceptable if it was "applied in a manner which would constitute a means of arbitrary or unjustifiable discrimination between

---

42. Perhaps another argument why the Appellate Body looked at too wide a body of measures for Article XX(g) is that this wider regime does not sit well with the "made effective" test, which presumably requires the measures affecting international trade to be in conjunction with domestic measures. This would direct attention to the importer rules.

43. This had in fact been the logic behind the Panel's rejection of the measures under the relationship test, and again shows how important the ambit of facts selected was to the Appellate Body's contrary conclusion.

countries where the same conditions prevail, or a disguised restriction on international trade . . . ." The Appellate Body saw the chapeau as addressing the manner of application of the measures rather than their specific content.[44] It noted that the purpose of the chapeau was to "prevent abuse of the exceptions." Thus the exceptions "should not be so applied as to frustrate or defeat the legal obligations of the holder of the right under the substantive rules of the General Agreement." The measures must be "applied reasonably," taking into account all parties' rights if the exceptions are not to be abused or misused. As will be pointed out below, the latter comment constitutes the most important aspect of the Appellate Body's decision and the one which should prove the most contentious.

Before examining the Appellate Body's reasoning, it is worth noting that the key terms—"arbitrary," "unjustifiable discrimination," and "disguised restriction"—are also ambiguous and call for a degree of refinement by an adjudicative body. In particular, the term "unjustifiable" is problematic because it does not refer to the criteria by which to determine justification. For example, is justification to be found within the environmental policy goal, or through the absence of a measure less violative of GATT, or through a combination of both? Is an arbitrary measure one which has no justification or one which has no *reasonable* justification? If reasonableness is imported, how will it be determined?

Because of the lack of clear direction within the provisions themselves, it is worth examining the drafting history, as the Appellate Body did, to see what guidance can be gleaned from this source. The Appellate Body did not quote from the history but concluded that it showed the purpose to be to prevent "abuse." The chapeau wording was taken from an earlier document and developed over time, so that there is no easy site to inspect to determine intent. In the GATT context, and from the perspective of *travaux préparatoire*, the wording originated with the draft International Trade Organization Charter during the Preparatory Committee's London session.[45] One delegation had referred to the problem of "indirect protection" and specifically, the ability of such exemptions to be "misused for indirect protection."[46] Another spoke of the fact that the exceptions "were rational but were open to widespread abuse" and as a result a "provision to prevent abuse" should be incorporated.[47] The

---

44. Appellate Body report, *supra* note 5, at 22.

45. World Trade Organization, *Analytical Index: Guide to GATT Law & Practice*, vol. 1, 563 (6th ed. 1995) [hereinafter GATT *Analytical Index*].

46. GATT *Analytical Index*, *supra* note 45, at 563, *citing Note of the Netherlands and the Belgo—Luxembourg Economic Union, 30 October 1946*, E/PC/T/C.II/32.

47. Comment of South African delegate, E/PC/T/C.II/50 at 6.

Rapporteur of the Procedures Subcommittee stated that "one of the main objectives of Article 30 (the precursor to Article XXIII) was to prevent evasion of the provisions of Chapter IV . . . . It was almost impossible to draft exceptions which could not be abused, if good faith was lacking."[48] The Technical Subcommittee of Committee II at London proposed that:

> In order to prevent abuse of the exceptions of [the Article XX equivalent] . . . the following sentence should be inserted as an introduction: 'The undertakings . . . relating to import and export restrictions . . . shall not be construed to prevent the adoption or enforcement . . . of the following measures, provided that they are not applied in such a manner as to constitute a means of arbitrary discrimination between countries where the same conditions prevail, or a disguised restriction on international trade.'[49]

The GATT Analytical Index suggests that this proposal was generally accepted, subject to later review of the wording.[50] If the drafting intent behind the chapeau was to prevent "abuse," and if that term is to guide interpretation of its provisions, the key question is whether that term connotes an *intent* to subvert the terms of the General Agreement. If this is so, the chapeau would operate as an anti-avoidance or anti-circumvention type provision. Further, support for an intent-based interpretation could be gained from the phrase "which would constitute a means of . . . ." To speak of "means" can be linguistically suggestive of a purposive test.

This would bring the chapeau close to the contentious suggested principle of international law of "abuse of rights." This has been defined by one commentator as a "State exercising a right either in a way which impedes the enjoyment by other States of their own rights or for an end different from that for which the right was created, to the injury of another State."[51] It is contentious because it cannot be clearly identified as a "general principle of law recognized by civilised nations." Civil law countries adopt the principle, but where common law countries are concerned, proponents of the rule are forced to discern the principle from

---

48. E/P/CT/C.II/50 at 6.

49. GATT *Analytical Index, supra* note 45, at 564, *citing* E/P/CT/C.II/50, at 7.

50. GATT *Analytical Index, supra* note 45, at 564. This author's research has not uncovered why "unjustified" was included. The term was not in the review report of EPCT/C.II/54 at 33.

51. Alexandre Kiss, *Abuse of Rights, in* ENCYCLOPEDIA OF PUBLIC INTERNATIONAL LAW, vol. I, at 6 (R. Bernhardt ed., 1992) [hereinafter Bernhardt].

a range of specific precedents in substantive law areas, thus leaving themselves open to the argument that the principle does not exist or that the real principle being applied is that of reasonableness, good faith, or equity.[52]

There would be a number of difficulties with some kind of anti-avoidance interpretation as the key to the chapeau. These include questions such as the appropriate degree of intent, the evidence to be used to determine intent, and the sensitivity of accusing governments of intentionally seeking to frustrate their obligations. On the other hand, to interpret the term in a different manner raises the question of what that manner should be and how this is to be justified under an acceptable interpretation of the provisions themselves. This will be discussed further after examining the way that the Appellate Body utilized the terms.

The Appellate Body provided some further elaboration of the terms "arbitrary discrimination," "unjustifiable discrimination," and "disguised restriction." It felt that as they sit side-by-side, they impart meaning to each other and that, as a result, disguised restriction includes disguised discrimination. It refused to interpret "disguised" in a narrow way that would only cover hidden regulations, but held instead that one of the things covered by the term is any measure that imposes "restrictions amounting to arbitrary or unjustifiable discrimination in international trade taken under the guise of a measure formally within the terms of an exception listed in Article XX." Under this interpretation, factors relevant to the determination of arbitrary or unjustified discrimination could also be used for the determination of disguised restriction. "The fundamental theme is to be found in the purpose and object of avoiding abuse or illegitimate use of the exceptions to substantive rules available in Article XX."[53]

From a policy perspective, it would clearly have been an unreasonable restriction to interpret "disguised" to only mean "hidden." On the other hand, a wider meaning would be difficult to apply because it would draw attention to the intent of the legislator and would allow for complaints that some offending measures are technically acceptable but have improper ulterior purposes. At first sight, this might seem to be the intended meaning in the eyes of the Appellate Body, given its references to "guise," "formally," "purpose," and "object." However, when the Appellate Body expressly elaborated on the meaning of these terms and proceeded to examine the actual measures in terms of the chapeau, it arguably used a

52. Anthony D'Amato, *Good Faith*, *in* Bernhardt, *supra* note 48, at 599–601.

53. Appellate Body report, *supra* note 5, at 25.

different meaning. The Appellate Body, after holding that the chapeau is aimed at preventing abuse, formulated the proposition that to satisfy this requirement, "the measures falling within the particular exceptions must be applied reasonably, with due regard both to the legal duties of the party claiming the exception and the legal rights of other parties concerned."[54] Thus, in this most central and crucial proposition, it interpreted an anti-abuse test as some form of *reasonableness* test.

That the Appellate Body utilized an effective reasonableness test is further supported by the factors it considered relevant to the application of the chapeau. It examined alternative approaches that could have been taken by the United States in order to achieve its stated policy goals, and went on to critically examine the U.S. reasons for not adopting those measures. Nondiscriminatory alternative measures that were considered included the use of either statutory or historical baselines for all without distinctions between local and foreign refiners, and a two-tiered approach that would allow historical for all, subject to appropriate verification. These arguments had also been raised before the Panel, and the Appellate Body agreed with the Panel's rejection of the U.S. justifications.[55] In particular, U.S. concerns with extraterritorial verification were not thought to be compelling given that country's vast experience with foreign analysis and verification in antidumping proceedings. The United States' stated concern of forcing some domestic refiners to reform too quickly if statutory baselines were applied across the board, showed that it failed to properly consider the same problem facing foreign refiners. Furthermore, the Appellate Body considered that attempts could have been made to enter into cooperative arrangements with the governments of Venezuela and Brazil. By looking in detail at the discriminatory aspects and comparing them to less restrictive alternatives, it is also arguable that the Appellate Body looked at a narrower range of measures with respect to the chapeau than with respect to Article XX(g).

The final reasoning of the Appellate Body on this most crucial issue was not outlined in any further detail, perhaps because of the concern for the political sensitivity of the decision to apply the chapeau. It stated that these two failures on the part of the United States went well beyond what was necessary to find a violation of Article III:4, and further, that the resulting discrimination "must have been foreseen and was not merely inadvertent or unavoidable." This supported the conclusion that the baseline establishment rules constituted "unjustifiable discrimination" and

54. Appellate Body report, *supra* note 5, at 22.A.

55. Panel report, *supra* note 5, at para. 6.28.

a "disguised restriction on international trade."[56] In effect, the decision adopts both a foreseeability and a reasonableness test and concludes that unreasonable discrimination in terms of the desired policy goals that is more than merely inadvertent or unavoidable offends against the chapeau.

The linguistic and policy concerns with an intent-based, anti-avoidance interpretation of the chapeau were pointed out above. There are also a number of linguistic and policy concerns with a test that looks at foreseeability, avoidability, and reasonableness. It could imply that the chapeau provisions are to be interpreted as relating only to intentional, rather than inadvertent, abuse. Yet, something can surely be unjustifiable even if accidental. Secondly, both this and a foreseeability test draw attention away from the actual form, manner, and extent of discrimination, to the expectations or reasonable expectations of the party employing the measure.

An avoidability or reasonableness test that specifically calls attention to the respective rights of the parties will also tend to equate to the "necessity" test as employed under Article XX(b), given that a reasonably available, GATT-legal alternative will show avoidability. This leads to the linguistic question of the justification of an interpretation that holds that the effect of the chapeau is to indirectly superimpose a necessity test on Article XX(g), notwithstanding the Appellate Body's conclusion that the words "relating to" suggested a different intent. Critics could ask why the Appellate Body did not see an intended distinction in meaning through the use of the distinct terms—"arbitrary," "unjustifiable," and "disguised"—as opposed to the term "necessary."

Secondly, such an interpretation would have very little if any operation on those subparagraphs that refer to "necessary" and "essential." Because the Appellate Body took such a wide approach to the relationship test in Article XX(g), there was nothing remarkable in the dual finding that the measures came within Article XX(g), but that they fell afoul of the chapeau. If the case had instead dealt with Article XX(b), it would be rare for facts to allow such a dual finding to be made, as a measure which satisfied the current interpretation of the necessary test would surely not be seen as arbitrary or unjustified. An earlier panel had suggested that an inconsistent measure could not be seen as "necessary" "if an alternative measure which it could reasonably be expected to employ and which is not inconsistent with other GATT provisions is available to it."[57] This

---

56. Appellate Body report, *supra* note 5, at 28–29.

57. *United States—Section 337 of the Tariff Act of 1930, Report by Panel*, GATT Doc. L/6439 (Nov. 7, 1989) 345, 392 para. 5.26.

would be so unless the chapeau was limited to administrative or judicial applications, as discussed further below.[58] When this is looked at alongside the above-mentioned effect on the Article XX(g) relationship test of undermining any intended difference to the other tests, the chapeau is left with an unusual sphere of operation.

Thirdly, a reasonableness test is not the natural way to rephrase the term "abuse." It could also be asked whether the phrase "disguised restriction" should best be interpreted to be identical to that of "unjustifiable discrimination." The more natural and distinct field for its operation would be provisions that on their face purport to have the required goal, but which on more careful analysis cannot be explained on any basis other than being aimed at trade restriction per se.

The above criticisms are essentially technical and linguistic. From a policy perspective, a reasonableness test may certainly be the first choice as a guiding parameter in some part of the treaty provisions, whether the chapeau or otherwise. It is certainly desirable to try to encourage governments to select the most efficient and legally valid policy alternative. From a legal administration perspective, the difficulties with such a goal arise from the problems facing the drafters in providing tests of reasonableness, and for adjudicators in applying such provisions in a rigorous and consistent manner. Some of the factors raised by the Appellate Body as problems with an "effectiveness" test under Article XX(g) would again be relevant, although "reasonableness" may be an easier concept for an adjudicator to determine confidently than long-term effectiveness. A further problem would be that many highly contentious factors would need to be considered in the context of a determination of reasonable political behavior.

These factors were less of a problem in the context of the *Reformulated Gasoline* case, simply because both adjudicative bodies discounted all of the United States' argued justifications. The more difficult questions will have to be resolved when, for example, a case arises concerning a slightly less restrictive provision, where arguments may be made either way about its relative efficiency in terms of the mandated policy goal.

---

58. Nevertheless, that such a dual finding was at least possible in the minds of the original drafters is shown by the record of discussions of the Third Committee at the Havana Conference, which said:

> The Committee agreed that quarantine and other sanitary regulations are a subject to which the Organization should give careful attention with a view to preventing measures 'necessary to protect human, animal or plant life or health' from being applied in a manner which would constitute a means of arbitrary or unjustifiable discrimination or a disguised restriction in international trade . . . .

GATT, *Analytical Index*, *supra* note 45, at 565, *citing* E/CONF.2/C/3/SR.35.

I. *Onus and Burden of Proof Under the Chapeau*

A second practical and conceptual question relating to the application of the chapeau is that of the requisite onus and burden of proof. The Appellate Body followed the standard approach of requiring the party invoking the exception to satisfy all elements of it, including the chapeau requirements, but went on to say that this "is, of necessity, a heavier task than that involved in showing that an exception, such as Article XX(g) encompasses the measure at issue."[59] No reasons were given in support of this conclusion. In upholding the view that parties seeking to rely on an exempting provision should have the onus of proof under it, the Appellate Body's approach is consistent with that of previous panels.[60] The reference to a heavier burden is a novel one. It would not seem justified if the chapeau was held to only cover intentional abuse, as there would need to be some evidence of this before a tribunal could make an adverse finding. If a *reasonableness* test is to apply, it is an understandable comment, as it is more difficult to show that a measure is a reasonable one in the context of all parties' rights, than to merely show that it is a measure "relating to" a particular subject area.[61]

CONCLUSION

The first Appellate Body decision has been examined in detail in order to analyze the reasoning and methodology adopted, both in the context of the particular dispute and also for the broader purpose of drawing conclusions about adjudicative methodology in general under the WTO system. Where the latter aim is concerned, it would of course be foolish to try to draw too many concrete conclusions as a result of the very first Appellate Body decision. It will take a number of cases before any clear trends and principles will emerge. The article has thus concentrated on the key questions and their competing policy arguments.

---

59. Appellate Body report, *supra* note 5, at 22.

60. *See, e.g.*, Canada—Administration of the Foreign Investment Review Act, May 26, 1983, GATT B.I.S.D. (30th Supp.) at 164 (1984); United States—Section 337 of the Tariff Act of 1930, Nov. 7, 1989, GATT B.I.S.D. (36th Supp.) at 393 (1990); United States—Restrictions on Imports of Tuna, Feb. 18, 1992, GATT B.I.S.D. (39th Supp.) at 197 (1993).

61. The Appellate Body next considered the qualifying words "between countries where the same conditions prevail" that limit the proscribed "arbitrary" and "unjustifiable discrimination" and noted that these words were ambiguous. The Body was concerned with the potential argument that the words were only aimed at discrimination between different exporting countries, in which case it would apply to Article I disputes but not Article III disputes. The general wording of the *chapeau*, the failure to limit the express provisions to "foreign" or "third" countries and the failure of the United States to rely on this argument led the Appellate Body to conclude that no ruling was necessary and that it could rely on the parties' acceptance of its relevance to the instant dispute. Appellate Body report, *supra* note 5, at 22–24.

As far as the Appellate Body's broad approach is concerned, both the stated and unstated principles in the *Reformulated Gasoline* case show that there is likely to be an ongoing and inevitable conflict between, on the one hand, the desire to be conservative in approach and not be seen to be developing new legal norms, and, on the other hand, the need to display in a logical and clear fashion the reasons for the decision and the responses to the panel decision appealed from. This tension will be exacerbated as WTO disputes gravitate more and more towards the greyer areas of the rules, such as the key terms considered in the *Reformulated Gasoline* case.

A key thesis of this article is that because the legal norms and tests to be employed in these areas are less certain, any adjudicator will need to expand on the meaning of those tests and will as a result be open to superficial accusations of lawmaking. In reality, this is an inevitable corollary of a binding adjudicatory model operating on such difficult provisions. Rather than question or criticize this in a broad sense, it is preferable to consider the many subsidiary questions that face such adjudicators. An understanding of the problems of adjudication should lead to a number of results and principles.

First, international adjudicatory reports do not and should not constitute binding precedents, although correct and logical decisions are likely to be consistent and past cases will always provide material for analysis and analogy. Adjudication is merely an alternative to successful dispute resolution by the parties themselves. Regardless of the inherent quasi-lawmaking function of adjudicators, adjudication thus remains a consensual event. The only common value needed in any international legal institution is respect for the law. In the case of treaties it is best encapsulated by the principle of *pacta sunt servanda*, which states that agreements are meant to be kept. An adjudicatory process is an essential element in any system which aims to protect the rights of parties against breaches of that principle.

An understanding of ambiguities and the choices they allow should lead to an increase in the resources given to drafting in the WTO, in order to minimize the number of unnecessary contentious disputes and to assist adjudicators in developing widely respected decisions. This is particularly important now that the WTO has a built-in reform agenda within its constituent documents.

WTO adjudication should be both encouraged and expected to follow the trend towards purposive interpretation, which builds on general notions of reasonableness and good faith, but which will inevitably lead to superficial cries of interference with sovereignty and will thus need to be strongly defended. WTO adjudication should, as a result, look for a coherent theory of the various agreements. Coherence was inherently an

issue in *Reformulated Gasoline* because of the interplay between the general and exempting provisions. This will be a growing problem given the wider number of agreements established in the Uruguay Round and their relationship to each other and to GATT 1994. More thought will need to be given to this issue in the future, both from a constitutional perspective and also from an optimal policy perspective.

In considering the proper policy relationship between Articles I and III on the one hand and Article XX on the other, there is a need to consider the possible measures that could justifiably fit within the latter and which must inevitably offend against the former. This in turn calls into question the various panel decisions on those articles, as the applicable tests are often unclear, unstated, or have been seen differently by different panelists and commentators. Changes to those tests will alter the relationship between the provisions and may even remove any inherent conflict. It has not been the purpose of this article to examine this question; however, as an example, it would be possible to read Article III so as to ignore *de minimus* differences or those which are reasonable in terms of sensible administrative practice. GATT Panels have already interpreted Article III:1 to allow product differentiation "for policy purposes unrelated to the protection of domestic production."[62] There is also the question of defining discrimination in either a technical or practical sense, an example of the substance-over-form problem. If these interpretations of Article III allow formalistic compliance, in many cases there will be no need to get to Article XX.[63]

A second issue is the means of determining the meaning of key phrases. The Appellate Body shied away from developing a paraphrased test of the Article XX(g) language, but effectively did so where the chapeau was concerned. It combined literal, historical, and purposive approaches to interpretation and made it very clear that the WTO is not separate from the general international law regime.

In these vague areas, the general approach that the Appellate Body takes towards judicial review will also be crucial. The first case showed that the Appellate Body was not shy to consider both legal and factual issues afresh. The related thesis was that in the context of a connecting test, the ambit of facts selected for consideration will be crucial to the outcome in many cases.

---

62. United States: Measures Affecting Alcoholic and Malt Beverages, Feb. 18, 1992, GATT B.I.S.D. (39th Supp.) at 276 (1993).

63. See GATT *Analytical Index*, *supra* note 45, at 151–55, which argues that there has been an expansion of the meaning of discrimination to encompass *de facto* discrimination.

Many commentators will no doubt argue whether the Panel was "right" or whether the Appellate Body was "right," but in one sense, this is an inappropriate way to examine the adjudicatory process. The discussion has shown that each approach has many arguments in support and many arguments against. This is not to say that either was wrong, but rather, that adjudication based on uncertain phrases will never allow for unassailable reasoning. In addition, it is too easy to see differing approaches to factual determinations as demanding vastly different interpretative principles. For this reason, a better approach is to ask a broader series of doctrinal and political questions. The doctrinal ones would include whether each approach was arguably correct on interpretative and policy grounds. Why did they diverge? What, if anything, in the language allowed for divergent approaches? Does this imply that the language should be changed? Here it can be concluded that the provisions facing the adjudicators were less than clear, but WTO Members would surely be supportive of an approach that attempted an integrated and policy-oriented approach to determining the role of a qualified exemption provision which relates to a broad policy goal. While the two adjudicative bodies differed in the application of the relationship test, they primarily did so because they looked at varying ambit of measures. However, both held that a measure which was discriminatory and which had no discernible means of advancing the policy goal, or was clearly inferior to other measures, could not satisfy Article XX(g) as qualified by the chapeau. Only a highly technical, narrow, and non-purposive approach to legal reasoning would have led to a wholly contrary conclusion.

In terms of the different approaches, the policy advantage of moving the discussion to the chapeau, as the Appellate Body did, is that it makes the analysis of Articles XX(b) and (d) more consistent with the Article XX(g) analysis. Certainly from a policy point of view, it is hard to justify a distinction between health and resources where the former has the more difficult test. At the same time, as the Appellate Body itself made clear, the differences in wording should not be ignored.

On balance, it is suggested that the Article XX(g) analysis might best be pursued by adopting the notion of "a significant aim" as the essence of the required relationship, and applying that test objectively and with due deference to sovereign rights and the problems of proof,[64] to the narrowest ambit of measures which offend against the positive norms and which still display a discrete object. However, even on such a basis, competing conclusions are still possible.

---

64. For a recent analysis of the question of standards of review and the policy of deference, see Steven P. Croley and John H. Jackson, *WTO Dispute Procedures, Standard of Review, and Deference to National Governments*, 90 AM. J. INT'L L. 193 (1996).

Perhaps the correct political questions to ask are whether the "right" party won, and if there was general happiness with the result. Should there be some reasonableness test, both from the perspective of emerging international environmental law principles, and from the perspective of a broad and integrated purposive interpretation of GATT provisions? It can be said that the right party won in terms of predictions, or if one utilizes a broad *prima facie* approach of strongly questioning the consistency of any discriminatory measure that has no powerful efficiency argument in its favor, or if one adopts the purposive approach discussed above. In addition, it appears that all parties and most Members have been supportive of the result. The convergence of principles dealing with the trade and environment interface will also help the future work program in this area.

Certainly, the first Appellate Body decision can lead to many debates about substance and methodology. But on the basis of positive answers to each of the above questions, both the Panel and Appellate Body decisions can be supported as reasonable responses to difficult theoretical questions and thus as positive elements of the body of GATT/WTO dispute settlement practice.

In addition, however, the dispute highlights a range of legal adjudicatory issues that are well worth analysis along with the substantive areas of international trade concerns. The dispute settlement function has grown in importance as the GATT and then the WTO have moved well away from concentrating primarily on tariff reductions and tariff bindings as a broad impetus for market access. The new agreements show how difficult the modern concept of market access is and the types of rules needed to foster it. Legal measures to identify, promote and support such access will continue to need a robust, careful-but-determined dispute settlement system to protect and expound upon the agreements reached from time to time.