

NORMATIVE VALUE JUDGMENTS IN WTO ADJUDICATION: THE ROLE OF MORAL REASONING AND
INTUITIONS IN THE SEALS DISPUTE

by

BENJAMIN PAUL CZAPNIK

A thesis submitted to the University of Birmingham for the degree of
DOCTOR OF PHILOSOPHY

Birmingham Law School
College of Arts and Law
University of Birmingham
April 2020

**UNIVERSITY OF
BIRMINGHAM**

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

This dissertation uses doctrinal and socio-legal methods to analyse the Article XX necessity test in WTO law. Using case studies from *Korea Beef*, *Brazil Tyres* and *Seals*, it analyses the mechanisms used by adjudicators to navigate the unresolved tension in WTO law between the retained and ceded regulatory autonomy of Members.

The analysis relies on deconstruction and behavioural techniques to make broader claims about judicial decision-making. Drawing from Thaler's work on "anomalies" in behavioural economics, the analysis looks at what the use of "double standards" can reveal about judicial decisions. I develop new techniques for distinguishing between random "double standards" in judicial reasoning and those which reflect systematic bias.

I further draw from "dual process theory" in cognitive and social psychology to assess how judges make decisions on matters where they may be unconsciously influenced by strong intuitions. I suggest that the discipline of behavioural law can identify when adjudicators are rationalising an intuitive judgment (as was the case in *Seals*), rather than using legal or moral reasoning to reach a decision. Finally, I investigate what the tendency for post-hoc rationalisation means for judicial decision-making and moral progress.

Acknowledgments

I would like to thank my PhD supervisors, Luca Rubini and Aleksandra Čavoški, for their guidance and support. I would also like to acknowledge a number of friends and colleagues for insightful comments and discussions, including Graham Cook, Jesse Kreier, Pierre Sauvé and Joost Pauwelyn. For their personal support over the years, I thank my partner, Vittoria, and my parents.

Contents

CHAPTER 1 – INTRODUCTION: QUESTIONS, ARGUMENTS AND CONCEPTUAL FRAMEWORK	1
1.1 Research Question(s)	1
1.2 Argument	2
1.3 Methods	3
1.3.1 Doctrinal Analysis.....	3
1.3.2 Socio-Legal Analysis	5
1.3.3 Case Studies.....	6
1.3.4 Deconstruction.....	6
1.4 Conceptual Framework (and Legal Theory).....	8
1.4.1 Legal Realism	9
1.4.2 Cognitive Psychology and Behavioural Economics.....	13
1.4.3 Moral Psychology and Normative Reasoning.....	17
1.5 My Analysis and the Broader Literature	23
1.5.1 WTO Doctrinal Law.....	23
1.5.2 Judicial Decision-Making.....	25
1.6 Chapter Summary	27
CHAPTER 2 - RATIONALE FOR THE GATT/WTO: BALANCING REGULATORY AUTONOMY AND PROTECTIONISM.....	32
2.1 Rationale for Creation of the GATT/WTO	32
2.1.1 The Standard Model	34
2.1.2 Practitioners' Story	36
2.1.3 Summary	40
2.2 Legitimate Regulation and the WTO	40
2.2.1 Trade Measures and Domestic Measures	40
2.2.2 The National Treatment Rule	43
2.2.2.1 Key Concepts under the NT Principle.....	43
2.2.3 SATAP, Regulatory Purpose and Protectionist Intent	46
2.2.3.1 SATAP	47
2.2.3.2 Regulatory Purpose: Article III or Article XX?	51
2.2.3.3 How to Assess Regulatory Purpose	53
2.2.4 General Exceptions (Article XX)	55
2.3 Legal Interpretation	55
2.3.1 Textual Interpretation.....	56

2.3.2 Object and Purpose	66
2.3.3 Legitimacy of the WTO System.....	70
CHAPTER 3 - THE ARTICLE XX LEGAL TEST AND STANDARD OF REVIEW	75
3.1 Introduction.....	75
3.2 Elements of the Legal Test	75
3.2.1 Provisional justification	76
3.2.1.1 Threshold Examination	76
3.2.1.2 Necessity Test.....	77
3.2.2 Chapeau	78
3.2.3 Policy Objective	79
3.3 Brief History of the Necessity Test(s)	81
3.3.1 Alternative Measures Test	81
3.3.2 The (Gradual) Introduction of the W&B Test.....	83
3.3.3 Summary	85
CHAPTER 4 - THE W&B AND AM TESTS: DOCTRINAL ANALYSIS AND CRITIQUE	86
4.1 Introduction.....	86
4.2 W&B Test: Doctrinal Analysis	86
4.2.1 Importance	86
4.2.2 Contribution and Trade Restrictiveness	89
4.3 AM Test: Doctrinal Analysis	93
4.3.1 Less Trade-Restrictive (or Less Treaty-Inconsistent).....	93
4.3.2 “Equivalent Contribution” (or “Desired Level of Protection”).....	94
4.3.3 Reasonable Availability	97
4.4 Interaction between the Tests	98
4.5 Summary of Doctrinal Critique.....	100
CHAPTER 5 - THEORETICAL CRITIQUE	102
5.1 Introduction.....	102
5.2 W&B Test (and Kaldor-Hicks Efficiency).....	102
5.3 AM Test (and Pareto Efficiency).....	107
5.4 Interaction between the Tests	108
5.4.1 Regulatory Autonomy and the Object of Review	109
5.4.2 Logical Contradictions	119
5.5 Summary of Theoretical Critique.....	122
CHAPTER 6 - DECONSTRUCTING THE NECESSITY TEST.....	124

6.1 W&B Test.....	124
6.1.1 Importance	124
6.1.2 Balancing Step.....	126
6.1.3 Purpose of W&B Test.....	127
6.1.4 Summary	129
6.2 AM Test: Korea Beef Case Study	129
6.2.1 Introduction.....	129
6.2.2 Critique 1 – Level of Protection	130
6.2.2.1 Re-characterising the Level of Protection	130
6.2.2.2 Lowering the Level of Protection	133
6.2.2.3 The Margin of Appreciation	135
6.2.3 Critique 2 - Financial and Administrative Burdens.....	137
6.2.4 Summary	143
CHAPTER 7 – THE BRAZIL TYRES CASE STUDY.....	146
7.1 Introduction to <i>Brazil Tyres</i>	146
7.2 Critique of <i>Brazil Tyres</i>	147
7.2.1 Critique 1: Three Different Benchmarks	147
7.2.2 Critique 2: The Objective is Too Narrow	152
7.2.3 Critique 3: The Benchmark Contained No “Magnitude”.....	155
7.2.4 Critique 4: The Benchmark was “Impermissible”	158
7.3 Behavioural Economics, “Anomalies” and Double Standards	159
7.3.1 Thaler, Anomalies and the Law.....	159
7.3.2 Replicating the <i>Brazil Tyres</i> Anomaly.....	163
7.3.3 Anomalies plus Psychology	165
7.4 Lessons from <i>Brazil Tyres</i>	167
7.4.1 The Policy Objective in WTO Law.....	167
7.4.2 WTO Adjudication and Value Judgments.....	169
7.4.3 <i>Brazil Tyres</i> and Doctrinal Law	171
7.5 Brazil Tyres Summary.....	173
7.6 Reflections: The Necessity Test and Regulatory Autonomy	174
CHAPTER 8 – THE PUBLIC MORALS JURISPRUDENCE	181
8.1 Introduction.....	181
8.2 Are Public Morals Different?	182
8.3 WTO Jurisprudence	189

8.3.1 Unilateralism and the Legal Test	190
8.3.1.1 Academic Views	190
8.3.1.2 Jurisprudence	192
8.3.2 Evidentiary Practice and External Standards	192
8.3.2.1 Originalism	193
8.3.2.2 Universalism and Transnationalism	194
8.3.2.3 Internationalism	195
8.4 Commentary	197
8.5 Expert Evidence	203
8.5.1 Science and/or Philosophy	203
8.5.2 Reconciling the Roles of Science and Philosophy	206
8.5.3 What Makes Public Morals Unique?	207
8.5.4 Summary	211
8.6 The Seals Dispute	212
8.6.1 Legal Test	212
8.6.2 Application of Legal Test	213
8.6.2.1 Public Concerns	213
8.6.2.2 Moral Standards	214
8.6.3 Summary	216
8.7 Conclusion	217
CHAPTER 9: THE WELFARIST DEFENCE	219
9.1 Moral Philosophy in WTO Law	219
9.2 Animal Welfarism	220
9.2.1 Introduction	220
9.2.2 The Suffering/Utility Standard	221
9.2.3 The Level of Protection	225
9.2.4 Non-Discrimination and Speciesism	228
9.2.5 Summary	233
9.3 Animal Rights	234
9.3.1 Introduction	234
9.3.2 The Standard of Review	236
9.3.3 Summary	240
9.4 Conclusion	240
CHAPTER 10: INTUITIONISM	242

10.1 Introduction.....	242
10.2 Intuitionism and the Seals Dispute	243
10.2.1 EU/Amici Arguments for Intuition.....	243
10.2.2 Perišin/Sellheim Opposition to Intuition.....	245
10.3 Philosophy: The Singer/Posner Debate.....	247
10.3.1 Posner's Defence of Intuitionism	248
10.3.1.1 Intuition Trumps Logical Argument	248
10.3.1.2 Intuition Trumps Consistency/Cohherence	249
10.3.1.3 Pure Intuitionism - Dog Product Bans	251
10.3.1.4 Summary	253
10.3.2 Welfarism's Rejection of Intuitionism	253
10.4 Is Posner a "Pure Intuitionist"?.....	254
10.4.1 Suffering Utility Standard.....	254
10.4.2 Posner's Unified Theory.....	256
10.5 WTO Jurisprudence	260
10.5.1 Does Article XX(a) Allow Posner's "Intuitionist" Measures?.....	260
10.5.2 Jurisprudential Reform	262
10.6 Conclusion.....	263
CHAPTER 11 - THE HYBRID APPROACH: WELFARISM MIXED WITH INTUITIONISM..	264
11.1 Introduction.....	264
11.2 Welfarism and Intuitionism are Mutually Exclusive	265
11.3 Logical Fallacy I: Equivocal Language	267
11.3.1 Adjudicators' Approach	267
11.3.2 EU's Approach.....	269
11.4 Logical Fallacy II: Contradictory Reasoning.....	271
11.4.1 The Amici and Instrumental Reasoning	271
11.4.2 Adjudicators: Science and Risk	273
11.5 Logical Fallacy III: Double Standards.....	280
11.5.1 Design Step	280
11.5.1.1 Public Concerns (Seal Welfare)	281
11.5.1.2 Moral Standards (Animal Welfare)	282
11.5.1.3 Summary.....	286
11.5.2 Weighing and Balancing Test	287
11.5.2.1 Importance (Animal Welfare)	287

11.5.2.2 Contribution (Seal Welfare).....	289
11.5.2.3 The Balancing Step	291
11.5.3 Summary	292
11.6 Conclusion.....	293
CHAPTER 12: WELFARISM AND REGULATORY CONSISTENCY	295
12.1 Regulatory Consistency	295
12.1.1 Consistency, Article XX(a) and Animal Welfare.....	295
12.1.2 Consistency Requirements in <i>Seals</i>	298
12.1.2.1 Academic Arguments.....	298
12.1.2.2 Canada's Consistency Argument.....	299
12.2 Avoiding the Consistency Requirement.....	301
12.2.1 Silence on Consistency and Policy likeness.....	301
12.2.2 The "Same Treatment" Argument	303
12.2.3 The Hypocrisy Argument	306
12.3 – Rejecting Consistency Requirements.....	307
12.3.1 Adjudicators Arguably Reject Consistency	307
12.3.2 Academic Arguments for Rejecting Consistency	312
12.3.3 Summary	316
12.4 Applying the Consistency Requirement.....	316
12.4.1 Seals Suffer more than Commodity Animals	317
12.4.2 Seal Slaughter is "Inherently Inhumane"	324
12.5 Conclusion.....	328
CHAPTER 13 - CONCLUSION AND FURTHER REFLECTIONS	330
13.1 Conclusion.....	330
13.2 Further Reflections on Law, Morality and Social Norms	338
13.2.1 "Coherent" Approaches: Reasoning and Intuition	339
13.2.2 The "Incoherent" Approach: Rationalisation	342
13.2.3 Rationalisation, Public Morals and the <i>Seals</i> Dispute	343
13.2.4 Law, Moral Progress and Behavioural Adjudication.....	347
TABLE OF GATT PANELS AND WTO DISPUTES.....	351
TABLE OF TREATIES AND INTERNATIONAL INSTRUMENTS	354
BIBLIOGRAPHY.....	355

List of Abbreviations

AB	Appellate Body
ADA	Anti-Dumping Agreement
AM	Alternative measure
CBA	Cost-benefit analysis
CITES	Convention on International Trade in Endangered Species of Wild Fauna and Flora
DSB	Dispute Settlement Body
DSU	Dispute Settlement Understanding
EFSA	European Food Safety Authority
EC	European Communities
EU	European Union
GATT	General Agreement on Tariffs and Trade
IMF	International Monetary Fund
LTI	Least treaty-inconsistent
LTR	Least trade-restrictive
OIE	Office International des Epizooties
MFN	Most favoured nation
NT	National treatment
SPS	Sanitary and Phytosanitary
TBT	Technical Barriers to Trade
UNCLOS	United Nations Convention for the Law of the Sea
US	United States
VCLT	Vienna Convention on the Law of Treaties
W&B	Weighing and balancing
WHO	World Health Organisation
WTO	World Trade Organisation

CHAPTER 1 – INTRODUCTION: QUESTIONS, ARGUMENTS AND CONCEPTUAL FRAMEWORK

1.1 Research Question(s)

This thesis seeks to address two questions and can be loosely understood as containing two parts. The first part, covering chapters 2-7, focusses primarily on WTO doctrinal law, especially the necessity test under GATT Article XX(a). It considers a perennial question in WTO law about where the line should be drawn between retained and ceded regulatory autonomy. I focus on this question, which underpins much WTO litigation, in the context of domestic regulatory measures which pursue social (non-trade) goals, but which have adverse impacts on foreign trade. This includes some landmark WTO disputes, such as *Korea Beef, Brazil Tyres and Seals*.¹ I will seek to explain how adjudicators have used the necessity test to define the policy space retained by Members to pursue social objectives through trade-restrictive measures.

The second part, covering chapters 8-12, uses WTO law as a site for making broader socio-legal claims about judicial decision-making. There is a strong tradition at the WTO which suggests that adjudicators should not base their decisions on normative value judgments which second-guess the regulator's decision. The mandate of WTO adjudicators requires them to make difficult judgment calls (about moral, health or environmental issues) but appears to equip them with little more than the tools of legal formalism. This thesis analyses how WTO adjudicators make

¹ Each of these disputes will be analysed in a dedicated case study.

(and justify) their decisions in “hard” cases when faced with constraints imposed by WTO law and the general principles of legal reasoning.²

1.2 Argument

I argue that WTO adjudicators have not been equipped with the mandate and tools they need to draw a clear dividing line between retained and ceded regulatory autonomy. Many difficult cases require adjudicators to make sensitive normative judgments (and to balance trade interests against other social concerns), but the Member-centric culture at the WTO prevents them from undertaking such intrusive analysis.³

In order to resolve hard cases, adjudicators make value judgments about the “necessity” of certain trade-restrictive social measures, but the culture of the WTO precludes them from making such judgments explicitly. This forces adjudicators to hide their value judgments (and therefore the real reasons for their decisions) behind reasoning which is ostensibly value-neutral.

This is an impossible task. In order to render their value judgments opaque, adjudicators must creatively manipulate their reasoning in ways which lead to logical fallacies and contradictions. The WTO maintains a façade of value-neutral adjudication at the expense of coherent reasoning. While reasoning errors might be

² While this thesis can loosely be understood as containing two major themes (WTO doctrinal law and judicial decision-making), I refrain from formally dividing it into two parts, primarily as there is significant overlap between the themes. For example, the *Brazil Tyres* case study (in the first part) includes techniques from behavioural economics to show that some legal reasoning was not “rational” and contained “systematic bias”. In the second part, chapter 8 contains some detailed WTO doctrinal analysis, including the test for “public morals” under Article XX(a).

³ This is in contrast with constitutional systems of adjudication such as the US. In the next chapter, I will compare the WTO’s “negative integration” model with more highly-integrated systems of centralised decision-making such as the US and EU.

less visible than explicit value judgments, they are not altogether invisible. They can be uncovered and exposed by legal scholars.

In order to critique the reasoning and decision-making of WTO adjudicators, I situate my analysis within the broader literature about decision theory, psychology and behavioural economics. I draw on the tools of behavioural economics to demonstrate when adjudicators adopt reasoning which is not “rational”. I rely on “dual process theory” from psychology (often referred to as System 1 and System 2) to distinguish between intuitive and rationalist modes of thinking.⁴

I apply these insights to the matter of animal exploitation (*in Seals*) as this is a sensitive matter where people tend to have powerful moral intuitions. I argue that in such disputes, the real reasons for a decision might be based on nothing more than intuition (System 1) which adjudicators subsequently seek to rationalise using moral reasoning (System 2). Since these rationalisations ultimately depend on incoherent reasoning, I seek to expose them by identifying widespread logical fallacies and contradictions in the *Seals* dispute.

1.3 Methods

My analysis will rely on two methods (doctrinal and socio-legal analysis) and two analytical “tools” (case studies and deconstruction). This section will describe these methods and tools.

1.3.1 Doctrinal Analysis

The thesis undertakes a detailed doctrinal analysis of WTO law, especially the necessity test under Article XX and the public morals exception. In the first part of

⁴ These concepts will be explained in section 1.4 below.

the thesis, I seek to explain how Article XX became the key site for reserving regulatory autonomy for domestic measures (despite the strident view of some Members and scholars that this should be addressed within Article III). I explain the key elements of an Article XX defence and focus on how the necessity test evolved in GATT panels and leading WTO disputes. The *Seals* case study also contains a dedicated chapter on the public morals jurisprudence.

The doctrinal analysis focusses on the leading "necessity" and public morals disputes, including case studies on *Korea Beef*, *Brazil Tyres* and *Seals*. I also draw extensively on the academic literature to identify the main critiques and theoretical justifications for the different approaches to necessity testing.

The doctrinal analysis lays the foundation for my subsequent socio-legal inquiries. A clear understanding of the content of WTO law is a pre-condition for my argument that adjudicators deviated from established legal rules or otherwise relied on incoherent reasoning. The doctrinal analysis also shows that some rules have developed in an equivocal manner which offers adjudicators a "menu" of options rather than constraining them to apply a clear rule.

Since I focus on the necessity test (and briefly on the public morals threshold examination), there are several important Article XX themes which I do not address. I refrain from detailed analysis of the chapeau and the extra-territorial application of regulatory measures.⁵ These are important issues, including in some of the disputes I focus on, but they fall outside the scope of my analysis.

⁵ That said, the chapeau will be introduced in general terms in section 3.2.2.

1.3.2 Socio-Legal Analysis

This thesis relies on socio-legal methods in two ways. First, drawing on the techniques of behavioural economists, I seek to demonstrate that (i) adjudicators sometimes rely on non-rational techniques such as “double standards” to justify their decisions; and (ii) this lack of rationality may be systematic (rather than random) in some cases.

It is not uncommon for legal scholars to critique an adjudicator’s decision by suggesting that it contains flawed reasoning and is therefore not fully rational. However, unlike psychology and behavioural economics, legal scholarship has devoted less attention to identifying systematic bias. Perhaps legal scholars assume that non-rational decisions are merely a random deviation from the rationalist norm of judicial decision-making. In my view, the possibility of systematic bias merits deeper investigation and I explore this in detail in the *Brazil Tyres* case study.

Second, I will draw on cognitive and moral psychology to seek to explain the underlying psychological drivers of the adjudicators’ decision in *Seals*. This analysis will replicate my approach to *Brazil Tyres* as I will show that (i) the adjudicators’ decision was not rational as it relied on logical fallacies and contradictions, and (ii) the use of flawed reasoning systematically favoured the same party (in that case, the EU regulator). However, the *Seals* case study will go one step further and seek to explain why adjudicators favoured the EU, even if this required them to circumvent WTO doctrinal law and general principles of legal reasoning.

Drawing from moral philosophy, I will show there is no coherent (or rationalist) argument which can justify the special protection of dogs/seals alongside the painful exploitation of commodity animals on factory farms. However, psychologists have

shown that humans have strong intuitions about which types of animal exploitation they consider to be moral. WTO adjudicators faced a tension between their (professional) obligation to undertake a rationalist assessment of the matter before them and their (human) intuition that some animals (such as seals) deserve special protection. I will argue that their human intuition trumped their legal, logical and moral reasoning.

1.3.3 Case Studies

My analysis further relies on two “tools” which infuse both parts of the analysis. In the doctrinal discussion, I undertake case studies of the *Korea Beef* and *Brazil Tyres* disputes. The main purpose of these case studies is to illustrate the key elements of the legal test, but also to illustrate the techniques that adjudicators use to disguise their value judgments in decisions. I argue that adjudicators have developed a suite of techniques to make value judgments against the regulator (*Korea Beef*) or in its favour (*Brazil Tyres*).

The second part contains a detailed case study of the *Seals* dispute. It includes doctrinal claims about the public morals jurisprudence and socio-legal claims about how judges make decisions. My socio-legal analysis will show that where adjudicators have strong intuitions about a matter, this can override their capacity for coherent legal reasoning.

1.3.4 Deconstruction

Textual deconstruction is a tool which involves “redescribing the commonplace in novel ways”.⁶ It is commonplace in critical theory, including legal scholarship, but its

⁶ Mark Tushnet, ‘Critical Legal Studies and Constitutional Law: An Essay in Deconstruction’ (1984) 36 Stanford Law Review 623, 629. In this sense, it shares the insight from behavioural economics that the way a matter is framed can significantly influence substantive opinions.

value is sometimes underestimated because it can appear to be a purely descriptive practice. Tushnet insists that descriptive analysis of legal cases can lead to useful insights although much legal analysis tends to bypass this step.⁷

Deconstruction requires a “close inspection” of the text to see what’s “at work below the surface”.⁸ It enables scholars to expose “unexpressed assumptions” or “various forms of arbitrariness” in a judgment.⁹ It might even reveal “apparent mendacity” by the court.¹⁰ Balkin notes that “there is generally great critical importance in discovering that a text says more than the author meant it to say, or that the logic of a text leads to an unexpected difficulty or contradiction”.¹¹ While deconstruction does not necessarily require analytical or normative claims, it can often inform such analysis (as it does in my thesis).

I will briefly signpost two examples of my use of deconstruction. First, since the *Brazil Tyres* adjudicators do not acknowledge their use of double standards to characterise the policy objective, I rely on deconstruction to identify and expose this important logical fallacy. At its core, this is not really an analytical task, as the double standard is embedded in the text, but a simple reframing or reordering of ideas can make it visible to readers who otherwise might miss it. This descriptive task lays the groundwork for subsequent analytical/normative work which would not have been possible without the deconstruction.

Second, in *Seals*, I use deconstruction to uncover the fact that adjudicators, the EU and some scholars rely on “intuitionist” arguments. Deconstruction is required because these actors do not explicitly acknowledge their reliance on intuitionism

⁷ ibid 630.

⁸ ibid 631.

⁹ ibid 646.

¹⁰ ibid 634.

¹¹ Jack Balkin, ‘Deconstructive Practice and Legal Theory’ (1987) 96 Yale Law Journal 743, 778.

(they purport to follow a welfarist view), but some of their substantive arguments can only be understood through an intuitionist lens. In this regard, the writings of Posner (an avowed intuitionist) are extremely useful for revealing which EU arguments are ultimately intuitionist in nature.

These are two prominent examples of the method, but I should note that deconstruction is an integrated part of my technique which is observable throughout the analysis.¹² Indeed, any time I suggest there is divergence between an author's stated and substantive positions, this represents a use of deconstruction.

1.4 Conceptual Framework (and Legal Theory)

This section describes my conceptual framework for analysing judicial decision-making. It briefly discusses the insights of the Realists who questioned the notion of "legal formalism" and noted the importance of "intuition" in judicial decisions.¹³

Thanks to recent advances in cognitive psychology (and the application of those insights by behavioural economists), modern legal scholars are equipped with tools to further develop the work of the Realists. If behavioural lawyers combine our analysis with insights from moral psychology, I argue that we can reach novel conclusions about how judges (and humans) engage in normative reasoning and/or the rationalisation of intuitive decisions.

¹² Indeed, I believe that all critical analysis of law relies on deconstruction to some extent.

¹³ My references to "Realism" in this section relate to American (rather than Scandinavian) Legal Realism.

1.4.1 Legal Realism

Over a century ago, the Realists considered the role of intuition and reasoning in judicial decision-making when they challenged the traditional “formalist” view of legal interpretation.¹⁴ Roscoe Pound observed:

“Legally, the judge's heart and conscience are eliminated. He is expected to force the case into the four corners of the pigeon-hole the books have provided. In practice, flesh and blood will not bow to such a theory. The face of the law may be saved by an elaborate ritual, but men, and not rules, will administer justice”.¹⁵

The Realists, many of them former judges, understood that the outcome of legal disputes was a function of human decision-making, not merely the formalist or rationalist application of an algorithm. They were not afraid to admit that judicial decisions could be based on a “hunch” or “intuition”, especially in difficult cases.¹⁶

They implied that judges might sometimes have an intuitive sense of the “desirable result”. Radin stated:

“If they [judges] act more frequently than otherwise by discovering the desirable result first... we really ought to know what test they employ in determining that a result is desirable”.¹⁷

Radin suggests that judges face a tension between their intuitions about the “desirable result” and their duty to find the legally-correct result. While the Realists convincingly argued that “legal formalism” could not explain judicial decision-making, they struggled to offer a systematic explanation of where judicial “hunches” come from or how judges know that their “intuition” is correct.

¹⁴ For example, Llewellyn challenged the formalist assumptions of judicial decision-making when he stated that “there is always more than one available correct answer” and “the court always has to select”. See Karl Llewellyn, ‘Remarks on the Theory of Appellate Decision and the Rules or Canons about How Statutes Are to Be Construed’ (1950) 3 Vanderbilt Law Review 395, 396.

¹⁵ Roscoe Pound, ‘Law in Books and Law in Action’ (1910) 44 American Law Review 12, 20.

¹⁶ See eg Joseph Hutcheson, ‘The Judgment Intuitive: The Function of the Hunch in Judicial Decision’ (1929) 14(3) Cornell Law Review 274.

¹⁷ Max Radin, ‘Theory of Judicial Decision or How Judges Think’ (1925) 11 ABA Journal 357, 359.

Frank sought to introduce a psychological perspective when he analysed “the effect on legal thinking of the assumption that judges behave like human beings”.¹⁸ He noted that “decision-making is a highly complicated psychological process” and that “there is little evidence as to how large a part the ‘rules’ play in that process”.¹⁹

Frank seemed to believe that a judge’s psychology could trump their legal training and expertise, but his approach gained little traction. No other Realists “came close to elevating the all-too-human psychological idiosyncrasies of individual judges to the status of the pivotal factor in legal decisionmaking”.²⁰ While the Realists successfully diagnosed an important problem in legal theory (the limitations of formalism), they were unable to develop “a systematic theory of the ‘social determinants’ of judicial decisions”.²¹

Frank identified two ways that psychology could shed light on and improve adjudication. First, he noted:

“Inexact as are the concepts of contemporary psychology, they... may be particularly helpful to the judge in teaching him to learn and control the worst effects of his own biases and prejudices”.²²

Frank therefore believed judges could improve their decision-making by using psychology to reduce “biases and prejudices”, provided they could figure out what their biases were. Second, he suggested that psychology could be used by observers to better predict judicial decisions (at least in theory), though he realised this tool was severely limited (in practice):

¹⁸ Jerome Frank, ‘Are Judges Human? Part One: The Effect on Legal Thinking of the Assumption That Judges Behave like Human Beings’ (1931) 80 University of Pennsylvania Law Review and American Law Register 17.

¹⁹ *ibid* 23.

²⁰ Brian Leiter, ‘Legal Realism and Legal Doctrine’ (2015) 163 University of Pennsylvania Law Review 1975, 1979.

²¹ *ibid* 1979.

²² Frank J, ‘Are Judges Human? Part Two: As Through a Glass Darkly’ (1931) 80 University of Pennsylvania Law Review and American Law Register 233, 245.

“only a rash man will assume that he ‘can put himself in the place of the judge or judges who will actually make the decision.’ Psychology is years away from telling us how to put ourselves in the mental shoes of the other fellow”.²³

In psychology (and behavioural economics), our knowledge of how humans (and economic agents) make decisions has improved drastically in recent decades. Has our knowledge advanced sufficiently (and have enough years passed) for legal scholars to put themselves in the “mental shoes” of judges without the risk of being “rash”?

One of the aims of WTO dispute settlement is to provide “security and predictability to the multilateral trading system”.²⁴ How do we measure if adjudicators are delivering on their mandate to provide “predictability”? If adjudicators are criticised for jurisprudence which is considered “unpredictable”, whose failure is this? Is a lack of predictability indicative of poor adjudication or of poor models for describing how judges make decisions?

The behavioural economists wisely decided to change their model of human decision-making rather than wait for all economic agents to start behaving perfectly rationally. *Homo economicus* was replaced by an imperfect human who makes decisions in a context of “bounded rationality”.²⁵

In law, what is our model for judicial decision-making? Are judges perfectly rational or are they human? If judges are less than fully rational, are their non-rational decisions systematic and predictable, or are they just random deviations from the

²³ Frank (n 18) 23.

²⁴ Dispute Settlement Rules: Understanding on Rules and Procedures Governing the Settlement of Disputes, Marrakesh Agreement Establishing the World Trade Organization, Annex 2 (adopted 15 April 1994, entered into force 1 January 1995) 1869 UNTS 401 (“DSU”) Art 3.2.

²⁵ Herbert Simon, *Models of Man: Social and Rational* (Wiley 1957).

rational norm? These are the types of questions which the “behavioural law” movement, inspired by behavioural economics, should seek to address.²⁶

Within behavioural law, there is a movement concerned with judicial decision-making.²⁷ This movement (which I will call “behavioural adjudication” or “judicial decision theory”) combines Frank’s questions about the psychology of judges with a modern understanding of how cognitive biases and heuristics influence human decision-making.²⁸ My thesis seeks to contribute to this nascent movement.

WTO law offers a fascinating forum to explore behavioural adjudication as WTO disputes ask difficult normative questions (about the legitimacy of moral, health and environmental measures), but adjudicators are precluded from explicitly relying on normative value judgments. This raises some interesting questions: How do they decide and how do they publicly justify their decisions?

I argue that WTO adjudicators often “disguise” the real reasons for their decisions behind a seemingly formalist, value-neutral and rational analysis. Realists would not be surprised by this type of behaviour:

“What the Legal Realists taught us is that too often the doctrine that courts invoke is not really the normative standard upon which they really rely”.²⁹

The Realists decried this phenomenon and argued for a reconciliation of the real reasons with the officially stated ones:

²⁶ For a comprehensive discussion of behavioural law, see Eyal Zamir and Doron Teichman (eds), *The Oxford Handbook of Behavioral Economics and the Law* (OUP, 2014).

²⁷ For a general discussion of the application of “behavioural” methods to normative reasoning and judicial decision-making, see Eyal Zamir and Doron Teichman, ‘Judicial Decision-Making: A Behavioral Perspective’ in Zamir and Teichman (*ibid*) and Cass Sunstein, ‘Moral Heuristics and Moral Framing Lecture’ (2003) 88 Minnesota Law Review 1556.

²⁸ It seems that Frank was fifty years ahead of his time. If he had been writing in the 1970s and 1980s, perhaps behavioural adjudication would have taken off as quickly as behavioural economics did.

²⁹ Leiter (n 20) 1975.

“It was central to Legal Realism to reform the law to make the actual doctrine cited by courts and treatise writers correspond to the actual normative standards upon which judges rely”.³⁰

If it is true that WTO adjudicators refrain from providing the real reasons for their decisions, are there “behavioural” tools which scholars can use to demonstrate that the real reasons are hidden or, even more ambitiously, to reveal what these real reasons are? The next section will outline my framework for addressing these questions, including a deeper discussion of behavioural economics and law.

1.4.2 Cognitive Psychology and Behavioural Economics

The discipline of psychology has made enormous strides in expanding our understanding of human decision-making. Kahneman was awarded a Nobel Prize for his work alongside Tversky³¹ on “human judgment and decision-making under uncertainty”, including how “human decisions may systematically depart from those predicted by standard economic theory” and “how human judgment may take heuristic shortcuts that systematically depart from basic principles of probability”.³²

Kahneman’s work is based on the “dual process theory” of human cognition.

According to this model:

“System 1 operates automatically and quickly, with little or no effort and no sense of voluntary control”.³³

“System 2 allocates attention to the effortful mental activities that demand it, including complex computations. The operations of System 2 are often associated with the subjective experience of agency, choice, and concentration”.³⁴

³⁰ ibid 1975.

³¹ Kahneman has stated that if Tversky had still been alive, he no doubt would have shared the Nobel Prize (which cannot be awarded posthumously).

³² Nobel Prize, ‘Press Release’ (9 October 2002) <<https://www.nobelprize.org/prizes/economic-sciences/2002/press-release/>> (consulted 3 February 2020).

³³ Daniel Kahneman, *Thinking Fast and Slow* (Farrar, Straus and Giroux, 2011) 20.

³⁴ ibid 21.

System 1 enables humans to quickly and efficiently make decisions in situations of uncertainty and incomplete knowledge. We do this by adopting heuristics which Sunstein describes as “rules of thumb, that work well most of the time, but that also systematically misfire”.³⁵

Kahneman cites the following famous example of System 1 misfiring:

“A bat and ball cost \$1.10. The bat costs one dollar more than the ball. How much does the ball cost?”³⁶

Kahneman tells the reader: “A number came to your mind. The number, of course, is 10: 10c”.³⁷ This example is designed to show how easily our System 1 can misfire by producing an answer which is “intuitive, appealing, and wrong”.³⁸ Kahneman notes:

“More than 50% of students at Harvard, MIT, and Princeton gave the intuitive—incorrect—answer”.³⁹

Kahneman is not critical of the fact that System 1 has misfired. After all, we cannot control our intuitive reaction (“the intuitive answer also came to the mind of those who ended up with the correct number—they somehow managed to resist the intuition”).⁴⁰ However, he does seem perturbed that “System 2 endorsed an intuitive answer that it could have rejected with a small investment of effort”.⁴¹

Humans need not always be slaves to the intuitive misfiring of System 1. Where we have the time and expertise to engage System 2, we can overrule certain mistakes even if they are “intuitive” and “appealing”⁴². If an expert mathematician (or even a

³⁵ Sunstein (n 27) 1558.

³⁶ Kahneman (n 33) 44.

³⁷ ibid.

³⁸ ibid. The correct answer is 5c.

³⁹ ibid 45.

⁴⁰ ibid.

⁴¹ ibid 44.

⁴² ibid.

mediocre one), was asked to reframe the “bat and ball” problem as an algebraic equation, and to write down each step in their reasoning, they would produce the correct answer. If we create the right decision-making structures, this can force us to use System 2 and avoid answers which are “intuitive, appealing, and wrong”.

Judicial decision-making seems different from other realms of human decision-making as it is built on structures which require the use of System 2. Judges are experts who must frame the issue under dispute in order to apply pre-established rules and reasoning and ultimately reach a decision. That decision must be justified by written reasons explaining how the judges arrived at their conclusion. Judges are equipped with the time and fact-finding resources (not to mention a professional obligation) to use System 2 to resolve the problem before them. Indeed, the whole judicial ritual appears designed to prevent System 1 from hijacking the decision-making process.

However, as behavioural economists have discovered, System 1 is a powerful force which can override the human capacity to think rationally. This is why behavioural economists have used insights from psychology to challenge the “rationality” assumptions in standard economic theory.⁴³ They have shown that heuristics and cognitive biases can systematically lead to non-rational outcomes.⁴⁴

In law, formalist legal theory resembles standard economic theory insofar as it assumes rational decision-making. The behavioural law movement is using insights from psychology to challenge this assumption, though we have made far less

⁴³ In my socio-legal analysis, I will focus on the “rationality” assumption which traditional legal theory (such as formalism) shares with standard economic theory. Behavioural economists also challenge the assumption that economic agents are “self-interested”, though I will not address this assumption as it seems less relevant to law. Traditional legal theory is based on an alternative assumption that judges are impartial and leave their personal interests outside the courtroom.

⁴⁴ I will introduce some of the leading biases, such as “loss aversion” and the “endowment effect” in the analysis of *Brazil Tyres*.

progress than behavioural economics. Indeed, there are some types of issue where behavioural law is better placed than behavioural economics to achieve novel insights, especially on matters related to normative reasoning.

Much analysis in psychology and behavioural economics focusses on intuitive judgments in situations where the correct answer can be verified (like the “bat and ball” problem). On such matters, it is possible to “demonstrate an error of fact or logic” and ultimately show that an intuitive judgment was wrong.⁴⁵

Psychologists have also studied intuitive “moral” judgments, such as when people consider it “fair” to lower an employee’s salary or raise product prices.⁴⁶ Behavioural economists have shown that people often make intuitive “moral” judgments which would be inconsistent with rationalist economic theory.

For example, even if nominal prices remain the same in both contexts, people might consider credit card surcharges unfair, yet consider cash discounts fair.⁴⁷ This suggests that intuitive judgments can vary depending on how the question is framed. System 2 judgments, on the other hand, are based on a substantive analysis and should not be affected by framing.

On normative questions, Sunstein recognises that “it is more difficult to produce a widely shared standard by which to test the question of mistake”.⁴⁸ While we can determine that it is not rational, can we really call it a “moral error” to be outraged by credit card surcharges but not cash discounts? Sunstein believes that moral heuristics “play a pervasive role in moral, political, and legal judgments” and that

⁴⁵ Sunstein (n 27) 1567.

⁴⁶ See eg Kahneman (n 33) 305-309 where he discusses “Loss Aversion in the Law”.

⁴⁷ Daniel Kahneman, Jack Knetsch and Richard Thaler, ‘Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias’ (1991) 5(1) *The Journal of Economic Perspectives* 193, 204.

⁴⁸ Sunstein (n 33) 1567.

“they sometimes produce significant mistakes”.⁴⁹ He seeks to catalogue some of the ways that moral heuristics have been captured in legal doctrine (and how they can misfire).⁵⁰

For example, Sunstein notes the following moral heuristic which generally works well: “people should not be permitted to engage in moral wrongdoing for a fee”.⁵¹ He notes that “there are no tradable licenses for rape, theft, or battery” because “the appropriate level of these forms of wrongdoing is zero”.⁵²

However, the heuristic can also misfire when people intuitively apply it to all situations. Sunstein argues that “pollution is an altogether different matter” as “some level of pollution is a byproduct of desirable social activities and products, including automobiles and power plants”.⁵³ Our moral heuristic suggests that people should not be allowed to pay to pollute, but our System 2 reasoning is capable of overruling this intuitive judgment.

My focus is on the influence of intuition and heuristics on legal and moral reasoning. Is it possible for a function which invariably belongs to System 2 (reasoning) to be hijacked by System 1? When considering questions about moral and legal reasoning, it is useful to consider the literature from moral psychology.

1.4.3 Moral Psychology and Normative Reasoning

Moral psychologists also rely on “dual process” theory to study the interplay between “intuitive” moral judgments (under System 1) and “rationalist” moral reasoning (under

⁴⁹ Cass Sunstein, ‘Moral heuristics’ (2005) 28 Behavioural and Brain Sciences 531, 531.

⁵⁰ ibid 535-541.

⁵¹ ibid 537.

⁵² ibid.

⁵³ ibid.

System 2).⁵⁴ Haidt's analysis discusses moral reasoning in ways which are highly-analogous to legal reasoning.⁵⁵

The "rationalist" approach would consider that "moral knowledge and moral judgment are reached primarily by a process of reasoning and reflection".⁵⁶ Under this approach, "the reasoner searches for relevant evidence, weighs evidence, coordinates evidence with theories, and reaches a decision".⁵⁷ In other words, the "reasoner" does not know what their moral conclusion is until after a conscious and effortful reasoning process. This resembles the rationalist model of judicial decision-making.

Alternatively, if the decision-maker engages in "one-step mental processes" where answers arise instantaneously as "sudden flashes of insight" or "gut reactions", they are not operating as a "reasoner".⁵⁸ This is the realm of intuitionism where "moral intuitions (including moral emotions) come first and directly cause moral judgments".⁵⁹ Moral decisions can be based on the fact that a person "feels a quick

⁵⁴ See eg Jonathan Haidt, 'The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment' (2001) 108(4) *Psychological Review* 814. Haidt writes from the perspective of moral psychology, not moral philosophy, therefore his findings are descriptive, not normative. He notes: intuitionism "is a descriptive claim, about how moral judgments are actually made. It is not a normative or prescriptive claim, about how moral judgments ought to be made" (at 815).

⁵⁵ Legal scholars would endorse the close links between legal and moral/philosophical reasoning. See eg Richard Posner, 'Animal Rights: Legal, Philosophical, and Pragmatic Perspectives' in Cass Sunstein and Martha Nussbaum (eds), *Animal Rights: Current Debates and New Directions* (OUP 2006). Posner notes that the "legal-formalist approach" has "distinct affinities with philosophical analysis" (at 51) and he further describes legal rights as the "cousin of deontological moral philosophy" (at 55).

⁵⁶ Haidt (n 54) 814.

⁵⁷ *ibid* 818.

⁵⁸ *ibid* (citing Kathleen Galotti).

⁵⁹ *ibid* 814.

flash of revulsion” or “knows intuitively that something is wrong”.⁶⁰ This resembles the “intuitions” and “hunches” invoked by the Realists.⁶¹

It might seem intuitively-correct, under System 1, that credit card surcharges are unfair, while cash discounts are fair. System 2, which focusses on the substance rather than the framing, would not endorse such a conclusion. If the nominal amounts were the same, System 2 would have the same normative view of credit card surcharges and cash discounts.

Haidt recognises that humans usually rely on a hybrid decision-making process which combines both intuitive and rationalist elements. For example, a decision-maker might make an intuitive moral judgment, but also be able to articulate seemingly cogent reasons for their decision. Haidt describes this as post-hoc rationalisation where a decision-maker produces reasoning “after a moral judgment is made” in order to “support an already-made judgment”.⁶² Because this type of reasoning seeks to justify a pre-ordained conclusion, Haidt compares it to “a lawyer trying to build a case”.⁶³

Post hoc rationalisation is designed to influence other people or respond to “a social demand for a verbal justification”,⁶⁴ but rationalisers might even fool themselves into believing their stated reasons were decisive or truly motivated their decision. After

⁶⁰ ibid.

⁶¹ In the *Seals* case study, I will consider both a normative/rationalist defence of the EU measure (in Chapter 9 where I discuss animal welfarism and Peter Singer) and a descriptive/intuitionist defence (in Chapter 10 where I consider intuitionism and Richard Posner).

⁶² Haidt (n 54) 818.

⁶³ ibid 814.

⁶⁴ ibid.

all, “intuition occurs quickly, effortlessly, and automatically, such that the outcome but not the process is accessible to consciousness.”⁶⁵

In certain cases, a decision-maker’s post-hoc search for justifications may fail to identify any credible reasons, a phenomenon known as “moral dumbfounding”⁶⁶:

“In the social intuitionist model it becomes plausible to say, ‘I don’t know, I can’t explain it, I just know it’s wrong’.”⁶⁷

The absence of cogent reasons does not necessarily lead the decision-maker to reassess their moral conclusion. Rather, they might hold firm to their moral intuition and somehow accept the fact that they cannot think of compelling reasons for holding it. In such cases, intuition trumps reason.

Haidt’s model suggests that where System 1 and System 2 operate simultaneously, intuition drives the moral judgment while moral reasoning merely provides the post hoc rationalisation. Haidt is generally sceptical of “the causality of reasoning in moral judgment” and believes that most judgments are driven by intuition rather than reasoning.

However, Haidt recognises a narrow set of contexts where reasoning can trump intuition:

“The reasoning process in moral judgment may be capable of working objectively under very limited circumstances: when the person has adequate time and processing capacity, a motivation to be accurate, no a priori judgment to defend or justify, and when no relatedness or coherence motivations are triggered.”⁶⁸

This reasoning ability is rare and:

⁶⁵ ibid 818.

⁶⁶ Sunstein (n 49) 533 (citing an unpublished manuscript by Haidt et al).

⁶⁷ Haidt (n 54) 814.

⁶⁸ ibid 822.

“may be common only among philosophers, who have been extensively trained and socialized to follow reasoning even to very disturbing conclusions (as in the case of Socrates or the more recent work of Peter Singer)”.⁶⁹

When moral reasoning has a causal effect and actually drives the moral decision, Haidt describes this as like “a judge searching for the truth”.⁷⁰ At first glance, the analogy seems apt. Judges meet his conditions for moral reasoning to have a causal effect as they have “adequate time”, “processing capacity”, “a motivation to be accurate” and “no a priori judgment”.

Haidt suggests that where there are widespread moral intuitions on a matter (as is the case with animal welfare), judges are among a select group of stakeholders (alongside philosophers) who remain capable of genuine moral reasoning. They therefore have a critical social function due to adverse consequences which can flow from failed reasoning. Haidt argues that governments and judges should rely on System 2 for important decisions as:

“moral intuitions often bring about non-optimal or even disastrous consequences in matters of public policy, public health, and the tort system. A correct understanding of the intuitive basis of moral judgment may therefore be useful in helping decision makers avoid mistakes”.⁷¹

Haidt suggests that judges (and philosophers) ultimately rely on System 2 “reasoning” to make normative decisions. Is this a reasonable assumption? Is Haidt seeking to describe how judges actually behave or is he merely invoking the rationalist stereotype of how they should behave?

When judges have a strong moral intuition, but their reasoning leads to a different conclusion, which should they follow? The Realists seemed to believe that judges often follow their “intuition” to achieve the “desirable result”. In my analysis, I

⁶⁹ ibid 829.

⁷⁰ ibid 814.

⁷¹ ibid 815. This echoes Frank’s call to overcome “biases and prejudices”.

summarise Haidt's framework with the following three approaches to moral decision-making⁷²:

Decision-making Approach	Real Reasons	Stated Reasons	Decision-making model
Moral Intuition	Intuition	Intuition	Moral dumbfounding
Post-Hoc Rationalisation	Intuition	Reasoning	Most human decisions, including lawyers
Moral Reasoning	Reasoning	Reasoning	Judges and philosophers

How can we determine which model actually describes judicial decision-making?

Since judges must offer detailed reasons for their decisions (and cannot admit to relying on mere intuition), we can eliminate “moral dumbfounding”. However, when we see a judge’s stated reasons for a decision, how can we tell if this is post-hoc rationalisation or true reasoning?

The Realists claimed that a judge’s stated reasons might not always match the real reasons and thus recognised the possibility of post-hoc rationalisation. Posner also recognises this phenomenon when he states that judges sometimes seek “a professionally respectable ground for rationalizing” decisions where they were actually “persuaded on other grounds”.⁷³

Is it possible to distinguish between intuitionist and rationalist decisions merely by looking at judicial reasoning for hidden clues? In the *Seals* case study, I will argue that it is possible to distinguish rationalisation from moral reasoning, at least in some cases.

⁷² In theory, there is a fourth approach where a decision is based on moral reasoning, but is publicly defended as intuitionist. I have yet to encounter this approach in judicial reasoning and therefore refrain from exploring it in detail.

⁷³ Posner (n 55) 58.

Sunstein recognises that “a great deal of progress remains to be made” on the law and moral heuristics and he invites further scholarship⁷⁴:

“One of my primary hopes is to help stimulate further research, testing when and whether people use moral heuristics that produce sense or nonsense in particular cases”.⁷⁵

I believe that legal questions about animal ethics provide an ideal site for testing such questions. As Herzog notes “ethical thinking about animals is rife with inconsistency and paradox”.⁷⁶ He further suggests:

“Sunstein’s heuristics may be illuminated even more when applied to understanding contradictions in how we think about animals than it is to human-focused moral quandaries”.⁷⁷

I will seek to show that the *Seals* adjudicators’ “reasoning”, on key animal issues, was based on System 1 intuitions, rather than legal, logical or moral reasoning from System 2. I will demonstrate this by showing that adjudicators relied on a range of logical fallacies and contradictions which were inconsistent with WTO law and basic principles of legal reasoning.

1.5 My Analysis and the Broader Literature

This section briefly describes how my analysis contributes to the existing literature on WTO doctrinal law and judicial decision-making.⁷⁸

1.5.1 WTO Doctrinal Law

My analysis contributes to the scholarly debate on WTO doctrinal law, especially with respect to the necessity test and public morals jurisprudence. There has been

⁷⁴ Sunstein (n 49) 532.

⁷⁵ *ibid.*

⁷⁶ Harold Herzog and Gordon Burghardt, ‘The Next Frontier: Moral Heuristics and the Treatment of Animals’ (2005) 28 Behavioural and Brain Sciences 554.

⁷⁷ *ibid.*

⁷⁸ This section refers to the existing literature in general terms only. Each of the matters which I briefly describe here will be explored in full (with citations to leading authors) in the relevant substantive chapters.

extensive academic debate about where adjudicators should assess the right to regulate (Article III or Article XX) and how it should determine necessity (AM or W&B testing).

However, much of the analysis overlooks the fact that adjudicators have been given an impossible task. In the vast majority of cases, it is impossible to determine if a social measure can justify a trade restriction without making a normative value judgment. While most academic scholarship (rightly) criticises the Appellate Body for reasoning which is incoherent or contradictory, these critiques tend to overlook the fact that WTO adjudicators do not necessarily have a better option. By recognising this underlying tension, I address a different question than other scholars. Rather than inquiring whether adjudicators should make value judgments, I seek to expose the techniques they use to surreptitiously embed their value judgments in their reasoning.

I also seek to contribute to the public morals debate. The current literature contains much discussion about (i) whether the moral exception is different to other exceptions; (ii) what makes moral measures different, and (iii) whether moral measures require a different jurisprudential approach. In my view, the literature has not yet provided clear answers to these questions. I will argue that there is a class of moral measures which appear to be different (as they are based on social intuitions rather than reasoning) and which may require unique jurisprudential approaches (such as evidentiary unilateralism).

Finally, I will consider the role played by moral philosophy in Article XX(a) disputes. While some authors have alluded to the utility of philosophical arguments in public

morals disputes, it has not been fully explored, especially considering the important role played by consequentialist moral philosophy in the *Seals* dispute.

In my view, where a regulator invokes moral philosophy to defend their claim, adjudicators should take advantage of the moral philosophy literature as a shortcut for recognising which arguments are coherent and which ones are implausible. Indeed, as long as adjudicators insist on a “moral standards” requirement as part of their public morals jurisprudence, this would suggest that only rationalist measures can be defended under Article XX(a). The “moral standards” requirement seems to rule out intuitionist defences, while opening the door to consequentialist and deontological moral arguments.

1.5.2 Judicial Decision-Making

There is an expansive literature on decision theory, cognitive psychology and behavioural economics which discusses how humans make decisions. This literature demonstrates that people do not always behave rationally, especially when they rely on judgments (or heuristics) from System 1 rather than reasoning from System 2.

Psychologists and behavioural economists have famously demonstrated that many incorrect or non-rational human decisions are not merely random, but can be attributed to systematic biases. Some legal scholars have used similar methods to show that judgment calls in a judicial context may also be systematically influenced by non-rational factors. For example, the size of a jury award or the choice between a murder or manslaughter finding, may be influenced by cognitive biases.

Psychology and behavioural economics have shown that System 1 judgments can systematically lead humans astray, but they tell us little about the reasoning behind

those decisions (apart from the fact that the decision-maker has chosen not to engage their System 2 reasoning). If we wish to understand how Systems 1 and 2 interact when humans seek to engage in normative reasoning, we must look beyond cognitive psychology and behavioural economics. On questions related to moral and legal reasoning, we should look to the disciplines of moral psychology and behavioural law. In my analysis, I will seek to show that seemingly rational reasoning can actually be shown to be based on intuition and systematic bias.

While it is common for legal scholars to identify “double standards” when critiquing a judgment, I believe there is value in equating these “double standards” with the “anomalies” which are found in behavioural economics. While the identification of double standards is a valuable contribution by legal scholars, we can follow the lead of behavioural economics and seek to develop methods to show if they are random or if they reflect systematic bias. I believe this area of legal scholarship is still nascent.

The most novel part of my analysis is the attempt to reverse engineer a reasoned judicial decision to see if it is based on true moral reasoning or mere post-hoc rationalisations. With the help of moral psychology, I believe that legal scholarship has the potential to produce insights about how humans “reason” which could be as significant as the work from behavioural economists about how people make judgments.

Admittedly, humans spend much more time deciding than reasoning (so our sample size is smaller) and reasons are harder to isolate than decisions (especially in a lab). However, if reason is the key to moral progress, as eloquently argued by Shermer,⁷⁹

⁷⁹ Michael Shermer, *The Moral Arc* (Henry Holt 2015).

it is important that scholars of normative reasoning (including scholars of law and ethics) take advantage of dual-process theory and moral psychology to hone our ability to distinguish between true reasoning and mere rationalisation. It seems that judges frequently engage in the latter. It is up to scholars of behavioural adjudication to study and expose this tendency.

1.6 Chapter Summary

Chapter 1 introduces my research question, argument, methods, conceptual framework and how this thesis contributes to existing literature.

Chapter 2 will provide broader context about the rationale for the GATT/WTO and how dispute settlement has evolved over the years from a diplomatic to an increasingly judicial approach. I discuss the competing views about the “object and purpose” of GATT/WTO, including the need to curtail protectionism while respecting the regulatory autonomy of Members. This underlying tension between protectionism and regulatory autonomy underpins much judicial and scholarly analysis of WTO law.

I explain the widely-held (but rejected) view that regulatory autonomy for domestic measures should be addressed under the national treatment rule (Article III) and how adjudicators have ultimately decided to situate this analysis under the general exceptions (Article XX). This chapter also addresses broader issues about legal interpretation and institutional legitimacy at the WTO.

Chapter 3 includes a dedicated analysis of Article XX. It will describe all aspects of the test, including a brief history of how the necessity test has evolved over time.

Chapter 4 introduces the two different approaches to necessity testing under Article XX: the weighing and balancing (W&B) test and the alternative measures (AM) test. It offers a detailed analysis of the jurisprudence, including some aspects of the legal tests which are disputed or lack clarity.

Chapter 5 provides a theoretical analysis of the necessity test. It will show that the AM test is often viewed as a search for Pareto efficiency, while the W&B test is viewed as a more intrusive search for Kaldor-Hicks efficiency (a type of cost-benefit analysis). The W&B test has sparked significant controversy because of its potential to make the regulator worse off.

Chapter 6 contains a case study of *Korea Beef*, the dispute which first introduced W&B testing and opened the debate about whether Kaldor-Hicks efficiency testing is appropriate at the WTO. The dispute was ostensibly decided under the AM test, but I will argue that adjudicators adopted techniques which enabled them to engage in “disguised balancing”.

Chapter 7 addresses *Brazil Tyres* where adjudicators ruled that an import ban on used tyres could be justified on public health grounds. My analysis will show that this finding in favour of Brazil relied extensively on “double standards”, especially with respect to the characterisation of the policy objective. I will argue that, instead of clearly defining Brazil’s policy objective, adjudicators used a number of different characterisations for different legal elements. My analysis will use “deconstruction” to make visible these double standards and I will introduce “behavioural” techniques to demonstrate the presence of bias by adjudicators.

The *Seals* case study, covering chapters 8-12, forms the core of my socio-legal analysis.⁸⁰ It starts with a doctrinal analysis of the public morals jurisprudence (chapter 8). This analysis introduces the Article XX(a) legal test, including the requirement that such measures be based on “moral standards”. It inquires whether these standards can be met purely by reference to a society’s own values (unilateralism) or whether a regulator must also provide external evidence to prove that a measure is legitimate.⁸¹ I further consider whether the moral exception is different from the health/environmental exceptions and, if so, what makes it different.

Chapters 9-10 build on *Brazil Tyres* by showing that adjudicators adopted an equivocal definition of the EU’s policy objective. Chapter 9 will describe the first meaning of “seal welfare” where the seal products ban is understood as an “animal welfare” measure motivated by the “welfarist” movement of Singer and Bentham. I will introduce the Suffering/Utility Standard which is a consequentialist reasoning tool for establishing when animal exploitation is ethical.

I further introduce a second key welfarist principle: the proscription against “speciesism”. Singer argues that any discrimination between different animals should be based on morally-relevant criteria, rather than species membership. In this sense, welfarism contains a “consistency requirement”.⁸²

Chapter 10 establishes the second conception of “seal welfare”: the “intuitionist” view that seals deserve to be exempted from exploitation on the grounds that they are special and people care more about them. I will draw on the literature of Posner, who has publicly defended intuitionism as a basis for offering total protection to

⁸⁰ Although chapter 8 is primarily a doctrinal chapter focussed on the public morals jurisprudence.

⁸¹ For example, a regulator might provide evidence from other states, international law or moral philosophy.

⁸² For a discussion of consistency and speciesism, see Peter Singer, *Animal Liberation* (3rd edition, HarperCollins 2002); James Rachels, ‘Drawing Lines’ in Cass Sunstein and Martha Nussbaum (eds), *Animal Rights: Current Debates and New Directions* (OUP 2006).

charismatic species. This intuitionist approach was ultimately relied on (without any explicit acknowledgment) by adjudicators, the EU and a number of scholars.

Intuitionism differs significantly from welfarism as it is not rationalist and therefore does not rely on moral, legal or logical reasoning. In addition to Singer's philosophical objection to intuitionism, a number of legal scholars have opposed the legitimacy of intuitionism under WTO law, including Perišin and Sellheim.⁸³

Chapter 11 will highlight that the welfarist and intuitionist claims are mutually-exclusive approaches to animal regulation. Despite this fact, adjudicators combined them into a hybrid claim which relied on logically contradictory views on key issues, such as the role of risk, science, sentience and standards.

Chapter 12 addresses the question of consistency. Welfarist philosophers, such as Singer, argue that "all animals are equal" and that they must be treated consistently.⁸⁴ Adjudicators, however, seem to lack a clear view on whether Article XX(a) contains a consistency requirement. They adopt three different approaches to addressing the question of consistency: (i) they sidestep the matter; (ii) they hint that there is no consistency requirement (using unconvincing reasons); and (iii) they undertake a consistency analysis in favour of the EU.

Ultimately, the *Seals* adjudicators were able to hold contradictory views on whether the seal products ban was welfarist or intuitionist. They were also able to hold contradictory views on matters of jurisprudence, such as whether Article XX(a) contains a consistency requirement.

⁸³ Their views will be described in detail in chapter 9.

⁸⁴ Singer (n 82).

Chapter 13 concludes by considering why it is problematic for judges to rationalise decisions which were actually reached on the basis of intuition. Since judges are one of very few stakeholders who possess the mandate and resources to engage in System 2 reasoning, they have a critical social role. Where judges claim that questionable social practices are based on moral standards, they act as barriers to moral progress.

CHAPTER 2 - RATIONALE FOR THE GATT/WTO: BALANCING REGULATORY AUTONOMY AND PROTECTIONISM

2.1 Rationale for Creation of the GATT/WTO

This chapter will frame the tension between trade and non-trade values at the WTO.

It will explain the competing reasons for the creation of the WTO, how it seeks to discipline protectionism while respecting its Members' regulatory autonomy and some different judicial approaches for assessing the legitimacy of trade-restrictive measures.

Jackson notes there are “divergent alternative views about the policy goals of the [GATT/WTO] system”.¹ From a narrow trade perspective, it seeks to enhance welfare by removing trade distortions. However, at a broader level, a functioning global market relies on “predictability and stability”, an “institutional framework”, “rules” and “a dispute settlement system of some kind”.² The WTO provides this regulatory infrastructure to ensure the efficient functioning of the multilateral trading system.

The WTO’s goals are arguably broader than trade. It was established alongside the other Bretton Woods institutions for economic governance (the World Bank and IMF) and sought to complement the UN’s stewardship of political and security governance arrangements. Baldwin argues that the objectives of the GATT drafters was “international political stability” more so than maximizing “economic welfare”.³

¹ John Jackson, ‘Power and Rules in the Changing Economic Order: The Case of the World Trade Organization’ (2008) 84(3) International Affairs 437, 444.

² ibid 438.

³ Richard Baldwin cited in Douglas Irwin, Petros Mavroidis and Alan Sykes, *The Genesis of the GATT* (CUP 2008) 188.

Ruggie describes the post-WWII system of international economic governance as “embedded liberalism”.⁴ While the WTO facilitated *free* trade (as well as transparency and non-discrimination), it certainly did not pursue unbridled *free* trade.⁵ Governments rejected “unimpeded multilateralism”⁶ and sought to design a system which left them sufficient autonomy to manage “socially disruptive domestic adjustment costs”.⁷

Elsig, Hoekman and Pauwelyn seek to reconcile these competing objectives through a cross disciplinary analysis of the WTO’s *raison d’être*.⁸ They narrow the list down to 13 key objectives for the multilateral trading system, including multiple objectives from law⁹, economics¹⁰ and international relations.¹¹ Some important non-trade goals include “protection of the environment”, “promotion of human rights” and the “reduction of poverty not only at home but abroad”.¹² For interpreters of the WTO Agreement, who seek to know its “object and purpose”, the academic literature is unable to offer a single, comprehensive answer.

The next section will focus on two leading theories which seek to explain why governments enter into trade agreements: the “standard model” and the

⁴ John Ruggie, ‘International Regimes, Transactions, and Change: Embedded Liberalism in the Postwar Economic Order’ (1982) 36(2) International Organization 379.

⁵ In fact, even after eight rounds of continuous trade liberalisation, the Marrakesh Agreement claims to pursue the “substantial reduction” of tariffs and trade barriers rather than their elimination. Marrakesh Agreement Establishing the World Trade Organization (adopted 15 April 1994, entered into force 1 January 1995) 1867 UNTS 154 (“Marrakesh Agreement” or “WTO Agreement”), Preamble (3rd recital).

⁶ Ruggie (n 4) 393.

⁷ *ibid* 399.

⁸ Manfred Elsig, Bernard Hoekman and Joost Pauwelyn, ‘Thinking about the performance of the World Trade Organization: A discussion across disciplines’ (Robert Schuman Centre for Advanced Studies 2016, EUI Working Paper RSCAS 2016/13).

⁹ *ibid* 11 where the authors consider “compliance with WTO judgements” as a measure of success of the WTO.

¹⁰ *ibid* 7 where the authors note the main economic aims are to reduce trade barriers and manage political-economy pressures.

¹¹ *ibid* 3 where the authors highlight “macro” goals such as “rule of law”, acting as a buffer in “times of distress” and integrating new Members such as “powerful states” and LDCs.

¹² Jackson (n 1) 438.

“practitioners’ story”.¹³ Each theory agrees that trade agreements can increase global welfare, but they disagree about the significance of “protectionism”. The standard model suggests that unilateral protectionism can be rational and welfare-enhancing at the national level. The practitioners’ story views protectionism as a victory for vested interests at the expense of national welfare.

2.1.1 The Standard Model

The standard model is based on the idea that trade agreements “address the negative international externalities that would result if governments were to set their trade policies unilaterally”.¹⁴ This is because unilateral trade policy places “little or no weight on the well-being of foreigners” and therefore fails to consider “spillover effects” in other markets.¹⁵

However, the standard model makes a further claim that, in the absence of trade agreements, tariffs and other forms of protection are often “rational” and welfare-enhancing at the national level. This is a departure from classic economic theory which views tariffs as welfare-reducing for the imposing country.¹⁶ Krugman argues against the standard model (and supports the classic view) when he states that “a country serves its own interests by pursuing free trade regardless of what other

¹³ These terms (“practitioners’ story” and “standard model”) appear in Don Regan, ‘Explaining Trade Agreements: The Practitioners’ Story and the Standard Model’ (Robert Schuman Centre for Advanced Studies 2014, EUI Working Paper RSCAS 2014/113). Regan’s views will be discussed further below.

¹⁴ Gene Grossman and Henrik Horn, ‘Why the WTO? An Introduction to the Economics of Trade Agreements’ (Research Institute of Industrial Economics 2012, Stockholm, IFN Working Paper No 916) 5.

¹⁵ ibid 6-9.

¹⁶ Ricardo famously opposed Britain’s “Corn Laws” and served as intellectual inspiration for their eventual abolition in an act of unilateral liberalisation in 1846.

countries may do".¹⁷ In his view, the "national interest" should drive governments to pursue unilateral liberalisation.¹⁸

The standard model has several incarnations. The traditional version (the Bagwell-Staiger model) builds on the notion that, under certain conditions, tariffs can actually increase national income by distorting the terms-of-trade in favour of the tariff-imposing country.¹⁹ These are known as "optimum tariffs" and early versions of the standard model suggested unilateral tariffs were rational based on an assumption that all tariffs are "optimum".²⁰ Krugman is highly critical of this assumption and asserts that "the optimum tariff argument plays almost no role in real-world trade disputes".²¹

An alternative variant of the standard model uses a "social welfare function" to weight the preferences of different national actors (including non-economic preferences) to determine which policies are welfare-enhancing.²²

Grossman and Helpman further introduce a model based on political-economy factors which seeks to explain why tariff protection can be rational for the political decision-makers even where it reduces national income. Under this political-economy variant, welfare-reducing tariffs might sometimes be rational for a government which is more interested in its electoral prospects than "the aggregate

¹⁷ Paul Krugman, 'What Should Trade Negotiators Negotiate About?' (1997) 35(1) Journal of Economic Literature 113.

¹⁸ ibid.

¹⁹ For a general discussion of optimum tariffs and the Bagwell-Staiger model, see Grossman and Horn (n 14) 31-45. The standard model is also referred to as "terms-of-trade theory" or the "national market power model".

²⁰ ibid 19 where the authors note their analytical assumption that the "government always has a unilateral incentive to impose some positive tariff".

²¹ Krugman (n 17) 113.

²² Grossman and Horn (n 14) 12.

and socially-just welfare of society".²³ Governments may pursue protectionist policies for the purpose of maximising campaign contributions, keeping influential groups on-side, favouring electorates with small margins or favouring industries which are concentrated in swing states.²⁴

While this may raise moral questions about how political actors should behave, Grossman and Horn suggest that political-economy considerations clearly influence politicians who are not driven by "ethically-defensible preferences that they might have held in some best-of-all-possible worlds".²⁵ Even though the political-economy model is descriptive, they make the further claim that it can play a useful role in the "interpretation of existing agreements".²⁶

In summary, the standard model suggest that the unilateral use of tariffs, in the absence of international cooperation through a trade agreement, is rational. Depending on the variant, protection may be rational because it increases national income, improves social welfare or is politically-expedient for those in power.

2.1.2 Practitioners' Story

The "practitioners' story" agrees with certain variants of the standard model that protectionism is globally inefficient (in terms of income) and should be reduced through trade agreements such as the GATT. Rather than viewing protectionism as rational or welfare-enhancing at the national level, the practitioners' story sees it as a failure of domestic politics.

Krugman criticises the standard model's search for rationality to justify protectionism.

He believes economists should relinquish rational explanations of trade policy:

²³ ibid 12-13.

²⁴ ibid 13.

²⁵ ibid 29.

²⁶ ibid.

“The economist who wants to influence [trade] policy, as opposed to merely jeering at its foolishness, must not forget that the economic theory underlying trade negotiations is nonsense - but he must also be willing to think as the negotiators think, accepting for the sake of argument their view of the world”.²⁷

Krugman believes that the “mercantilist” approach of negotiators “does not make sense on any level” and therefore he does not seek to reconcile it with standard economic theory.²⁸ He concludes that governments often pursue an irrational trade policy and that trade agreements are designed “to protect us from ourselves”.²⁹

Criticism of the standard model also comes from trade lawyers, such as Regan, who refers to “the near impossibility of actually believing [it]”.³⁰ He agrees with Krugman that “practitioners do not regard terms-of-trade manipulation as a significant phenomenon in the real world”.³¹ Protectionism is irrational and welfare-reducing, but governments sometimes resort to it as a politically-expedient device.

Protectionism is ultimately a failure of the political process, rather than a rational expression of social choices. As Bhagwati notes, protectionist policies “reflect lobbying rather than social advantage”.³²

Hudec decries the oversized influence of “local economic interests which stand to reap immediate gain from trade protection” at the expense of “national export interests” and “foreign interests that have no representation at all in the process”.³³

Like the GATT drafters, Hudec is acutely aware of the infamous Smoot-Hawley tariffs based on uncontrolled “log-rolling”:

²⁷ Krugman (n 17) 114.

²⁸ ibid. Krugman further notes that if standard economic theory were correct, trade treaties would be unnecessary as governments would pursue “unilateral” free trade (at 113).

²⁹ ibid 118.

³⁰ Regan (n 13) 8.

³¹ ibid 5.

³² Jagdish Bhagwati, ‘Free Trade: What Now?’ (Keynote Address on the occasion of the International Management Symposium at which the 1998 Freedom Prize of the Max Schmidheiny Foundation was awarded, St Gallen, 25 May 1998) 7.

³³ Robert Hudec, ‘Circumventing Democracy: The Political Morality of Trade Negotiations’ (1993) 25 New York University Journal of International Law and Politics 311, 314.

“By the time it was enacted, Smoot-Hawley contained a tariff increase for just about every U.S. economic interest with enough wit to have its supporters place a telephone call to their representatives in Congress”.³⁴

Hudec cautions that we should “have second thoughts about both Smoot-Hawley and about the wisdom of direct democracy as a decision-making process for tariff policy matters.”³⁵ Rather than viewing Smoot-Hawley as an efficient use of optimum tariffs, or a democratic reflection of social choices, Hudec views it as a failure of domestic politics driven by vested interests.

The practitioners’ story views trade agreements as a mechanism for countering protectionism by shifting the site of regulatory authority (and decision-making) with respect to trade policy. This shift operates in two main ways. In the national context, the legislative power to regulate tariffs is checked (at least partly) by the Executive’s power to negotiate treaties on tariff policy.³⁶

More significantly, trade agreements shift regulatory authority from the national to the multilateral level. This addresses some of the deficiencies in the domestic political process by giving a voice to foreign exporters (represented by their trade negotiators) and by creating a platform for export interests (or even consumers) to support liberalisation. Importantly, these trade commitments “tie the hands” of governments and offer them a tool to better resist calls for protectionism.

Shifting regulatory authority to the supra-national level can eliminate some “negative externalities”, but there are still weaknesses at the international level. In addition to suffering from a “democracy deficit”, there is no hard barrier preventing vested

³⁴ ibid 313.

³⁵ ibid.

³⁶ Weiler argues that this is part of a broader trend where the Executive has captured power over foreign policy, including with respect to the WTO and trade negotiations. For further details, see Weiler JHH, ‘Law, Culture, and Values in the WTO – Gazing into the Crystal Ball’ in Daniel Bethlehem, Isabelle Van Damme, Donald McRae, and Rodney Neufeld (eds), *The Oxford Handbook of International Trade Law* (OUP 2009).

interests from capturing the multilateral process.³⁷ Weiler claims that “capture by special interests of governmental positions that are then enshrined in WTO obligations is notorious.”³⁸

Hudec is well aware of this risk. He highlights a “narrow” scenario where international cooperation can actually create new opportunities for rent-seeking behaviour which otherwise would not be available. Firms might seek to drive down standards in other countries, even where those standards serve a legitimate purpose:

“If there is a justified health restriction that impedes export business, we may expect that there will be export business interests trying to reduce the level of that health restriction”.³⁹

There are pros and cons to shifting decision-making authority from the national to the multilateral level. The multilateral level strengthens those voices which support free trade but who have little clout in the domestic process (foreign producers, export interests, consumers). However, the multilateral level suffers from a “democracy deficit” and creates new opportunities for rent-seeking firms to seek to weaken legitimate regulation in other countries.

Regan describes the practitioners’ story as “everyone’s informal view”.⁴⁰ This highlights the main strength and weakness of this model: it is widely-accepted by people involved in trade negotiations (at least off-the-record), but it is difficult to model. In any case, Regan observes that “there seems to be no danger that trade

³⁷ ibid 756.

³⁸ ibid.

³⁹ Hudec (n 33) 320

⁴⁰ Regan (n 13) 1.

negotiators and adjudicators will abandon the practitioners' story for the standard model.”⁴¹

2.1.3 Summary

Both the standard model and practitioners' story recognise that unilateral protection can have “spillover effects” which reduce global efficiency. Each model agrees that governments ceded some of their regulatory autonomy to the WTO in order to gain benefits in the form of increased national welfare. They also concur that there were limits to that ceding of regulatory authority and that governments retained their right to regulate with respect to trade policy and other social policy areas which may impact trade.

The key question is how much regulatory authority did national governments cede to the WTO and how much did they retain? At the level of theory, no model can give a clear answer. In order to address this question, the next section will look at WTO law and the role played by national treatment and the general exceptions.

2.2 Legitimate Regulation and the WTO

How does WTO law strike a balance between the substantive obligations of its Members and their right to regulate for legitimate purposes? WTO law approaches this matter differently for trade and domestic measures, so my analysis will highlight some key distinctions.

2.2.1 Trade Measures and Domestic Measures

Trade measures are primarily defined as tariffs and quantitative restrictions.⁴² For trade measures, the WTO has a norm-making function which it uses to develop

⁴¹ ibid 18. This implies that there is little risk of WTO adjudicators using the standard model to interpret WTO Agreements by determining which protectionist measures are “rational”.

prescriptive rules. The WTO's main tool is to prohibit tariffs in excess of a Member's bound rates.⁴³

If we consider legal obligations as falling along a spectrum from highly-prescriptive rules to open-ended standards,⁴⁴ the WTO's regulation of tariffs appears highly rule-like. The tariff negotiation process is time-consuming and resource-intensive, as each Members must commit to a specific bound rate for thousands of different products. In principle, the effort that goes into negotiating these prescriptive tariff should pay dividends by providing a high level of clarity to implementers (Parliaments and customs agencies) and adjudicators.⁴⁵

When governments impose tariffs within their bound rates, they may do so for any reason, including to protect domestic industry.⁴⁶ They are under no obligation to justify the distortionary effects of such tariffs by reference to a non-protectionist regulatory purpose.⁴⁷

Domestic measures can also act as barriers to trade. For example, a ban on flavoured cigarettes may well reduce consumption and trade in tobacco products, even if the regulatory purpose is health-related and the discriminatory effect on trade

⁴² Export subsidies are also trade measures, but my analysis will focus on tariffs.

⁴³ General Agreement on Tariffs and Trade 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1A (adopted 15 April 1994, entered into force 1 January 1995) 1867 UNTS 187 ("GATT") Art II:1. Each Member's specific commitments can be found in its schedule of concessions.

⁴⁴ See eg Louis Kaplow, 'Rules versus Standards: An Economic Analysis' (1992-1993) 42 Duke Law Journal 557.

⁴⁵ It is worth noting that even an issue as prescriptively-regulated as tariffs cannot be completely rule-like. Invariably some uncertainty will prevail, such as the customs classification of new or unusual products.

⁴⁶ Protection is not the only reason for a government to impose tariffs. Some developing countries continue to use tariffs as an important source of government revenue.

⁴⁷ This principle would not apply where a government imposes a tariff in excess of its bound rate in breach of Article II.

is merely incidental.⁴⁸ How does the WTO determine if such measures are legitimate?

The WTO does not have the mandate or expertise to adopt a rule-like approach to the regulation of health matters such as tobacco control. In order to prevent the abuse of domestic regulation, the WTO prohibits discrimination between domestic and foreign goods through the National Treatment (NT) rule.⁴⁹

Article III:1 establishes that measures “should not be applied to imported or domestic products so as to afford protection to domestic production” (the SATAP requirement).⁵⁰ In practice, this means that WTO Members retain autonomy to adopt their preferred regulatory approach on matters like tobacco control based on their society’s preferences.⁵¹ The WTO adopts a “negative integration” model,⁵² which allows for a “pluralism” of regulatory approaches,⁵³ provided Members respect the National Treatment rule.

In NT disputes about domestic measures, adjudicators must assess whether the measure has discriminatory effects to the detriment of imported products and whether it can be justified by a non-protectionist (legitimate) regulatory purpose.

⁴⁸ This type of measure was the subject of dispute in Appellate Body Report, *United States – Measures Affecting the Production and Sale of Clove Cigarettes*, WT/DS406/AB/R, adopted 24 April 2012 (US – Clove Cigarettes).

⁴⁹ The WTO also prohibits discrimination between third country Members through the Most-Favoured Nation (MFN) rule in GATT Art I.

⁵⁰ GATT Art III:1.

⁵¹ Though Members can choose to restrict their policy space in other forums which do have a norm-making function such as the WHO Framework Convention on Tobacco Control (opened for signature 16 June 2003, entered into force 27 February 2005) 2302 UNTS 166 (“WHO FCTC”).

⁵² For a general discussion of the WTO’s negative integration model, see Gene Grossman, Henrik Horn and Petros Mavroidis, ‘Legal and Economic Principles of World Trade Law: National Treatment’ (Research Institute of Industrial Economics 2012, Stockholm, IFN Working Paper No 917).

⁵³ For a discussion of pluralism regarding sensitive moral legislation, see Howse R and Langille J, ‘Permitting Pluralism: The Seal Products Dispute and Why the WTO Should Accept Trade Restrictions Justified by Noninstrumental Moral Values’ (2012) 37(2) Yale Journal of International Law 367. While I agree with Howse et al about the importance of “pluralism” in WTO law, chapters 9-12 will explain why I disagree with their application of this principle to the *Seals* dispute.

These issues will be analysed further under the NT rule (Article III) and general exceptions (Article XX).

2.2.2 The National Treatment Rule

This section will consider several key concepts under the NT rule including: (i) *de jure* and *de facto* discrimination; (ii) fiscal and regulatory measures; (iii) how to distinguish trade and domestic measures; and (iv) the coverage of Article III in terms of products and regulatory instruments. It will also consider how the SATAP requirement operates in practice to test for protectionist intent and legitimate regulatory purpose.

2.2.2.1 Key Concepts under the NT Principle

De Jure and De Facto Discrimination

WTO law distinguishes between measures which formally discriminate on the basis of origin (*de jure* discrimination)⁵⁴ and those which are formally origin-neutral but have “disparate impacts”⁵⁵ on imported and domestic products (*de facto* discrimination).⁵⁶

In the early commentary, Hudec considered *de jure* discrimination to be the main problem GATT was trying to address. He noted:

⁵⁴ De jure discrimination can be defined as “a measure that discriminates on its face as to the origin of products”; Panel Report, *Japan – Measures Affecting Consumer Photographic Film and Paper*, WT/DS44/R, adopted 22 April 1998 (*Japan Film*) para 10.380.

⁵⁵ For a discussion on the “aims and effects” test and origin-neutral measures with “disparate impacts”, See Don Regan, ‘Regulatory Purpose and “Like Products” in Article III:4 of the GATT (with Additional Remarks on Article III:2)’ (2002) 36(3) Journal of World Trade 443, Don Regan, ‘Further Thoughts on the Role of Regulatory Purpose Under Article III of the General Agreement on Tariffs and Trade: A Tribute to Bob Hudec’ (2003) 37(4) Journal of World Trade 737 and Robert Hudec, ‘GATT/WTO Constraints on National Regulation – “Aims and Effects” Test’ (1998) 32(3) The International Lawyer 619.

⁵⁶ De facto discrimination can be defined as “a measure that in its application upsets the relative competitive position between domestic and imported products”; *Japan Film* (n54) para 10.380.

“Historically, GATT has been principally occupied with border measures and explicitly discriminatory measures, with *de facto* discrimination only becoming a major concern relatively recently”.⁵⁷

Since Hudec’s analysis, *de facto* discrimination has become increasingly important.⁵⁸

In case there was any doubt, adjudicators have explicitly confirmed that *de facto* discrimination is covered by Article III. In *Korea Beef*, the Appellate Body noted “there may be cases where the application of formally identical legal provisions would in practice accord less favourable treatment to imported products”.⁵⁹

As the jurisprudence on *de facto* discrimination continues to grow, it has become clear that such discrimination can be hard to identify and that it raises complex legal issues. As Grossman notes:

“an inherent feature of domestic policy instruments is that they can be used both for protectionism, and for purposes that would be in the collective interest of GATT Members to allow”.⁶⁰

How should WTO law “sort the wheat from the chaff”?⁶¹

Fiscal and Regulatory Measures

Article III:2 addresses fiscal measures and adjudicators have recognised that it contains two separate substantive obligations. The first sentence prohibits non-discrimination for “like products” while the second sentence applies to “directly competitive or substitutable” (DCS) products.

⁵⁷ Hudec (n 55), 622. Hudec further notes that “of the first 207 legal complaints filed in GATT between 1948 and 1990, only a small handful involved claims of *de facto* discrimination by internal regulatory measures.”

⁵⁸ Indeed, the case studies which I consider in this thesis relate to *de facto* discrimination.

⁵⁹ Appellate Body Report, *Korea – Measures Affecting Imports of Fresh, Chilled and Frozen Beef*, WT/DS161/AB/R, WT/DS169/AB/R, adopted 10 January 2001 (*Korea Beef*) para 136 (citing the GATT Panel in *Section 337*).

⁶⁰ Grossman et al (n 52) 102.

⁶¹ *ibid* 32.

Article III:4 deals with regulatory measures which discriminate between imported and domestic “like products”. These regulatory measures are the main focus of my analysis, including in the case studies for *Korea Beef*, *Brazil Tyres* and *Seals*.⁶²

Distinguishing Trade and Domestic Measures

It is not always simple to draw a clear line between trade and domestic measures. Ad Article III establishes that any fiscal or regulatory measure which “applies to an imported product and to the like domestic product” will be treated as a domestic measure even if it is “collected or enforced in the case of the imported product at the time or point of importation”.⁶³ This means that an “import ban” will be treated as a domestic measures in situations where it merely seeks to operationalise a domestic sales ban for imported products at-the-border.⁶⁴

Coverage of Article III

Article III covers “internal taxes and other internal charges, and laws, regulations and requirements”.⁶⁵ It is therefore designed to discipline a wide range of domestic instruments. Certain types of domestic instrument, including subsidies and government procurement, are explicitly excluded from the coverage of the National Treatment rule.⁶⁶

⁶² It is important to note that at the panel stage, the EU is found to breach obligations under both the MFN Principle (Article I) and National Treatment Principle (Article III). At the appeal stage, the main focus of analysis was the EU’s MFN breach, though the Article XX analysis to justify the seal products ban would have been equally relevant to the defence of its Article III breach. In fact, since it entailed breaches under both Articles I and III, *Seals* highlights some of the benefits of addressing regulatory purpose under Article XX (in terms of reduced fragmentation), rather than having a separate SATAP analysis within Article III.

⁶³ GATT (n 43) Ad Art III.

⁶⁴ For example, in the *Seals* dispute, the domestic ban on the marketing and sale of seal products was accompanied by an import ban and the seal products ban was therefore treated as a domestic measure.

⁶⁵ GATT (n 43) Art III:1.

⁶⁶ ibid Art III:8.

Grossman argues that Article III is “an anti-circumvention device” designed solely to “safeguard the value of tariff concessions”.⁶⁷ If this was true, the National Treatment rule would only apply to products where Members have made tariff commitments. The Appellate Body has explicitly rejected this view: “Article III is not limited to products that are the subject of tariff concessions under Article II”.⁶⁸

2.2.3 SATAP, Regulatory Purpose and Protectionist Intent

The Appellate Body has stated that:

“The broad and fundamental purpose of Article III is to avoid protectionism in the application of internal tax and regulatory measures”.⁶⁹

This section will consider where adjudicators should determine if a regulatory purpose is “protectionist” (under Article III or Article XX) and how it should test this matter.

The SATAP requirement has been the source of much debate in WTO law. For proponents of the “aims and effects” test, such as Hudec and Regan, SATAP represents a substantive element of a National Treatment claim. Grossman, Horn and Mavroidis support this approach, arguing that Article III:1 should be treated as a “substantive restriction” requiring complainants to demonstrate that the regulator had protectionist intent.⁷⁰

An alternative interpretation of SATAP is that complainants merely need to show protectionist effects under Article III. If adjudicators find there is a substantive breach, regulators could subsequently defend their measure by showing they pursued a legitimate regulatory objective under Article XX.

⁶⁷ Grossman et al (n 52) 10.

⁶⁸ Appellate Body Report, *Japan – Taxes on Alcoholic Beverages*, WT/DS8/AB/R, WT/DS10/AB/R, WT/DS11/AB/R, adopted 1 November 1996 (*Japan – Alcoholic Beverages II*) 17.

⁶⁹ ibid 16.

⁷⁰ Grossman et al (n 52) 110.

This interpretation shifts the consideration of regulatory purpose from Article III to Article XX. It also shifts the onus onto the regulator to demonstrate a legitimate purpose rather than to merely defend itself from accusations of protectionist intent.

2.2.3.1 SATAP

In practice, the Appellate Body has taken a mixed (and sometimes confusing) approach to SATAP and regulatory purpose. It has confirmed that the SATAP requirement in Article III:1 merely establishes a “general principle” which informs the “substantive obligations” found in subsequent paragraphs.⁷¹ However, for different substantive obligations, it takes varying approaches to SATAP. Under Article III:2 (second sentence), it treats SATAP as an element of the legal test. However, for Article III:2 (first sentence) and for Article III:4, it defers discussion of regulatory purpose to Article XX.

SATAP as a Substantive Obligation

For claims under Article III:2 (second sentence), SATAP is a substantive legal element which complainants must demonstrate in order to prove a breach. The Appellate Body has justified this position on the basis that, unlike the other obligations in Article III, “the language of Article III:2, second sentence, specifically invokes Article III:1”.⁷²

SATAP as a Mere Principle

For Article III:2 (first sentence) and Article III:4, SATAP is not explicitly invoked in the legal text and adjudicators have therefore determined that “the presence of a

⁷¹ For this reason, complainants do not make substantive claims under Article III:1. Rather, they challenge fiscal measures under Article III:2 or regulatory measures under Article III:4.

⁷² Appellate Body Report, *Japan – Alcohol II* (n 68) 24. This is because Art III:2 explicitly rejects any measures which are “contrary to the principles set forth in paragraph 1” (the SATAP requirement).

protective application need not be established separately".⁷³ For domestic regulatory measures (the subject of my analysis), the Appellate Body has confirmed that "a determination of whether there has been a violation of Article III:4 does not require a separate consideration of whether a measure 'afford[s] protection to domestic production'."⁷⁴ As long as the other legal elements are established (such as "like products" and "less favourable treatment"), the measure is deemed to be applied "so as to afford protection".

For those obligations without an explicit SATAP requirement (including Article III:4), scholars and parties have sought other ways to introduce regulatory purpose into the Article III analysis. Grossman argues that the like product analysis should consider "not only market likeness, but also likeness from a policy point of view".⁷⁵ Under this approach, products should only be considered like "if there is no legitimate reason" for discriminating between them.⁷⁶ Regulators could rely on regulatory purpose to demonstrate that certain products are not like.

This view has found little favour in the case law.⁷⁷ Adjudicators generally undertake the like product analysis by reference to the four factors from *Border Tax Adjustments*.⁷⁸ The Appellate Body has clarified that the like products determination

⁷³ *ibid* 18.

⁷⁴ Appellate Body Report, *European Communities – Regime for the Importation, Sale and Distribution of Bananas*, WT/DS27/AB/R, adopted 25 September 1997 (*Bananas III*) para 216.

⁷⁵ Grossman et al (n 52) 110. This analysis is undertaken in the context of fiscal measures, but the principle is equally applicable to regulatory measures.

⁷⁶ *ibid*.

⁷⁷ The panel in *Clove Cigarettes* (n 48) flirted with this "legitimate regulatory distinction" (LRD) approach for a national treatment claim under the TBT Agreement, but it was rejected by the Appellate Body.

⁷⁸ GATT practice is for likeness to be determined on a "case-by-case basis" by reference to a number of factors including "the product's end-uses", "consumers' tastes and habits" and the "product's properties, nature and quality"; GATT Working Party Report, *Border Tax Adjustments*, L/3464, adopted 2 December 1970 para 18. Panels also have a longstanding practice of considering the customs tariff classification; GATT Panel Report, *Japan – Customs Duties, Taxes and Labelling Practices on Imported Wines and Alcoholic Beverages*, L/6216, adopted 10 November 1987 (*Japan – Alcoholic Beverages I*) para 5.6.

“is concerned with competitive relationships between and among products”.⁷⁹ This suggests that market likeness, rather than policy likeness, is the key factor.⁸⁰

Regulatory purpose has also been argued under the “less favourable treatment” element. This was the EU’s approach in *Seals* when it argued that there was a “legitimate regulatory distinction” which justified its seemingly discriminatory treatment towards Canadian seal products. The Appellate Body was dismissive of this argument and opted:

“against an interpretation of Articles I:1 and III:4 that requires an examination of whether the detrimental impact of a measure on competitive opportunities for like imported products stems exclusively from a legitimate regulatory distinction”.⁸¹

Mitchell argues that the Appellate Body has now “resolutely rejected” the notion of “legitimate regulatory distinction” under Article III.⁸² The “less favourable treatment” analysis relies on market, rather than policy factors. The Appellate Body has confirmed that less favourable treatment “involves an assessment of the implications of the contested measure for the equality of competitive conditions between imported and like domestic products”.⁸³ According to Mitchell, *Seals* has “definitively established” that, in order to demonstrate less favourable treatment, “mere detrimental impact on competitive conditions for imports is sufficient”.⁸⁴

⁷⁹ Appellate Body Report, *European Communities – Measures Affecting Asbestos and Asbestos-Containing Products*, WT/DS135/AB/R, adopted 5 April 2001 (*Asbestos*) para 103.

⁸⁰ However, in *Clove Cigarettes* (n 48) para 119, the Appellate Body stated that: “the regulatory concerns underlying a measure, such as the health risks associated with a given product, may be relevant to an analysis of the ‘likeness’ criteria under Article III:4 of the GATT 1994, as well as under Article 2.1 of the TBT Agreement, to the extent they have an impact on the competitive relationship between and among the products concerned.” This suggests that policy likeness may be indirectly relevant.

⁸¹ Appellate Body Reports, *European Communities – Measures Prohibiting the Importation and Marketing of Seal Products*, WT/DS400/AB/R / WT/DS401/AB/R, adopted 18 June 2014 (*Seals*) para 5.125.

⁸² Andrew Mitchell, David Heaton and Caroline Henckels, *Non-Discrimination and the Role of Regulatory Purpose in International Trade and Investment Law* (Elgar 2016) 116.

⁸³ Appellate Body Report, *Seals* (n 81) 5.101.

⁸⁴ Mitchell et al (n 82) 116. Mitchell also summarises of some of the other approaches which adjudicators had used prior to the definitive ruling in *Seals*.

Despite decades of debate, WTO jurisprudence seems to have established clearly that Article III:4 is concerned with market factors rather than regulatory purpose. It contains no separate SATAP requirement and its key elements (such as like products and less favourable treatment) focus on the competitive relationship between imported and domestic products.

The Appellate Body has stated that "the fundamental purpose of Article III... is to ensure equality of competitive conditions between imported and like domestic products".⁸⁵ Subsequent panels have confirmed that "a finding on the WTO-consistency of the measure is not based on any consideration of the rationale or justification for the measure".⁸⁶

If we wish to find the balance between substantive obligations and retained sovereignty for domestic regulatory measures, it is clear that Article XX has become the key site for considering regulatory purpose. Indeed, in *Seals*, the Appellate Body confirmed that "under the GATT 1994, a Member's right to regulate is accommodated under Article XX".⁸⁷

Protectionist intent (under the SATAP requirement) and regulatory purpose (under Article XX) appear to assess very similar questions. Does it really matter that adjudicators have chosen to assess it under Article XX rather than Article III?

⁸⁵ Appellate Body Report, *Canada – Certain Measures Concerning Periodicals*, WT/DS31/AB/R, adopted 30 July 1997 (*Canada Periodicals*) 18.

⁸⁶ Panel Reports, *Brazil – Certain Measures Concerning Taxation and Charges*, WT/DS472/R, Add.1 and Corr.1, WT/DS497/R, Add.1 and Corr.1, adopted 11 January 2019, as modified by Appellate Body Reports WT/DS472/AB/R, WT/DS497/AB/R (*Brazil Taxation*) 7.153. This statement was made in the context of Article III:2 (first sentence) but is equally applicable to Article III:4.

⁸⁷ Appellate Body Report, *Seals* (n 81) para 5.125.

2.2.3.2 Regulatory Purpose: Article III or Article XX?

For advocates of the “aims and effects” test, the choice between Articles III and XX is highly significant. They argue that deferring the “aims” discussion to Article XX “is likely to impose a more restrictive regime”⁸⁸ which will “often yield different outcomes”.⁸⁹

In addition to shifting the burden of proof to the regulator, Article XX only provides access to an exhaustive list of exceptions. Regan suggests that this could lead to “draconian” results.⁹⁰ Grossman considers it “fairly easy” to identify legitimate objectives not on the list, such as “promotion of culture”, but he concedes that in practice this interpretive approach has not “unduly constrained Members”.⁹¹

Regan and Hudec suggest there is a “public relations” imperative: Members oppose their measures being branded a “violation”, even if they can ultimately defend it under Article XX.⁹² This objection may have become outdated. Recent WTO observers seem to be more interested in the final dispute outcome than intermediate findings of violation.

Some scholars also object to the current approach as it creates fragmentation between Article III:2 second sentence (which considers regulatory purpose under the substantive SATAP requirement) and other fiscal or regulatory measures (which defer this matter to Article XX). Mitchell further highlights that the rejection of

⁸⁸ Grossman et al (n 52) 116

⁸⁹ ibid 103-4.

⁹⁰ Regan (n 55) p752.

⁹¹ Grossman et al (n 52) 121-122.

⁹² Regan (n 55) 749.

“legitimate regulatory distinctions” under Article III:4 leads to fragmentation between Article III and the TBT case law.⁹³

It seems as though some fragmentation is unavoidable due to the convoluted drafting of the different Agreements. In *Seals*, the EU was found to be in breach of both non-discrimination obligations (under Articles I and III). If adjudicators applied the “aims and effects” test under Article III:4, the EU’s regulatory purpose would have been considered under Article III (for the NT claim) and Article XX (for the MFN claim). In that case, the rejection of “aims and effects” reduced fragmentation between different provisions.

In my view, the most sanguine analysis of the “aims and effects” debate comes from Weiler who notes that, as a textual matter, there are legitimate arguments in support of both SATAP interpretations. He considers the real world consequences of the “aims and effects” debate to be marginal. He rejects the view that there are important implications for the onus of proof⁹⁴ or for which policy purposes may be used to justify disparate impacts.⁹⁵

Weiler considers the debate about “aims and effects” to be a symbolic battle about the “normative hierarchy”, rather than an issue with meaningful consequences for doctrinal law.⁹⁶ The decision to reject the “aims and effects” test established the primacy of “the market and liberalized trade”,⁹⁷ while it “relegated all competing values to ‘exceptions’ with deep symbolic and cultural implications.”⁹⁸

⁹³ See eg Mitchell (n 82).

⁹⁴ Weiler (n 36) 766. Weiler describes such claims as “overstated”. . .

⁹⁵ ibid 767. Weiler states: “from a practical point of view this difference seems more illusory than real. Article XX is broad and sufficiently open-textured to cater for most exigencies”.

⁹⁶ ibid 768.

⁹⁷ ibid.

⁹⁸ ibid 771.

2.2.3.3 How to Assess Regulatory Purpose

The rejection of the “aims and effects” test has led to fragmentation within Article III. For those fiscal measures covered by Article III:2 second sentence, adjudicators must assess if the measure was applied “so as to afford protection”. For the other National Treatment obligations, any consideration of regulatory purpose occurs under the Article XX general exceptions.⁹⁹

Leading commentators suggest that adjudicators should avoid analysing the “subjective motive of legislators and regulators”¹⁰⁰ (or their *mens rea*)¹⁰¹ when considering regulatory purpose. The Appellate Body confirmed this view when it stated that "the subjective intentions inhabiting the minds of individual legislators or regulators do not bear upon the inquiry, if only because they are not accessible to treaty interpreters."¹⁰² This suggests that any attempt to identify subjective motives is not only inappropriate, but even impossible.

However, there have been cases where adjudicators took into account governmental statements about intent. In *Canada – Periodicals*, the Appellate Body cited statements of protectionist intent by the government including “we must protect ourselves against split-runs coming from foreign countries” and references to “the long-standing policy of protecting the economic foundations of the Canadian periodical industry”.¹⁰³

⁹⁹ In theory, if a regulator fails the SATAP test under Article III, it could still seek to raise a defence under Article XX.

¹⁰⁰ Hudec (n 55) 642. Hudec describes this as “traditionally a no-man’s-land for courts”.

¹⁰¹ Weiler acknowledges this as a theoretical option but dismisses it as a “hard” approach; Weiler (n 36) 764.

¹⁰² Appellate Body Report, *Chile – Taxes on Alcoholic Beverages*, WT/DS87/AB/R, WT/DS110/AB/R, adopted 12 January 2000 (*Chile – Alcohol*) para 62. The AB was discussing the notion of “SATAP” under Article III:2 second sentence, however this principle appears to apply equally to other parts of GATT including Article XX.

¹⁰³ Appellate Body Report, *Canada Periodicals* (n 85) 31.

In *Mexico Soft Drinks*, the panel claimed that “the Appellate Body has confirmed that statements made by government representatives of a Member, admitting to the protective intent of a measure, may be relevant”.¹⁰⁴ However, they “cautioned against ascribing too much importance to the subjective legislative intent of legislators and regulators... particularly when that declared intent is that protectionism was not an objective”.¹⁰⁵

Adjudicators generally rely on “objective indicators of purpose”,¹⁰⁶ an approach which is consistent with the DSU’s requirement that panel’s undertake an “objective assessment”.¹⁰⁷ The Appellate Body has confirmed that “the statutory purposes or objectives” might be discernible “to the extent that they are given objective expression in the statute itself”.¹⁰⁸ In key SATAP cases, the Appellate Body was able to find “objective” evidence of protectionism in the “magnitude of the dissimilar taxation”¹⁰⁹ or “the truncated nature of the line of progression of tax rates”.¹¹⁰

Weiler states that failure to offer “a plausible explanation” for a domestic measure which discriminates against imported products should lead to “a constructive presumption of bad purpose”.¹¹¹ This suggests that a positive finding of legitimate regulatory purpose can be used to construe that there was no protectionist intent. They appear to be two sides of the same coin.

¹⁰⁴ Panel Report, *Mexico – Tax Measures on Soft Drinks and Other Beverages*, WT/DS308/R, adopted 24 March 2006, as modified by Appellate Body Report WT/DS308/AB/R (*Mexico – Soft Drinks*) para 8.91.

¹⁰⁵ ibid.

¹⁰⁶ Hudec (n 55) 642.

¹⁰⁷ Dispute Settlement Rules: Understanding on Rules and Procedures Governing the Settlement of Disputes, Marrakesh Agreement Establishing the World Trade Organization, Annex 2 (adopted 15 April 1994, entered into force 1 January 1995) 1869 UNTS 401 (“DSU”) Art 11.

¹⁰⁸ Appellate Body Report, *Chile – Alcohol* (n102) para 62.

¹⁰⁹ Appellate Body Report, *Japan – Alcohol II* (n 68), 29.

¹¹⁰ Appellate Body Report, *Chile – Alcohol* (n102), para 71.

¹¹¹ Weiler (n 36) 764.

2.2.4 General Exceptions (Article XX)

In order to show that a domestic regulatory measure pursues a legitimate purpose, the onus is on the respondent to defend its measure under the Article XX general exceptions. The relevant provision reads as follows:

“Subject to the requirement that such measures are not 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, nothing in this Agreement shall be construed to prevent the adoption or enforcement by any contracting party of measures:

- (a) necessary to protect public morals;
- (b) necessary to protect human, animal or plant life or health;
- ...
- (d) necessary to secure compliance with laws or regulations which are not inconsistent with the provisions of this Agreement, including those relating to customs enforcement, the enforcement of monopolies operated under paragraph 4 of Article II and Article XVII, the protection of patents, trademarks and copyrights, and the prevention of deceptive practices;
- ...
- (g) relating to the conservation of exhaustible natural resources if such measures are made effective in conjunction with restrictions on domestic production or consumption”¹¹²

Adjudicators have determined that Article XX is the key provision which protects the regulatory autonomy of Members for domestic measures. The next chapter will describe the Article XX legal test in detail.

2.3 Legal Interpretation

Considering its critical role safeguarding the right to regulate, Article XX is a relatively short provision. In *Seals*, the key question under Article XX was whether

¹¹² GATT (n 43) Art XX. I have highlighted the four sub-paragaphs which have arisen most often in WTO disputes, including the three exceptions which are based on a “necessity” standard. These are the four general exceptions which are relevant to my subsequent analysis.

the Seal Regime was “necessary” to protect “public morals”, but neither of these terms is clearly defined in the GATT text.¹¹³

In terms of legal interpretation, Jackson notes that “the VCLT goes through a series of about ten or twelve (depending on how you count) techniques in articles 31 and 32, but there are a myriad of other possible techniques not mentioned therein.”¹¹⁴ This section will briefly consider some of the key issues which adjudicators face when interpreting Article XX according to its “ordinary meaning” as well as its “object and purpose”.¹¹⁵

2.3.1 Textual Interpretation

Article XX provides guidance to Members about which public interests may justify a violation of GATT obligations.¹¹⁶ Public interests such as “public morals” and “health” are explicitly named, but the text provides little guidance on the standard of review. Bown and Trachtman note that “the only word that is being interpreted is the word ‘necessary’.”¹¹⁷ Since this term essentially determines “how much deference should be given to the nation-state”,¹¹⁸ it is imperative that regulators understand how necessary a measure must be for Article XX to protect it?

There are two main approaches to interpreting “necessity”. Under the “alternative measures” (AM) test, adjudicators must determine if the regulator’s objective could

¹¹³ The “necessity test” will be analysed extensively in chapters 3-7. “Public morals” will be addressed in chapter 8.

¹¹⁴ Jackson (n 1) 449-50. Jackson also questions whether “the VCLT is really appropriate in the WTO context” and claims that it “has not been accepted by large numbers of WTO members”.

¹¹⁵ These methods are both referred to in the Vienna Convention on the Law of Treaties (entered into force 27 January 1980) 1155 UNTS 331 (“VCLT”) Art 31(1).

¹¹⁶ I use the term “public interest” to describe the ten general exceptions listed in Article XX such as public morals and health. This approach comes from Diebold who considers it a useful way to distinguish the listed public interests from the specific “policy objectives” which arise in each case; See Nicolas Diebold, ‘The Morals and Order Exceptions in WTO Law: Balancing the Toothless Tiger and the Undermining Mole’ (2007) 11(1) Journal of International Economic Law 43.

¹¹⁷ Chad Bown and Joel Trachtman, ‘Brazil – Measures Affecting the Import of Retreaded Tyres: A Balancing Act’ (2009) 8 World Trade Review 85, 89.

¹¹⁸ Jackson (n 1) 448.

have been achieved in a manner which was less trade-restrictive or WTO-inconsistent.¹¹⁹ The more recent “weighing and balancing” (W&B) test requires adjudicators to balance the social utility of the measure relative to its trade-restrictiveness. Some commentators describe this as a type of cost-benefit analysis, though there are competing views about the exact nature of the test.¹²⁰ The precise content of these two tests has evolved over time and there is debate about how they interact with one another.

Some authors suggest that textual interpretation can be used to determine the meaning of the term necessary, especially the choice of whether to use an AM or W&B test. Regan claims that the search for a “less-restrictive alternative” under the AM test is what is “suggested most naturally by the word ‘necessary’.”¹²¹ Bown and Trachtman appear to agree that the term “would most naturally suggest... a least-treaty-inconsistent-alternative test”.¹²²

They further argue that:

“on the precedent side, prior to *Korea–Beef* and *EC–Asbestos*, there was an entrenched understanding that ‘necessary’ meant least-treaty-inconsistent-alternative-reasonably-available”.¹²³

WTO negotiators knew this was the contemporary interpretation of the term “necessary” when they left Article XX untouched in the Uruguay Round. Jackson does not agree that the term “necessary” provides clear guidance. He observes “a

¹¹⁹ In some early disputes, adjudicators used a “less treaty-inconsistent” standard. This will be discussed further in section 3.3.1.

¹²⁰ Trachtman’s framework for understanding the W&B test will be explained in chapter 5.

¹²¹ Regan (n 55) 348.

¹²² Bown and Trachtman (n 117) 89.

¹²³ ibid.

certain amount of ambiguity in the treaty language".¹²⁴ This is consistent with HLA Hart's view on the "open-textured" nature of language.¹²⁵

The appropriate standard of review under Article XX is one of the most heated debates in WTO law. This critical term ultimately determines how much deference adjudicators must pay to regulators. It does not seem possible to answer this question merely by textual interpretation of the term "necessary".¹²⁶

The open-texture of language

The "indeterminate" nature of the term "necessary" can be understood by looking to Hart's analysis of the "open texture" of words and rules. According to Hart, "there is a limit, inherent in the nature of language, to the guidance which general language can provide."¹²⁷ There may be "unchallenged" or "easy" cases, but there will also be "hard" cases without certain answers. Hart argues that "canons of 'interpretation' cannot eliminate, though they can diminish, these uncertainties".¹²⁸

The Appellate Body has recognised that the term "necessary" refers to "a range of degrees of necessity"¹²⁹ or a "continuum".¹³⁰ If we apply Hart's framework to the

¹²⁴ Jackson (n 1) 448. Jackson further notes that "this kind of ambiguity is unavoidable, incidentally, as it is nearly impossible to negotiate a treaty text with one hundred nations without leaving some gaps and ambiguities."

¹²⁵ See HLA Hart (OUP 2012), Chapter VII(1) ("The Open Texture of Law").

¹²⁶ The debate about the standard of review for "necessary" resembles the debate about SATAP. While some authors suggest that the text is clear, it seems more likely that both terms are capable of multiple interpretations. The debate about how to define the terms is ultimately about deeper ideological differences rather than the literal meaning of the terms.

¹²⁷ Hart (n 125) 126.

¹²⁸ ibid.

¹²⁹ Appellate Body Report, *Korea Beef* (n 59) para 161.

¹³⁰ Panel Report, *European Communities – Conditions for the Granting of Tariff Preferences to Developing Countries*, WT/DS246/R, adopted 20 April 2004, as modified by Appellate Body Report WT/DS246/AB/R (*EC – Tariff Preferences*) para 7.211.

term “necessary”, this means there may be certain situations where it has a clear meaning and others where it falls into the “penumbra of uncertainty”.¹³¹

The Appellate Body has engaged in some analysis to identify and delineate the “clear” cases. In *Korea Beef*, it stated that measures which are “indispensable or of absolute necessity... certainly fulfil the requirements” of the necessity test.¹³² At the other end of the continuum, in *Brazil Tyres*, it rejected the notion that a measure which only makes a “marginal or insignificant contribution” to its objective could be considered necessary.¹³³

While this adds some clarity to the meaning of “necessary”, work remains to shed light on the “penumbra of uncertainty”. Hart argues that these “hard” cases offer the adjudicator “something in the nature of a choice between open alternatives”.¹³⁴

There may not be “one uniquely correct answer to be found”¹³⁵, but rather:

“The discretion thus left to [the adjudicator] by language may be very wide; so that if he applies the rule, the conclusion, even though it may not be arbitrary or irrational, is in effect a choice”.¹³⁶

Bix understands this discretion to mean that there may be “a range of possible right answers” and that we should reject the formalist approach of seeking “one right answer”.¹³⁷ If this is correct, the grey area between “indispensable” and “insignificant contribution” actually offers much discretion to the Appellate Body. Perhaps the AM and W&B tests are both justifiable as “right answers” (or they can even be combined into a composite test).

¹³¹ Hart (n 125) 12.

¹³² Appellate Body Report, *Korea Beef* (n 59) para 161.

¹³³ *ibid* 150.

¹³⁴ Hart (n 125) 127.

¹³⁵ *ibid* 132.

¹³⁶ *ibid* 127.

¹³⁷ Brian Bix, *Law, Language, and Legal Determinacy* (OUP 1995) 26.

Constructive Ambiguity

In addition to the uncertainties inherent in language, WTO law suffers from an additional problem of intentional ambiguity. In a proposal for DSU reform, the US argued that WTO law contains “constructive ambiguity”:

“where the negotiators leave unresolved particular issues by agreeing on language that does not resolve the issue and is capable of more than one interpretation. Constructive ambiguity can serve as a placeholder marking an area where negotiators accept that it may be appropriate to agree on disciplines but where further negotiation is necessary before those disciplines can be specified”.¹³⁸

This may explain some of the ambiguity in WTO Agreements. If GATT drafters intended for the term “necessary” to require the least restrictive means (or a cost-benefit analysis), they certainly could have used language which clarified the standard of review they were seeking.

The US position draws on Hart’s view that certain terms might be “capable of more than one interpretation”. However, where Hart believes that this ambiguity provides discretion to adjudicators (as there is more than one right answer), the US argues that it takes discretion away from adjudicators (as there is no right answer). The US suggests such provisions are mere “placeholders” for future negotiation: adjudicators must wait for further guidance from legislators.

But, how should adjudicators deal with such ambiguities in practice? Dealing with constructive ambiguity is one of the major challenges facing the “juridification” of the

¹³⁸ US Proposal, ‘Further Contribution of the United States on Improving Flexibility and Member Control in WTO Dispute Settlement’ TN/DS/W/82/Add.1 (25 October 2005) 2.

system, as the diplomatic, consensus-based GATT panel system has evolved into the WTO's quasi-judicial dispute settlement system.¹³⁹

Evolution from Diplomacy towards Law

Constructive ambiguity might have been plausible under the original GATT dispute settlement system. As Jackson notes, GATT disputes appeared more like a "diplomatic" than a "judicial" process where contracting parties had safeguards to ensure they were not exposed to intolerable decisions.¹⁴⁰

Disputes were heard by a Working Party (rather than an independent panel of experts) where the contracting party defending its measure helped produce the report. Due to the reverse consensus rule, any contracting party could block an adverse decision which it considered intrusive on its regulatory autonomy. In this system, contracting parties were the lawmakers (through GATT negotiations), the subject of legal obligations¹⁴¹ and the adjudicators (in Working Parties).¹⁴²

Over time, dispute settlement became increasingly "judicial" and governments gradually lost their veto power over the process and outcome. In the 1950s, the procedure changed from Working Parties to panels of experts. The most significant changes happened when the DSU was introduced under the Marrakesh Agreement.¹⁴³ The DSU introduced a "reverse consensus" rule which meant that

¹³⁹ For a discussion of the increasingly judicial culture of WTO adjudication, see JHH Weiler, 'The Rule of Lawyers and the Ethos of Diplomats: Reflections on the Internal and External Legitimacy of WTO Dispute Settlement' (Harvard Law School 2000, Harvard Jean Monnet Working Paper 9/00).

¹⁴⁰ For a detailed discussion of the evolution of WTO adjudication from "diplomatic" to "quasi-judicial", see Jackson (n 1).

¹⁴¹ For the purpose of this claim, I treat the State as a unitary body, even though in practice treaty negotiations might involve the lawmaking being done by one arm (such as the Federal Executive) which creates obligations for other parts of the government (such as a State/Provincial legislature).

¹⁴² Under a GATT panel system, the respondent had to consent to the establishment of a Working Party or panel and it had to consent to the adoption of the final report.

¹⁴³ DSU (n 107).

Members could no longer block the establishment of a panel¹⁴⁴ or the adoption of panel and Appellate Body reports.¹⁴⁵

Officially, the process remains “Member-driven” insofar as the dispute reports have no binding effect until adopted by the Member-driven Dispute Settlement Body (DSB). Official interpretations of the covered agreements can only be made by the General Council or Ministerial Conference (bodies which are also composed of Members).¹⁴⁶

In practice, the reverse consensus rule ensures that all dispute reports are adopted by the DSB as a matter of course. As Weiler notes, “the circumstances would have to be utterly unique to envisage a consensus in the General Council and/or Ministerial Conference to overturn an interpretation or decision of the Appellate Body.”¹⁴⁷

Unlike Working Parties, the Appellate Body is not equipped to deal with “constructive ambiguity”. If the covered agreements adopt ambiguous language on politically-sensitive issues, the Appellate Body has no mandate to defer these questions until further negotiations occur. Former Appellate Body Member Ehlermann strongly rejects “all manner of judicial activism”, but he astutely observes that “neither a panel nor the Appellate Body is entitled to refuse to address a claim because the panelists or the Appellate Body members want to avoid deciding a legal question that has delicate political consequences.”¹⁴⁸

¹⁴⁴ ibid Art 6(1).

¹⁴⁵ ibid Arts 16(4) and 17(14).

¹⁴⁶ Marrakesh Agreement (n 5) Art IX:2.

¹⁴⁷ Weiler (n 139) 10.

¹⁴⁸ Claus-Dieter Ehlermann, ‘Tensions between the dispute settlement process and the diplomatic and treaty-making activities of the WTO’ (2002) 1(3) World Trade Review 301, 305.

Due to the negotiation “stalemate” at the WTO, the legislative function appears unlikely to remove ambiguity from WTO texts anytime soon. In the meantime, adjudicators have an obligation to “address each of the issues raised” by parties.¹⁴⁹ Adjudicators have no mandate to classify certain legal language as “constructive ambiguity” and “remand” it to negotiators for resolution.

In any case, even if we accept that “constructive ambiguity” exists in WTO law, it does not necessarily explain the vagueness of Article XX.¹⁵⁰ It is plausible that Article XX was intentionally drafted as an open-ended standard rather than a prescriptive rule.

Rules and Standards

If Article XX was intended to operate as a legal standard, it means that GATT drafters chose not to define the precise contours of their retained regulatory autonomy; instead, they delegated this task for future adjudicators to resolve.

There’s a robust literature on the differences between rules and standards, including when each approach is preferable.¹⁵¹ The advantage of a standards-based approach is that it contains flexibility to deal with new situations as they arise. Legislators/negotiators are not required to foresee each situation or determine a rule in advance. This is particularly useful for GATT drafters seeking to determine the

¹⁴⁹ DSU (n 107) Art 17.12.

¹⁵⁰ The US has most aggressively pursued the notion of “constructive ambiguity” in the context of the Agreement on the Implementation of Article VI of GATT 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1A (adopted 15 April 1994, entered into force 1 January 1995) 1868 UNTS 201 (“Anti-Dumping Agreement”). The US has advocated for a highly-deferential standard of review under Art 17(6)(ii) which states: “Where the panel finds that a relevant provision of the Agreement admits of more than one permissible interpretation, the panel shall find the authorities’ measure to be in conformity with the Agreement if it rests upon one of those permissible interpretations.” WTO adjudicators have failed to operationalise this principle, reasoning that there is (almost) always a single “permissible interpretation”. This is one of the leading causes of US frustration with the Appellate Body which has played out in the protracted series of “zeroing” disputes.

¹⁵¹ See eg Kaplow (n 44).

regulatory autonomy of Members to introduce moral, health or environmental regulations.

The disadvantage of a standard is that it is less predictable. When governments set tariff rates, they generally know if their law is WTO-consistent. When they pass domestic regulations, such as a seal products ban, it is harder to be certain about WTO-consistency.

Since standards are less predictable, this gives greater authority to adjudicators to serve an interpretative function (which opponents may describe as “judicial activism” or “law-making”). In some respects, the legislators transfer their power to make the rules (as a general matter, before the fact) to adjudicators to do so (in specific cases, after the fact).

If Article XX is a standard, this suggests the GATT drafters intentionally kept the provision vague, not as a placeholder for further negotiation, but rather to let adjudicators resolve specific problems as they arose. This approach would permit adjudicators to interpret the necessity standard and apply it to real-world scenarios in accordance with basic principles of judicial interpretation.

In early GATT disputes, there may have been little need to clearly distinguish between “constructive ambiguity” (requiring further negotiation) or the use of legal standards (requiring adjudication). After all, it was ultimately the same actors making determinations in either case. Whether the contracting parties convened through the negotiation function (in a subsequent trade round) or the quasi-judicial function (as a Working Party), both mechanisms required consensus.

Ehlermann described the GATT's dispute settlement as "weak" because "it was based on the same principle as was the political process, that is the rule of consensus."¹⁵² Under either function, a government could not have its regulatory autonomy taken away without its consent.

However, as the WTO dispute settlement system has become increasingly "judicial", especially following the Uruguay Round, there are now a pool of independent adjudicators (at both the panel and appellate stage) who can make legal decisions which require compliance from Members. WTO law no longer permits Members to block adverse decisions,¹⁵³ though some Members have resorted to the extra-judicial option of merely being non-compliant.¹⁵⁴

The use of standards (rather than rules) raises an important issue about the role of precedent in WTO law.¹⁵⁵ Strictly speaking, the WTO's quasi-judicial system purports to shun the notion of *stare decisis*. The DSU states that "recommendations and rulings of the DSB cannot add to or diminish the rights and obligations provided in the covered agreements".¹⁵⁶

In practice, however, the WTO dispute system possesses many of the hallmarks of precedent. The Appellate Body has stated that adopted rulings "create legitimate expectations among WTO Members, and, therefore, should be taken into account where they are relevant to any dispute".¹⁵⁷ Where similar issues arise, panels

¹⁵² Ehlermann (n 148) 301.

¹⁵³ Blocking a dispute report would require consensus from all Members, including the "winning party", which means it is exceedingly unlikely to occur.

¹⁵⁴ When Members fail to implement an adverse dispute ruling within a reasonable period of time, they are potentially subject to other consequences such as "retaliation" by the complainant.

¹⁵⁵ Kaplow argues that "precedent may be seen as the (perhaps partial) conversion of a standard into a rule"; Louis Kaplow, 'General Characteristics of Rules' in Bouckaert B and De Geest G (eds), *Encyclopedia of Law and Economics* (Edward Elgar, 2000) 502, 511.

¹⁵⁶ DSU (n 107) Art 3.2.

¹⁵⁷ Appellate Body Report, Japan — Alcohol II (n 68) 107-108.

should follow earlier Appellate Body rulings “absent cogent reasons”.¹⁵⁸ These principles promote consistent jurisprudence and thus enable the DSU to fulfil its mandate of contributing “security and predictability to the multilateral trading system”.¹⁵⁹

Bacchus and Lester note:

While there is no formal system of precedent, there is an informal practice of taking into account past rulings to help ensure certainty and foreseeability for those who are affected by the rulings.¹⁶⁰

Jackson asserts it is “simply not credible to argue that there is no precedent effect in the WTO dispute settlement system” and believes we should focus on “how much weight is given to precedent”.¹⁶¹ In his view, “the WTO is not very far below the *stare decisis* top rung of the precedent ladder.”¹⁶²

2.3.2 Object and Purpose

Perhaps the object and purpose of the GATT/WTO can shed light on where the line is located between substantive obligations and retained autonomy. I have already described some competing views from economics, law and politics. Some of these competing explanations are referred to in the first recital of the Marrakesh Agreement preamble which states:

“Recognizing that their relations in the field of trade and economic endeavour should be conducted with a view to raising standards of living, ensuring full

¹⁵⁸ Appellate Body Report, United States – Final Anti-Dumping Measures on Stainless Steel from Mexico, WT/DS344/AB/R, adopted 20 May 2008 (*US – Stainless Steel (Mexico)*) para 160. For a deeper analysis, see James Bacchus and Simon Lester, ‘Of Precedent and Persuasion: The Crucial Role of an Appeals Court in WTO Disputes’ (12 September 2019) 74 *Free Trade Bulletin* 1.

¹⁵⁹ DSU (n 107) Art 3.2.

¹⁶⁰ Bacchus and Lester (n 158) 1.

¹⁶¹ Jackson (n 1) 449.

¹⁶² *ibid.*

employment and a large and steadily growing volume of real income and effective demand, and expanding the production of and trade in goods and services, while allowing for the optimal use of the world's resources in accordance with the objective of sustainable development, seeking both to protect and preserve the environment and to enhance the means for doing so in a manner consistent with their respective needs and concerns at different levels of economic development...”¹⁶³

Even if we focus narrowly on the GATT's trade and economic objectives, the WTO appears to have conflicting aims. The preamble endorses the economist's view that the WTO seeks to increase global welfare and maximise efficiency. It calls for “a large and steadily growing volume of real income” as well as “expanding the production of and trade in goods and services”.

However, the preamble echoes Ruggie's “embedded liberalism” by referring to distributional equity. It recognises the importance of “raising standards of living” and “ensuring full employment”. Economists are well-placed to address questions of efficiency, but the matter of “distributional” equity is better handled through political consensus, usually at the national level.

Over the life of GATT, objectives beyond trade and efficiency have become increasingly important. In the Uruguay Round, “poverty reduction” and “environmental concerns” were explicitly added to the list of objectives.¹⁶⁴ Adjudicators have recognised that the “WTO Agreement contains multiple policy objectives and all of these objectives are important”.¹⁶⁵

The preamble has been drawn on to interpret Article XX, including through “dynamic interpretation” to give broad meaning to non-trade values.¹⁶⁶ How do governments

¹⁶³ Preamble to the Marrakesh Agreement (n 5) 1st recital. Subsequent recitals provide further detail about other the aims of the WTO.

¹⁶⁴ Elsig et al (n 8) 3.

¹⁶⁵ Panel Report, *EC – Tariff Preferences* (n 130) para 7.52.

¹⁶⁶ In Chapter 8, I will discuss how the preamble was relied on by adjudicators to give a broad meaning to the term “exhaustible natural resources” to cover environmental protection writ large.

reconcile competing goals when the benefits of global efficiency come at the expense of local jobs (or environmental concerns)?

The existence of multiple, competing objectives in the WTO preamble may explain why Elsig describes it as “too ambitious”.¹⁶⁷ The preamble spells out the WTO’s main objectives, but provides little guidance on the critical question of how regulators should make specific trade-offs between competing values. This is precisely the difficult question which often finds its way to adjudication under Article XX.

Deference and Circumvention

The doctrine of “constructive ambiguity” suggests that where treaty language is vague, adjudicators should be highly-deferential to regulators. Dworkin would likely support this deferential approach, at least for a negative integration system like the WTO. He has advocated the view “that if a statute uses vague language it must be taken to have changed the legal status quo ante only to the extent justified by the indisputable core of the language employed”.¹⁶⁸

For GATT drafters, the status quo prior to negotiation was sovereign states possessing an “inherent power” to regulate”.¹⁶⁹ Dworkin’s approach suggests that to the extent the term “necessary” is vague, it should be interpreted to mean that governments did not cede their right to regulate.

However, this view has its weaknesses. The WTO’s approach to domestic regulation co-exists alongside its regulatory approach for trade measures. An

¹⁶⁷ Elsig et al (n 8) 2.

¹⁶⁸ Ronald Dworkin, *A Matter of Principle* (Harvard University Press 1985) 129.

¹⁶⁹ This principle was confirmed in Appellate Body Report, China – Measures Affecting Trading Rights and Distribution Services for Certain Publications and Audiovisual Entertainment Products, WT/DS363/AB/R, adopted 19 January 2010 (China Audiovisuals) para 222 where the Appellate Body stated: “we see the ‘right to regulate’, in the abstract, as an inherent power enjoyed by a Member’s government, rather than a right bestowed by international treaties such as the WTO Agreement”

overly-deferential approach to domestic regulation could enable circumvention of trade commitments. As Grossman notes:

“Governments have more ways to provide protection than only with tariffs. They often can replicate the effects of such policies, or at least nearly so, with combinations of other fiscal instruments, with quantitative restrictions, and with administrative and other impediments to trade”.¹⁷⁰

If tariff protection can be replicated through domestic measures, how can we explain the prescriptive rule-like approach for trade measures alongside a negative integration approach for domestic regulation? Why would GATT drafters take a more lenient approach to domestic measures if they can replicate the effects of a tariff?¹⁷¹

Regan appears wary of economic theories about circumvention. He notes that even where two measures may have the same economic effect, other “irrational” factors may render some approaches feasible and others politically-difficult or impossible. For example, in the case of subsidies, it is politically more feasible for governments to offer tax exemptions (revenue foregone) than grants (direct expenditure), even though this distinction would “appear as irrationality” under “any formal model that looks only at the balance in the treasury.”¹⁷²

For some “irrational” reason, the political calculus may vary for different types of measure even where they produce the same economic effect. The circumvention argument from economic theory would require stricter controls of domestic regulation, but Regan implies that “political” safeguards to prevent abuse already

¹⁷⁰ Grossman and Horn (n 14) 27.

¹⁷¹ If domestic regulation can be used to circumvent prescriptive trade commitments, should adjudicators take a strict approach to national treatment in order to prevent circumvention and ensure that trade commitments are meaningful?

¹⁷² Regan (n 13) 13. This discussion resembles the example of credit card surcharges in chapter 1 where a different framing leads to different perceptions of fairness/appropriateness even if the substantive effect remains the same under both approaches.

exist. While Regan argues convincingly that there are limits to the economic theory, he cannot create a model which clearly delineates which types of circumvention are politically feasible and which ones are implausible.¹⁷³

2.3.3 Legitimacy of the WTO System

Neither legal textualism nor the “object and purpose” can produce consensus on the appropriate standard of review for Article XX. Further, the Appellate Body appears constrained by extra-judicial considerations related to the WTO’s institutional legitimacy. As Weiler states:

“the legitimization strategy practised by the Appellate Body (whether express or implicit) has been one of hermeneutic prudence and institutional modesty with a keen eye on balancing internal and external legitimacy”.¹⁷⁴

Ehlermann believes there is a “widely perceived and more or less aggressively criticized ‘legitimacy’ deficit of the WTO” which he attributes to a dual “stalemate” of the Appellate Body and the WTO’s political/negotiating function.¹⁷⁵ Proponents of a strong and legitimate WTO face a more daunting task than Hart who was prepared to discuss the matter of legal interpretation unconstrained by practical concerns such as institutional legitimacy. At the WTO, however, the question of legitimacy is never far below the surface, especially when a regulator receives an adverse ruling.

Hart’s approach would offer a broad discretion to the Appellate Body due to the vague language chosen by GATT drafters. The Appellate Body appears to be aware that “legitimacy” concerns limit this discretion. Perhaps Dworkin’s “status quo” approach, which defers to the sovereignty of Members, is preferable.

¹⁷³ After all, it is difficult to model behaviours which are “irrational”.

¹⁷⁴ Weiler (n 139) 15.

¹⁷⁵ Ehlermann (n 148) 305.

The problem for the WTO is that there are two conflicting objectives at play. Article XX seeks to protect the right to regulate for legitimate public purposes, but it also seeks to curtail illegitimate protectionism. If WTO adjudicators adopt an interpretation which grants more regulatory autonomy to Members, the corollary is that this creates a higher risk of protectionist abuse.

Which objective is more important? Does the WTO create a right to trade or does it merely protect the equality of competitive opportunities? On trade measures, the WTO regulates market access conditions. For domestic measures, however, it merely seeks to protect “the equal competitive relationship between imported and domestic products”.¹⁷⁶

The Appellate Body seeks to find an appropriate balance between its “Members’ right to regulate” and “the desire to avoid creating unnecessary obstacles to international trade”.¹⁷⁷ Pirker astutely points out that “a ‘right’ is balanced against a ‘desire’, and not another right”.¹⁷⁸ Does this imply that respect for regulatory autonomy has a higher status in WTO law than the proscription against protectionism?

Today’s respondent (seeking more policy space) is tomorrow’s complainant (seeking to curtail another’s protectionism). The Appellate Body is engaged in a difficult balancing act. If it is too restrictive on the right to regulate, Members might accuse it of “judicial activism” or failing to respect the original GATT contract. However, granting too much policy space to Members would open the door to increased

¹⁷⁶ Appellate Body Report, *Japan – Alcohol II* (n 68) 16.

¹⁷⁷ Appellate Body Report, *Clove Cigarettes* (n 48) para 96.

¹⁷⁸ Benedikt Pirker, *Proportionality Analysis and Models of Judicial Review: A Theoretical and Comparative Study* (Europa Law Publishing 2013) 290.

protectionism which would undermine the WTO's ability to provide the regulatory infrastructure for an open trading system.

Contract or Constitution

There's an existential question which underpins the Appellate Body's dilemma: is the WTO constitutional or contractual in nature (or precisely where does it sit on the spectrum from contract to constitution)? Grossman sees the WTO as a contract between sovereign governments, describing it as a "negative integration" system where other social preferences "trump" trade obligations¹⁷⁹ and domestic policies can be "defined unilaterally".¹⁸⁰ Other commentators argue that the WTO has evolved from its contractual origins. Steger claims it is a "myth" that the WTO is merely a "contract" which "countries can withdraw from at any time."¹⁸¹

Much discussion of the WTO's status seeks to compare it with EU integration or US Federalism. While the US is a constitutional system, the EU is harder to classify. In some respects, it has taken steps in the direction of constitutionalism, for example through the judicial doctrine of "direct effects" and through the legislative approach of qualified majority voting (rather than consensus) for some subject matter areas.¹⁸²

The WTO may be moving in a similar direction, albeit at a slower pace. As Ehlermann states:

"The development from the original GATT to today's WTO has not been as revolutionary as the transformation of Europe from the original Coal and Steel Community to the EU of 25 or more Member States. However, for a variety of reasons, today's WTO is also very different, both in terms of quantity and of

¹⁷⁹ Grossman et al (n 52) 106.

¹⁸⁰ ibid 121.

¹⁸¹ Debra Steger, 'The Culture of the WTO: Why It Needs to Change' (2007) 10(3) JIEL 483, 490.

¹⁸² Weiler (n 36) 754.

quality, from the original club of CONTRACTING PARTIES that composed the old GATT".¹⁸³

The WTO appears to be more than a pure contract and to have made gradual steps, over the decades, towards deeper integration. However, it is hard to assess precisely where it sits on the spectrum from contract to constitution.

Pirker suggests that the WTO lacks "constitutional adjudication" and asserts that "adjudicators appear to feel poorly legitimized to engage in intrusive review".¹⁸⁴

Nagy argues that judicial "balancing" is appropriate for the "internal market regimes of federal states" based on "shared sovereignty" such as the US, EU or Australia.¹⁸⁵

While W&B testing is appropriate for systems of constitutional adjudication, Nagy notes that there are "remarkable differences between federal states and the world trade system" and that the use of balancing "finds no warrant in WTO law".¹⁸⁶

Jackson observes that WTO Members have always displayed "a measure of deference" to dispute rulings.¹⁸⁷ He argues that "consent" by Members should not be compulsory "for every small detail, or for every resolution of ambiguity or gap-filling by a dispute settlement institution."¹⁸⁸ Thus, he supports at least some discretion for adjudicators, even if this sometimes leads to adverse rulings which a "losing" party would not consent to. But, Jackson further notes that there must be

¹⁸³ Ehlermann (n 148) 308.

¹⁸⁴ Pirker (n 178) 281.

¹⁸⁵ Csondor Nagy, 'Clash of Trade and National Public Interest in WTO Law: The Illusion of "Weighing and Balancing" and the Theory of Reservation' (2020) 23(1) Journal of International Economic Law 143, 144-5. He offers as examples the "US Constitution's Dormant Commerce Clause, the Australian Constitution's Free Trade Clause, and the EU Internal Market".

¹⁸⁶ Ibid 161.

¹⁸⁷ Jackson (n 1) 452.

¹⁸⁸ ibid.

limits to the discretion of adjudicators, especially due to the WTO's weak legislative capacity to override.¹⁸⁹ Where should we draw the line?

How much regulatory autonomy did Members cede under the GATT/WTO and how much did they retain? Adjudicators seek to answer this question every time they assess if a domestic measure is "necessary" under Article XX. It is an unenviable task but, as Ehlermann observes, it cannot be avoided.

¹⁸⁹ Jackson (*ibid* 454) states "the criticism that the WTO system does not have a counterweight or checking and balancing procedure analogous to those found in nation-states (such as a legislature) has some merit".

CHAPTER 3 - THE ARTICLE XX LEGAL TEST AND STANDARD OF REVIEW

3.1 Introduction

The Appellate Body has stated that “Members have a large measure of autonomy to determine their own policies”.¹ The previous chapter established that this autonomy for domestic regulation is reserved in Article XX. The Appellate Body acknowledges that the Article XX exceptions are “important and legitimate in character”.² This chapter will briefly describe the elements of an Article XX defence before focussing in detail on the necessity test.

3.2 Elements of the Legal Test

The Article XX defence is “two-tiered” involving “provisional justification” under the sub-paragraph as well as “further appraisal of the same measure” under the chapeau.³ In addition to these two tiers, the characterisation of the objective also plays a critical role (and can influence the substantive analysis)⁴ which is why I address it as a key element in its own right.⁵ This section will therefore address the legal tests for the provisional justification, chapeau and characterisation of the objective.⁶

¹ Appellate Body Report, *United States – Standards for Reformulated and Conventional Gasoline*, WT/DS2/AB/R, adopted 20 May 1996 (*Gasoline*) 22.

² Appellate Body Report, *United States – Import Prohibition of Certain Shrimp and Shrimp Products*, WT/DS58/AB/R, adopted 6 November 1998 (*Shrimp*) para 121.

³ The “two-tiered” test was introduced by the Appellate Body in *Gasoline* (n 1) 21. It has been consistently followed by WTO adjudicating bodies. Some commentators consider the “two-tiered” description to be overly-simplistic. Diebold divides an Article XX defence into five key steps: policy objective, scope definition (“design step”), risk/risk analysis, necessity test and chapeau; Nicolas Diebold, ‘The Morals and Order Exceptions in WTO Law: Balancing the Toothless Tiger and the Undermining Mole’ (2007) 11(1) *Journal of International Economic Law* 43, 47–48.

⁴ The characterisation of the objective will be analysed in detail in the *Brazil Tyres* and *Seals* case studies.

⁵ This is consistent with the views of Diebold (n 3) that the policy objective is a separate element.

⁶ Since my main case study (*Seals*) is a public morals dispute, I will rely primarily on Article XX(a) jurisprudence, including *Colombia Textiles* (which in my view provides some of the Appellate Body’s clearest statements of

3.2.1 Provisional justification

The Appellate Body has divided the provisional justification into “two steps”. In a public morals dispute, the measure must be (i) “designed to” protect public morals (“design step”) and (ii) “necessary” to protect public morals (“necessity test).⁷ I will discuss each “step” separately.

3.2.1.1 Threshold Examination

The first step is a “threshold examination”⁸ or “design step”⁹ to determine if the measure addresses a relevant public interest under Article XX.¹⁰ The Appellate Body has noted that this is not “a particularly demanding step” especially compared to the necessity test which “entails a more in-depth, holistic analysis of the relationship between the measure and the protection of public morals.”¹¹ The measure must merely be “designed to”, “not incapable of” or have “a relationship to” the protection of public morals.¹² This element generally leads to little controversy though it can play an important role in framing the issues under dispute.¹³

Complainants sometimes concede the threshold examination and focus their artillery elsewhere,¹⁴ as adjudicators rarely dismiss Article XX defences at the threshold

law) and *Seals*. It will also draw on other significant “necessity” disputes under Articles XX(b) and (d) and some Article XX(g) disputes.

⁷ Appellate Body Report, *Colombia – Measures Relating to the Importation of Textiles, Apparel and Footwear*, WT/DS461/AB/R and Add.1, adopted 22 June 2016 (*Colombia Textiles*) para 5.67.

⁸ *ibid* 5.68.

⁹ *ibid* 5.68.

¹⁰ There does not seem to be a clear convention on how to name this step, so I will generally call it the “threshold examination” or “design step”. Under the TBT Agreement, the equivalent test inquires if the regulator pursues a “legitimate objective”. Diebold (n 3) refers to this as the “scope definition”.

¹¹ Appellate Body Report, *Colombia Textiles* (n 7) para 5.70.

¹² *ibid* paras 5.67-5.68.

¹³ Chapter 8 will provide a detailed doctrinal analysis of the “design step” for public morals disputes. It will show how the framing of key issues under the “design step” (including the policy objective) significantly affected the subsequent analysis.

¹⁴ Doyle notes that this occurred in *Audiovisuals* where “the United States was willing to concede that China’s measures were designed to protect public morals, thus satisfying the first element of the defense”; Christopher Doyle, ‘Gimme Shelter: The Necessary Element of GATT Article XX in the Context of the Audiovisuals Products Case (2011) 29 Boston University International Law Journal 143, 151.

stage.¹⁵ Fontanelli suggests that it would take “some temerity” for a panel to reject a defence at this threshold stage and assess that it “does not even deserve to reach the necessity test”.¹⁶

In my view, there remains a lack of clarity about whether the object of review for the threshold examination is the policy objective or the measure. In *Colombia Textiles*, the Appellate Body sought to find “the existence of a relationship between the measure and the protection of public morals”.¹⁷

However, in *Brazil Taxation*, the panel appeared to accept both objects of review. It sought to ensure that “the measure is not incapable of contributing to the objective”¹⁸ and that “the alleged *public policy objective* at issue is indeed a public moral objective”.¹⁹ The Appellate Body should provide clearer guidance on the object of review under the design step.

Adjudicators may sub-divide the design step into separate questions. For example, adjudicators took a two-part approach in *Seals* when they inquired (i) whether there were “public concerns” and (ii) whether those concerns were “moral”.²⁰

3.2.1.2 Necessity Test

Under an Article XX defence, the regulator must also demonstrate that its measure has the relevant “nexus” or “degree of connection” with the relevant public interest.

¹⁵ Fontanelli mentions *Mexico – Soft Drinks* as a rare exception to this rule; Filippo Fontanelli, ‘Necessity Killed the GATT - Art XX GATT and the Misleading Rhetoric about Weighing and Balancing’ (2012) 5 European Journal of Legal Studies 39, 50.

¹⁶ *ibid* 50.

¹⁷ Appellate Body Report, *Colombia Textiles* (n 7) para 5.68.

¹⁸ Panel Reports, *Brazil – Certain Measures Concerning Taxation and Charges*, WT/DS472/R, Add.1 and Corr.1, WT/DS497/R, Add.1 and Corr.1, adopted 11 January 2019, as modified by Appellate Body Reports WT/DS472/AB/R, WT/DS497/AB/R (*Brazil Taxation*) para 7.583.

¹⁹ *ibid* 7.558

²⁰ Appellate Body Reports, *European Communities – Measures Prohibiting the Importation and Marketing of Seal Products*, WT/DS400/AB/R / WT/DS401/AB/R, adopted 18 June 2014 (*Seals*). The *Seals* jurisprudence under the design step will be addressed in detail in chapter 8.

The ten sub-paragaphs of Article XX use different terms to describe the required “nexus”. In *Gasoline*, the Appellate Body noted that it “does not seem reasonable to suppose that the WTO Members intended to require, in respect of each and every category, the same kind or degree of connection”.²¹ This suggests that different general exceptions have a different standard of review depending on the “nexus” term employed.

The Appellate Body has categorised these different “nexus” terms as:

“necessary’ – in paragraphs (a), (b) and (d); ‘essential’ – in paragraph (j); ‘relating to’ – in paragraphs (c), (e) and (g); ‘for the protection of’ – in paragraph (f); ‘in pursuance of’ – in paragraph (h); and ‘involving’ – in paragraph (i).”²²

Under Articles XX(a)(b) and (d), a measure must be “necessary” and the second step is therefore described as a “necessity test”. In *Seals*, the Appellate Body described the necessity test as follows:

“a necessity analysis involves a process of "weighing and balancing" a series of factors, including the importance of the objective, the contribution of the measure to that objective, and the trade-restrictiveness of the measure. The Appellate Body has further explained that, in most cases, a comparison between the challenged measure and possible alternatives should then be undertaken”.²³

The necessity test can be divided into two complementary tests: the W&B and AM tests. These two tests form the core of the analysis in chapters 3-7.

3.2.2 Chapeau

The chapeau requires that:

²¹ Appellate Body Report, *Gasoline* (n 1) 17.

²² *ibid.*

²³ Appellate Body Report, *Seals* (n 20) para 5.169.

“measures are not 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”.²⁴

The Appellate Body has described the role of the chapeau as preventing “abuse of the exceptions”²⁵ and “but one expression of the principle of good faith”.²⁶ Measures which restrict trade “must be exercised *bona fide*” and “reasonably”.²⁷ The chapeau has been used in a variety of ways, including to ensure procedural fairness of administrative requirements²⁸ and to challenge measures with multiple (seemingly-contradictory) objectives.²⁹ The *Seals* dispute raises interesting chapeau questions on both these fronts but, due to limitations of space, this thesis will not address these questions.

3.2.3 Policy Objective

The policy objective (or regulatory goal/purpose) plays a key role throughout the Article XX analysis. Despite its important role, there is little jurisprudence clarifying how panels should characterise the objective. In *Seals*, adjudicators stated:

“A panel should take into account the Member's articulation of the objective or the objectives it pursues through its measure, but it is not bound by that Member's characterizations of such objective(s)”.³⁰

The Appellate Body has further stated that “a panel cannot... simply accept, without question, a Member's characterization of its measure”.³¹ In *Energy Supply*, the

²⁴ General Agreement on Tariffs and Trade 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1A (adopted 15 April 1994, entered into force 1 January 1995) 1867 UNTS 187 (“GATT”) Art XX (chapeau).

²⁵ Appellate Body Report, *Gasoline* (n 1) 22.

²⁶ ibid 22 (where it drew insight from the “drafting history of Article XX”). For a discussion of the purpose for (and legal test under) the chapeau, see Lorand Bartels, ‘The Chapeau of the General Exceptions in the WTO GATT and GATS Agreements: A Reconstruction’ (2015) 109(1) American Journal of International Law 95.

²⁷ Appellate Body Report, *Shrimp* (n 2) para 158.

²⁸ For example, in *Shrimp* (ibid), one of the reasons the US measure failed under the chapeau was the fact that it offered an insufficient (and discriminatory) compliance period to certain countries.

²⁹ In *Brazil Tyres*, one of the reasons Brazil's ban failed under the chapeau was that the justification for the exceptions had “rational connection” to the main policy objective (the so-called “rational connection” test; Appellate Body Report, *Brazil – Measures Affecting Imports of Retreaded Tyres*, WT/DS332/AB/R, adopted 17 December 2007 (*Brazil Tyres*) 89-90.

³⁰ Appellate Body Report, *Seals* (n 20) para 5.144.

panel noted that the policy objective should be defined with “a certain minimum level of clarity” but it also stated that “we do not believe that the required level of clarity should be overly demanding”.³² In that dispute, the panel found that the EU’s objective had been expressed with “the level of clarity required to assess its defence in a meaningful manner”.³³

In the absence of significant jurisprudence on this matter, some academics have sought to fill the gap. McGrady notes:

“the Appellate Body has previously held that a regulatory goal should be determined objectively and that a member’s characterization of its goal as evidenced by texts of statutes, legislative history, pronouncements of government agencies and officials may be taken into account. A panel may also find guidance in the structure and operation of the measure, in its effects and in any evidence offered by the complaining party to contradict the respondent’s stated objective”.³⁴

Mitchell endorses some deference towards Members when he suggests that their “assertions of regulatory purpose should be considered and given weight”.³⁵ However, he concludes that “regulatory purpose is ultimately a matter of evidence and judgement” for the adjudicator. Once the policy objective has been characterised, it plays a critical role for legal elements such as the threshold test, importance, contribution and the AM test.³⁶

³¹ Appellate Body Reports, *China – Measures Related to the Exportation of Rare Earths, Tungsten, and Molybdenum*, WT/DS431/AB/R / WT/DS432/AB/R / WT/DS433/AB/R, adopted 29 August 2014 (*Rare Earths*) para 5.95.

³² Panel Report, European Union and its member States – Certain Measures Relating to the Energy Sector, WT/DS476/R and Add.1, circulated to WTO Members 10 August 2018 [appealed by the European Union 21 September 2018 – the Division suspended its work on 10 December 2019] (*EU – Energy Supply*) para 7.1153.

³³ *ibid* paras 7.1152-7.1153.

³⁴ Benn McGrady, ‘Necessity Exceptions in WTO Law: Retreaded Tyres, Regulatory Purpose and Cumulative Regulatory Measures’ (2008) 12 *Journal of International Economic Law* 153, 156.

³⁵ Andrew Mitchell, David Heaton and Caroline Henckels, *Non-Discrimination and the Role of Regulatory Purpose in International Trade and Investment Law* (Elgar 2016) 97.

³⁶ While outside the scope of this analysis, it is worth noting that the characterisation of the policy objective is also relevant under the chapeau. For example, in *Seals* (n 20) para 5.133, the Appellate Body stated: “the Panel’s characterization of the objective of the measure has implications both in respect of the analysis under subparagraph (a), as well as under the chapeau”.

3.3 Brief History of the Necessity Test(s)

In GATT panels and early WTO disputes, adjudicating bodies adopted the AM test which focussed on identifying the least trade-restrictive (or treaty-inconsistent) means of achieving the regulator's objective.³⁷ In *Korea Beef*, the Appellate Body complemented this approach with the W&B test which is based on seeking an appropriate balance between the social utility of a given measure and its trade-restrictiveness. This section will provide a brief history of how the two tests evolved.

3.3.1 Alternative Measures Test

The meaning of “necessary” in Article XX was first addressed by GATT panels.³⁸ In *Section 337*, the US regime for patent enforcement was challenged as it had stricter timeframes, procedures and remedies for foreign patent infringers than domestic ones.³⁹ The US’ Article XX(d) defence was rejected as:

“[A] contracting party cannot justify a measure inconsistent with another GATT provision 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”.⁴⁰

The GATT panel found that the *Section 337* enforcement mechanism failed the AM test as the regular court mechanism (used for domestic infringers) was also reasonably available for foreign infringers.

³⁷ There are a number of abbreviations in the literature for this test including LTR (least trade-restrictive) and LTI (least treaty-inconsistent). I prefer calling it the AM test as it captures both the LTR and LTI approaches.

³⁸ For a general discussion of the pre-WTO arrangements for dispute settlement, see John Jackson, ‘Power and Rules in the Changing Economic Order: The Case of the World Trade Organization’ (2008) 84(3) International Affairs 437.

³⁹ While domestic infringers could only be sued in courts, foreign infringers could be sued either in court or under the *Section 337* mechanism.

⁴⁰ GATT Panel Report, *United States Section 337 of the Tariff Act of 1930*, L/6439, adopted 7 November 1989 (*Section 337*) para 5.26.

The next major dispute, *Thai Cigarette*, saw a challenge to Thai measure which had the practical effect of banning foreign tobacco products. Thailand's attempt to defend the measure for "public health" reasons under Article XX(b) failed as:

"the import restrictions imposed by Thailand could be considered to be "necessary" in terms of Article XX(b) only if there were no alternative measures consistent with the General Agreement, or less inconsistent with it, which Thailand could reasonably be expected to employ to achieve its health policy objectives".⁴¹

Thailand could have achieved its health objectives in a less trade-restrictive way by regulating foreign tobacco companies rather than excluding them entirely from its market.

Whereas, the *Section 337* panel was silent on whether the AM test was the only way of testing for necessity, the *Thai Cigarette* panel appeared to close the door to other methods. It stated that the necessity standard could be met "only if there were no alternative measures consistent with the General Agreement" or less treaty-inconsistent.

In the first dispute following the establishment of the WTO,⁴² *Gasoline*,⁴³ the US sought to defend the necessity of its measure (and implementation methods) for

⁴¹ Panel Report, *United States – Standards for Reformulated and Conventional Gasoline*, WT/DS2/R, adopted 20 May 1996, as modified by Appellate Body Report WT/DS2/AB/R (*Gasoline*) 40 (citing GATT Panel Report, *Thailand – Restrictions on Importation of and Internal Taxes on Cigarettes*, DS10/R, adopted 7 November 1990 (*Thai Cigarettes*)).

⁴² The WTO was established on 1 January 1995 under the Marrakesh Agreement Establishing the World Trade Organization (adopted 15 April 1994, entered into force 1 January 1995) 1867 UNTS 154 ("Marrakesh Agreement" or "WTO Agreement"). It marked a significant development for WTO adjudication as WTO dispute settlement had a dedicated agreement for the first time: Dispute Settlement Rules: Understanding on Rules and Procedures Governing the Settlement of Disputes, Marrakesh Agreement Establishing the World Trade Organization, Annex 2 (adopted 15 April 1994, entered into force 1 January 1995) 1869 UNTS 401 ("DSU"). GATT panels were replaced by WTO panels and, most significantly, the Appellate Body was established as a standing appeals mechanism.

⁴³ Strictly speaking, *Gasoline* (n 1) was the second WTO dispute, but it was the first that actually reached the adjudication stage.

gasoline emission standards under Article XX(b) and (d).⁴⁴ The US' policy objective on emission standards was not controversial and adjudicators confirmed the autonomy of the US "to regulate in order to obtain whatever air quality it wished".⁴⁵

However, there was disagreement regarding the necessity of the "baseline establishment rules" which discriminated between domestic and foreign firms for the purpose of calculating gasoline emission reductions. The Appellate Body endorsed the panel's use of the AM test to find that

"the baseline establishment rules had not been shown by the United States to be "necessary" under Article XX(b) since alternative measures either consistent or less inconsistent with the General Agreement were reasonably available to the United States for achieving its aim".⁴⁶

Together, these two GATT panels and one WTO dispute represent the pre-W&B era for necessity cases.⁴⁷ They were based on "an entrenched understanding that 'necessary' meant least-treaty-inconsistent-alternative-reasonably-available".⁴⁸

3.3.2 The (Gradual) Introduction of the W&B Test

In *Korea Beef*, the Appellate Body introduced a new test into WTO jurisprudence when it stated:

"determination of whether a measure... [is] "necessary" ... involves in every case a process of weighing and balancing a series of factors which prominently include the contribution made by the compliance measure to the enforcement of the law or regulation at issue, the importance of the common

⁴⁴ The US also raised a defence under Article XX(g) which I will not analyse as it was not based on a necessity standard.

⁴⁵ Appellate Body Report, *Gasoline* (n 1) 15 (citing the *Gasoline* panel).

⁴⁶ *ibid* 16.

⁴⁷ *Shrimp* included a defence under Article XX(b) but it was not analysed for reasons of judicial economy. The decisive analysis was conducted under Article XX(g) which does not have a necessity requirement. During the GATT era, there was also an unadopted GATT panel in *Tuna - Dolphin*. For a discussion of this dispute, see Alan Sykes, 'The Least Restrictive Means' (2003) 70 University of Chicago Law Review 403.

⁴⁸ Chad Bown and Joel Trachtman, 'Brazil – Measures Affecting the Import of Retreaded Tyres: A Balancing Act' (2009) 8 World Trade Review 85, 89.

interests or values protected by that law or regulation, and the accompanying impact of the law or regulation on imports or exports".⁴⁹

In *Korea Beef*, the Appellate Body merely mentioned the W&B test as *obiter dictum*.⁵⁰ In subsequent disputes, the W&B test was increasingly applied as a core part of the necessity test. In *Asbestos*, the Appellate Body expanded on the element of "importance". It stated:

"the objective pursued by the measure is the preservation of human life and health through the elimination, or reduction, of the well-known, and life-threatening, health risks posed by asbestos fibres. The value pursued is both vital and important in the highest degree".⁵¹

This quote may seem unremarkable, as it merely confirms the regulator's assessment that protection from asbestos is important (an intuitively obvious conclusion). However, no previous adjudicator had ever had the temerity (or perceived mandate) to make normative value judgments about the importance of a regulator's policy objective.

In *Asbestos*, the Appellate Body undertook no independent analysis of the contribution and trade-restrictiveness elements. After a cursory analysis of importance, it reverted to the familiar AM test and concluded that France's measure was necessary. It determined that the proposed alternative (controlled use of products containing asbestos) would achieve a lower level of protection than France was seeking.

After presenting the W&B test rhetorically (in *Korea Beef*) and analysing the importance element (in *Asbestos*), the next major developments occurred in

⁴⁹ Appellate Body Report, *Korea – Measures Affecting Imports of Fresh, Chilled and Frozen Beef*, WT/DS161/AB/R, WT/DS169/AB/R, adopted 10 January 2001 (*Korea Beef*) para 164.

⁵⁰ This is my interpretation as adjudicators did not appear to analyse any of the elements of the W&B test.

⁵¹ Appellate Body Report, *European Communities – Measures Affecting Asbestos and Asbestos-Containing Products*, WT/DS135/AB/R, adopted 5 April 2001 (*Asbestos*) para 172.

Gambling which “laid out the detailed procedural steps”⁵² and *Brazil Tyres* which introduced some significant analysis of contribution and trade-restrictiveness.⁵³

Though there have been some minor refinements, the W&B test as enunciated in *Gambling* and *Brazil Tyres* provides the basic framework for current adjudication.

3.3.3 Summary

During the GATT and early WTO years, adjudicators introduced the AM test in order to determine the meaning of necessity under Article XX.⁵⁴ In *Korea Beef*, the Appellate Body introduced the W&B test as a further method of necessity testing and gradually started applying this test in subsequent cases. The next chapter will undertake a doctrinal and normative critique of both tests.

⁵² Michael Ming Du, ‘Autonomy in Setting Appropriate Level of Protection under the WTO Law: Rhetoric or Reality?’ (2010) 13 Journal of International Economic Law 1077, 1092.

⁵³ Though, as I will argue below, *Brazil Tyres* failed to provide detailed guidance on the “importance” or “balancing” elements.

⁵⁴ Especially Articles XX(b) and (d).

CHAPTER 4 - THE W&B AND AM TESTS: DOCTRINAL ANALYSIS AND CRITIQUE

4.1 Introduction

This chapter will provide a doctrinal analysis and critique of the W&B and AM tests.

For each test, it will explain how the test has evolved in WTO jurisprudence, including any aspects of the tests which are vague or contradictory. It will further describe the lack of clarity about how the two tests interact.¹

4.2 W&B Test: Doctrinal Analysis

The W&B test analyses three key factors: “the importance of the objective, the contribution of the measure to that objective, and the trade-restrictiveness of the measure” (abbreviated to “importance”, “contribution” and “trade-restrictiveness”).²

Further, it entails a balancing step to see if the measure’s social utility outweighs its adverse trade effects.

4.2.1 Importance

Although, it is a key element, adjudicators provide very little guidance on the methodology for assessing “importance”.³ Further, there is little clarity about the object of review. In different disputes, adjudicators have assessed the importance of the “policy objective”, the “societal interest” or the “public interest”. These shifts in language have had a real impact on the substantive analysis.

¹ This chapter focusses on a doctrinal critique of the W&B and AM tests. Subsequent chapters will undertake a theoretical critique of the appropriateness of W&B and AM testing (chapter 5), a “deconstruction” of how the application of these tests deviates from the theory (chapters 6 and 7) and broader socio-legal analysis of judicial decision-making (chapter 7). For ease of analysis, I have separated each family of critique into separate chapters.

² Appellate Body Reports, *European Communities – Measures Prohibiting the Importation and Marketing of Seal Products*, WT/DS400/AB/R / WT/DS401/AB/R, adopted 18 June 2014 (*Seals*) para 5.169.

³ In contrast to contribution and trade-restrictiveness where adjudicators have established detailed methodologies (described further below).

In *Audiovisuals*, the panel selected the public interest (public morals) as the object of review. It stated that "the protection of public morals ranks among the most important values or interests pursued by Members as a matter of public policy".⁴ It justified this view by stating:

"We do not consider it simply accident that the exception relating to "public morals" is the first exception identified in the ten sub-paragraphs of Article XX. We therefore concur that the protection of public morals is a highly important value or interest".⁵

This suggests that the W&B test requires panels to assess (and rank) the importance of the ten public interests in Article XX. Why should panels create a hierarchy among these interests and base it on the order in which they're listed? According to this reasoning, "importations or exportations of gold or silver" (Article XX(c)) would seem more important than "exhaustible natural resources" (Article XX(g)) because it appears higher in the Article XX list. Shouldn't adjudicators respect the express guidance of GATT drafters and accept that all of the listed public interests are sufficiently important to justify trade-restrictive measures? It seems absurd that a panel would seek to justify the importance of "public morals" (as a general category) rather than the specific moral norm under dispute.

This strange approach to "importance", which was tacitly accepted by the Appellate Body in *Audiovisuals*, has been criticised by subsequent adjudicators. In *Brazil Taxation*, the panel considered that:

"in determining the importance of a particular objective, it is more pertinent to assess the importance of the particular societal interest being protected,

⁴ Appellate Body Report, *China – Measures Affecting Trading Rights and Distribution Services for Certain Publications and Audiovisual Entertainment Products*, WT/DS363/AB/R, adopted 19 January 2010 (*China Audiovisuals*). para 243.

⁵ Panel Report, *China – Measures Affecting Trading Rights and Distribution Services for Certain Publications and Audiovisual Entertainment Products*, WT/DS363/R and Corr.1, adopted 19 January 2010, as modified by Appellate Body Report WT/DS363/AB/R (*China Audiovisuals*) 7.817.

rather than assuming that by virtue of its status as a 'public moral' objective the interest is per se vital or important to the highest degree".⁶

While the *Brazil Taxation* panel correctly chastised the approach in *Audiovisuals*, it was not immune from holding confused views about the object of review. In the paragraph cited above, the panel uses two different objects of review by referring to the importance of both the "objective" and the "societal interest".

This reflects a general confusion in the way the legal test has been enunciated by adjudicators over the years. For example, in the original formulation in *Korea Beef*, the Appellate Body referred to "the importance of the common interests or values". In other cases, such as *Seals*, the Appellate Body referred to the "importance of the objective".

One explanation could be that the terms "objective" and "societal interest" are interchangeable, but this view seems implausible. As discussed above, the policy objective is a central concept which must be defined and characterised by adjudicators in every dispute. The "societal interest", by contrast, is an amorphous and undefined concept. Its use seems confusing and redundant.

In *Colombia Textiles*, the Appellate Body arguably reconciled this confusing language when it found that "the objective of combating money laundering reflects societal interests that can be described as vital and important in the highest degree."⁷ Under this formulation, it is the "societal interest" which must be important, while the "objective" must merely "reflect" it. In my view, this formulation fails to

⁶ Panel Reports, *Brazil – Certain Measures Concerning Taxation and Charges*, WT/DS472/R, Add.1 and Corr.1, WT/DS497/R, Add.1 and Corr.1, adopted 11 January 2019, as modified by Appellate Body Reports WT/DS472/AB/R, WT/DS497/AB/R (*Brazil Taxation*) para 7.591.

⁷ Appellate Body Report, *Colombia – Measures Relating to the Importation of Textiles, Apparel and Footwear*, WT/DS461/AB/R and Add.1, adopted 22 June 2016 (*Colombia Textiles*) para 5.105.

resolve the underlying ambiguity about the object of review. The *Brazil Tyres* case study will show that this ambiguity can have a serious bearing on a dispute.

4.2.2 Contribution and Trade Restrictiveness

This section will describe the methodology for analysing contribution and trade-restrictiveness. Since they share much in common, I will address them as a single package and, where necessary, note any key differences between them.⁸

Adjudicators must assess “the extent of the measure's contribution to the end pursued, rather than merely ascertaining whether or not the measure makes any contribution.”⁹ The same is true for trade-restrictiveness, where adjudicators must assess “the degree of a measure's trade-restrictiveness”.¹⁰ The analysis of these elements requires an adjudicator to answer the question “how much” or “to what degree”.

In this sense, contribution and trade-restrictiveness are not pass/fail tests.¹¹ Rather, adjudicators must quantify these concepts. The Appellate Body has stated, “the greater the contribution, the more easily a measure might be considered to be 'necessary'.”¹² However, there are no fixed thresholds: “the flexibility of such an exercise does not allow for the setting of pre-determined thresholds in respect of any particular factor.”¹³

In *Brazil Tyres*, adjudicators found that an import ban was highly trade-restrictive and that the measure must therefore make a “material contribution”:

⁸ In theory, there is no reason why the “importance” analysis should not also be subject to the considerations discussed below. However, I will show in the next chapter that, in practice, adjudicators tend to address “importance” in a cursory manner, rather than applying a rigorous methodology.

⁹ Appellate Body Report, *Colombia Textiles* (n 7) 5.72.

¹⁰ *ibid* 5.73.

¹¹ Unlike the design step.

¹² Appellate Body Report, *Colombia Textiles* (n 7) 5.72.

¹³ Appellate Body Report, *Seals* (n 2) 5.215.

“when a measure produces restrictive effects on international trade as severe as those resulting from an import ban, it appears to us that it would be difficult for a panel to find that measure necessary unless it is satisfied that the measure is apt to make a material contribution to the achievement of its objective”.¹⁴

In *Seals*, the complainants argued that the import ban should be required to make a “material contribution” as had been the case in *Brazil Tyres*. The Appellate Body rejected the notion that there was a “generally applicable pre-determined threshold” such as “material contribution”.¹⁵ It accepted the panel’s use of a lower threshold:

“we do not consider that the Panel erred in concluding that the EU Seal Regime “is capable of making and does make some contribution” to its objective, or that it makes a contribution “to a certain extent”. ”¹⁶

Since both *Seals* and *Brazil Tyres* involved the same level of trade-restrictiveness (import bans), does the lower contribution threshold in *Seals* suggest that the objective (seal welfare) was more important?

While adjudicators must determine the extent/degree of contribution and trade restrictiveness, they are not obliged to do so in numerical terms. They need not quantify that a measure saves X lives, produces X “utils” or restricts \$X of trade. Some adjudicators have called for greater clarity in the analysis, suggesting that “a determination of the trade-restrictiveness of a particular measure should be as precise as possible”.¹⁷ However, in practice, the W&B factors tend to be assessed “impressionistically and qualitatively”.¹⁸

¹⁴ Appellate Body Report, *Brazil – Measures Affecting Imports of Retreaded Tyres*, WT/DS332/AB/R, adopted 17 December 2007 (*Brazil Tyres*) para 150.

¹⁵ Appellate Body Report, *Seals* (n 2) para 5.216.

¹⁶ ibid para 5.228.

¹⁷ Panel Report, *Brazil Taxation* (n 6) para 7.607.

¹⁸ Alan Sykes, 'The Least Restrictive Means' (2003) 70 University of Chicago Law Review 403, 415.

The Appellate Body has confirmed that panels “should enjoy a certain latitude in designing the appropriate methodology” for the contribution analysis.¹⁹ It has further confirmed that adjudicators can choose to adopt “a qualitative or quantitative manner” for both the contribution and trade-restrictiveness elements.²⁰ Even though panels are offered some discretion on methodology, the Appellate Body has cautioned that their “latitude” is not “boundless”.²¹

The Appellate Body has instructed panels to “always assess the actual contribution made by the measure to the objective pursued”,²² but this does not mean that the contribution must have already occurred. Where the contribution of the measure is not yet “immediately observable”, a panel may conclude that a measure is “apt to produce a material contribution to the achievement of its objective”.²³

The “actual contribution” requirement is intended to preclude assessments based on a desired or aspirational contribution. Adjudicators must assess what is likely to happen not what regulators hope will happen. They can rely on “quantitative projections in the future” or “qualitative reasoning based on a set of hypotheses that are tested and supported by sufficient evidence.”²⁴

There have been rare cases where adjudicators have been unable to analyse certain W&B factors due to “a lack of sufficient clarity” or data in the factual record.²⁵ In *Colombia Textiles*, the Appellate Body noted:

¹⁹ Appellate Body Report, *Brazil Tyres* (n 14) para 145.

²⁰ Appellate Body Report, *Colombia Textiles* (n 7) para 5.72. The Appellate Body first established the notion that panels can use qualitative or quantitative reasoning in the context of assessing “risk” in disputes such as *Asbestos*.

²¹ Appellate Body Report, *Brazil Tyres* (n 14) para 145.

²² Appellate Body Report, *Audiovisuals* (n 4) para 252.

²³ Appellate Body Report, *Brazil Tyres* (n 14) para 151.

²⁴ *ibid.*

²⁵ Appellate Body Report, *Colombia Textiles* (n 7) para 5.116.

“Without sufficient clarity in respect of these factors, a proper weighing and balancing that could yield a conclusion that the measure is 'necessary' could not be conducted”.²⁶

In *Colombia Textiles*, a lack of data hampered the regulator's defence and prevented it from demonstrating “necessity”. In other disputes, adjudicators have overcome a lack of meaningful data and proactively used qualitative reasoning to fill in the gaps.

In *Seals*, the Appellate Body recognised that the panel report “does not provide much information as to the precise degree or extent of the contribution” and that its analysis was “not very detailed”.²⁷ However, it recognised “significant limitations” regarding “the nature, quantity, and quality of evidence” so it ultimately accepted the panel’s vague conclusion that the Seal Regime “is capable of making and does make some contribution” to seal welfare.²⁸

The Appellate Body accepted the panel’s qualitative approach (in the absence of definitive data) and stated that “it is not clear what greater clarity or precision the Panel could have achieved in the circumstances of this case”.²⁹

Adjudicators appear to have a choice when faced with poor data on contribution (and trade restrictiveness). They can either refuse to make findings (to the detriment of the regulator’s defence) or they can use qualitative reasoning to reach the best possible conclusion, even in trying conditions. The EU was the beneficiary of proactive adjudicators in *Seals*, whereas Colombia suffered at the hands of a panel which seemingly lacked confidence in its capacity for qualitative reasoning.

The methodology for assessing contribution and trade-restrictiveness is critical. They represent two key factors under the W&B test. Further, they both serve as

²⁶ Appellate Body Report, *Colombia Textiles* (n 7) para 5.116.

²⁷ Appellate Body Report, *Seals* (n 2) para 5.228.

²⁸ *ibid.*

²⁹ *ibid.*

benchmarks under the AM test, since any proposed alternative measure must make an “equivalent contribution” to the objective while being less “trade-restrictive”.

4.3 AM Test: Doctrinal Analysis

Under the AM test, the onus is on the complainant to “identify any alternative measures that, in its view, the responding party should have taken”.³⁰ The proposed alternative should be “less trade restrictive”, make an “equivalent contribution” and be “reasonably available”. This section will consider each of these three conditions.

4.3.1 Less Trade-Restrictive (or Less Treaty-Inconsistent)

In early disputes, GATT panels and the *Gasoline* adjudicators relied on a “less treaty-inconsistent” standard.³¹ Since *Korea Beef*, adjudicators have tended to apply a “less trade-restrictive” standard.³²

Weiler is critical that the Appellate Body is “extraordinarily loose with its own use of language”,³³ however Sykes seems unperturbed, explaining away the different approaches as “just a linguistic variant”.³⁴ The “less trade-restrictive” standard appears to have a major advantage as it enables adjudicators to use their trade-restrictiveness analysis from the W&B test as a benchmark for testing the proposed alternative.

³⁰ Appellate Body Report, *Seals* (n 2) para 5.169 (citing *Gambling*).

³¹ Appellate Body Report, *United States – Standards for Reformulated and Conventional Gasoline*, WT/DS2/AB/R, adopted 20 May 1996 (*Gasoline*) 16 and 26-27. Adjudicators found that any proposed alternative should be “either consistent or less inconsistent with the General Agreement”.

³² However, the transition from LTI to LTR has not been seamless. In *Brazil Tyres*, for example, the panel reverted to a “less treaty-inconsistent” standard when it considered “a measure, other than the import ban, that is compatible, or less incompatible, with the WTO Agreement”; Panel Report, *Brazil – Measures Affecting Imports of Retreaded Tyres*, WT/DS332/R, adopted 17 December 2007 (*Brazil Tyres*) para 7.157.

³³ JHH Weiler, ‘Comment on Brazil - Measures Affecting Imports of Retreaded Tyres’ (2009) 8 World Trade Review 137, 138.

³⁴ Sykes (n 18) 406.

Under a “less treaty inconsistent” standard, adjudicators would have to quantify the level of treaty-inconsistency of the contested measure to serve as a point of comparison for the proposed alternative. Kapterian rightly questions whether it is even possible to measure or quantify the level of GATT-inconsistency.³⁵

That said, some scholars also question whether trade-restrictiveness is really capable of being quantified. Voon argues that “no single definition of trade-restrictiveness exists” which undermines its utility as a standard.³⁶ She describes it as a “black box” where adjudicators “are free to inject into trade-restrictiveness whatever assessments and judgments they see fit.”³⁷

The issue of whether a proposed alternative is less trade-restrictive generally raises little controversy. Even if Voon is correct that there is no clear way to assess the magnitude of trade-restrictiveness, complainants have a strong interest in only proposing alternatives which are less trade-restrictive.³⁸ While the “less trade-restrictive” approach might not be perfect, it appears to be more suitable than the “less treaty-inconsistent” standard.

4.3.2 “Equivalent Contribution” (or “Desired Level of Protection”)

Adjudicators have applied two different benchmarks regarding how effective the proposed alternative must be. Sometimes, the Appellate Body searches for

³⁵ Gisele Kapterian, ‘A Critique of the WTO Jurisprudence on ‘Necessity’’ (2010) 59 International & Comparative Law Quarterly 89, 105.

³⁶ Tania Voon, ‘Exploring the Meaning of Trade-Restrictiveness in the WTO’ (2015) 14(3) World Trade Review 451, 451.

³⁷ *ibid* 453.

³⁸ However, the situation can be more complicated when complainants oppose the discriminatory aspects of a measure (as opposed to its trade-restrictiveness). In order to remove discrimination, regulators can choose to either reduce the regulatory burden on imports (thus making the measure less trade-restrictive) or increase the regulatory burden on domestic products (thus making the measure more trade-restrictive). This situation occurred in *Clove Cigarettes* where the panel (somewhat mischievously) characterised Indonesia as arguing that the contested measure was “less trade-restrictive than necessary to fulfil its objective”; Panel Report, *United States – Measures Affecting the Production and Sale of Clove Cigarettes*, WT/DS406/R, adopted 24 April 2012 (*US – Clove Cigarettes*) para 7.395.

alternatives which provide “an *equivalent contribution* to the achievement of the objective.”³⁹ This language (“equivalent contribution”) strongly suggests that the benchmark can be drawn directly from the “actual contribution” analysis, under the W&B test, where adjudicators establish what the contested measure actually achieves.⁴⁰

However, the Appellate Body has also suggested that the proposed alternative must “preserve for the responding Member its right to achieve its desired level of protection with respect to the objective pursued”.⁴¹ This language (“desired level of protection”) suggests that the benchmark is based on the regulator’s aspirations, rather than the contested measure’s “actual contribution”. This would require adjudicators to construct a new benchmark for the AM test based on the regulator’s desires.

Why are there two different ways of expressing the benchmark? In my view, this can be explained by looking to the evolution of the two necessity tests and, especially, the introduction of the W&B test. The “level of protection” has been part of WTO law since GATT panels first applied the AM test. However, the regulator’s “objective” and the measure’s “contribution” only became legal elements in the WTO era when they were introduced under the W&B test.⁴²

For GATT panels, the concepts of “objective” and “contribution” existed as conceptual notions in the background, but they were not explicit legal elements to be

³⁹ Appellate Body Report, *Brazil Tyres* (n 14) para 178.

⁴⁰ The panel takes a similar approach in *China Audiovisuals* when it discusses “the *contribution* that the alternative would make to the objective of protecting public morals (“at least *equivalent to*”). (Panel, *Audiovisuals*, para 246).

⁴¹ Appellate Body Report, *Brazil Tyres* (n 14) para 156.

⁴² For the purpose of comparative analysis, it is useful to note that the concept of “contribution” seems highly similar to the concept of “effectiveness” which is sometimes used in US Supreme Court cases; see eg Trachtman J, ‘Trade and... Problems, Cost-Benefit Analysis and Subsidiarity’ (1998) 9 European Journal of International Law 32, 47.

analysed. Under the AM test, adjudicators had no obligation to explicitly characterise the objective or to determine the measure's actual contribution.

This point can be illustrated by the *Gasoline* dispute where the panel searched for an alternative measure which would "allow the United States to achieve its desired level of clean air without discriminating against imported gasoline".⁴³ The benchmark for testing proposed alternatives was expressed by reference to an aspirational level of protection ("desired level of clean air"). The panel was not required by WTO jurisprudence (as formulated at that time) to characterise the US' policy objective even though it was strongly implicit in the level of protection: the US' objective was to improve/protect air quality.

The panel looked at the US' "desired level of clean air". In the absence of the W&B test, the panel had not been required to establish the "actual contribution" of the contested measure and therefore constructed an aspirational level of protection to serve as the benchmark. This seems like a poor analytical approach. Why didn't they determine the contribution which the measure was apt to make to air quality?

The artifice of constructing a "desired level of protection" has become redundant. The W&B test requires adjudicators to characterise the policy objective and quantify the actual contribution of the measure to that objective. This analysis is best placed to serve as the benchmark for the AM test.

However, in practice, two methodologies remain on the books: "equivalent contribution" and "desired level of protection". This might just be a case of "loose"

⁴³ Panel Report, *United States – Standards for Reformulated and Conventional Gasoline*, WT/DS2/R, adopted 20 May 1996, as modified by Appellate Body Report WT/DS2/AB/R (*Gasoline*) para 6.27.

language but, as will be demonstrated in the *Brazil Tyres* case study, this unclear jurisprudence can lead adjudicators to apply two substantively different standards.

In the context of *Korea Beef*, Regan explains why the “equivalent contribution” approach is superior:

“Korea cannot assert a ‘desired level of protection’ that is higher than the level achieved by the actual measure and then hold proposed alternatives to the standard Korea asserts as opposed to the standard its actual measure achieves”.⁴⁴

The appropriate benchmark for the AM test is “the contested measure’s contribution to the objective”. This means there are two elements (“contribution” and “objective”) which work together to establish a single benchmark. Both elements are questions of fact which should have been clearly established by adjudicators prior to undertaking the AM analysis.

4.3.3 Reasonable Availability

Any proposed alternative must be “reasonably available” to the regulator. This notion originated from the Section 337 GATT panel which only considered alternatives which regulators “could reasonably be expected to employ”.⁴⁵

A proposed alternative can be available even if it exposes the regulator to “some change or administrative cost”.⁴⁶ In *Asbestos*, the Appellate Body noted that “an alternative measure did not cease to be 'reasonably' available simply because the alternative measure involved administrative difficulties for a Member.”⁴⁷

⁴⁴ Don Regan, 'The Meaning of 'Necessary' in GATT Article XX and GATS Article XIV: The Myth of Cost-Benefit Balancing' (2007) 6 *World Trade Review* 347, 359.

⁴⁵ GATT Panel Report, *United States Section 337 of the Tariff Act of 1930*, L/6439, adopted 7 November 1989 (*Section 337*) para 5.26.

⁴⁶ Appellate Body Report, *Audiovisuals* (n 4) para 327.

⁴⁷ Appellate Body Report, *European Communities – Measures Affecting Asbestos and Asbestos-Containing Products*, WT/DS135/AB/R, adopted 5 April 2001 (*Asbestos*) para 169.

However, there are limits to the resource burden which proposed alternatives can impose. An alternative measure is not “reasonably available” if it imposes an “undue burden” through “prohibitive costs or substantial technical difficulties.”⁴⁸ This issue will be considered further in the *Korea Beef* case study where “reasonable availability” was a key issue.

4.4 Interaction between the Tests

When the W&B test was introduced in *Korea Beef*, the Appellate Body claimed that the necessity test “encapsulates the general considerations” of the W&B test.⁴⁹ It suggested “the weighing and balancing process we have outlined is comprehended” by the AM test.⁵⁰ This is a dubious claim as, prior to *Korea Beef*, there was no jurisprudence which required “balancing” by adjudicators. It seems more likely that the W&B test represented a major shift, rather than a method which was lying dormant within the AM test.

In *Audiovisuals*, the Appellate Body appeared to re-write this early jurisprudence when it stated:

“as part of an overall evaluation of “necessity” using the “weighing and balancing” process, a panel must examine whether the responding party could reasonably be expected to employ an alternative measure”.⁵¹

This inverted the hierarchy between the tests. In *Korea Beef*, the W&B test was “encapsulated” or “comprehended” within the AM test. By *Audiovisuals*, the W&B test was driving the “overall evaluation” and the AM test was merely “part of” it.⁵²

⁴⁸ Appellate Body Report, *United States – Measures Affecting the Cross-Border Supply of Gambling and Betting Services*, WT/DS285/AB/R, adopted 20 April 2005 (*US – Gambling*) para 308.

⁴⁹ Appellate Body Report, *Korea – Measures Affecting Imports of Fresh, Chilled and Frozen Beef*, WT/DS161/AB/R, WT/DS169/AB/R, adopted 10 January 2001 (*Korea Beef*) para 166.

⁵⁰ *ibid*.

⁵¹ Appellate Body Report, *Audiovisuals* (n 4) para 318.

⁵² This represents a significant shift from the Appellate Body’s hesitant reasoning in *Korea Beef* that the W&B test was subsumed within the pre-existing AM test. Both claims are equally false, but the new way of

This distinction is not highly significant for substantive analysis of the tests, but it is revealing of the stealthy way the W&B test entered, and became entrenched within, WTO jurisprudence.

In *Audiovisuals*, the Appellate Body suggested both tests were part of a "single, integrated, yet multifaceted inquiry".⁵³ However, it also kept the two tests separate by describing the necessity test as a "two-step process".

The Appellate Body has suggested that if a measure passes the W&B test, "this result must be confirmed by comparing the measure with possible alternatives".⁵⁴ This statement clarifies the sequencing of the two tests (W&B, followed by AM), but it tells us little about their substantive relationship. By saying that the W&B analysis is "confirmed" by the AM test, does this imply that they must produce the same result?

The Appellate Body has also introduced jurisprudence suggesting that it is not always compulsory to undertake both steps. In *Solar Cells*, it stated that "in most cases, a comparison between the challenged measure and reasonably available alternative measures should then be undertaken."⁵⁵ This suggests there may be cases where the W&B analysis is decisive and no AM analysis is required.

Where both steps are undertaken, is it not redundant or "inutile" to have two different tests which merely confirm the same result? The Appellate Body has often recalled "the internationally recognized interpretive principle of effectiveness" which requires that "provisions of the *WTO Agreement* should not be interpreted in such a manner

describing the test may reflect a growing courage by the Appellate Body to promote the legitimacy of the W&B test. Whereas it once tried to hide the W&B test (within the AM test, in obiter dicta), the W&B test has now come to the forefront, at least as a matter of discourse.

⁵³ Appellate Body Report, *Audiovisuals* (n 4) para 237.

⁵⁴ Appellate Body Report, *Audiovisuals* (n 4) para 241 (citing *Brazil Tyres*).

⁵⁵ Appellate Body Report, *India – Certain Measures Relating to Solar Cells and Solar Modules*, WT/DS456/AB/R and Add.1, adopted 14 October 2016 (*Solar Cells*) para 5.59.

that whole clauses or paragraphs of a treaty would be reduced to redundancy or inutility.”⁵⁶

Shouldn’t the same principle apply to legal tests developed by the Appellate Body? If two tests merely confirm the same outcome, this creates redundancy. If the tests produce different results, adjudicators should clarify the comparative advantage of each test. Why are there two tests? In the case of conflicting results, which one prevails? The relationship between the tests is poorly explained. As we will see in the next section, even though both tests are aimed at determining “necessity”, they are each based on a very different theoretical foundation.

4.5 Summary of Doctrinal Critique

The doctrinal analysis reveals that some aspects of the W&B and AM tests are ambiguous, contradictory or poorly-clarified. This occurs with respect to such issues as the object of review for the importance analysis (“policy objective” or “societal value”) and the AM benchmark (actual contribution or desired level of protection).

This unclear use of terminology and concepts receives scant attention from WTO scholars even though it can seriously undermine the coherence of the adjudicators’ legal analysis.⁵⁷ Weiler notes that the Appellate Body is “extraordinarily loose with its own use of language”.⁵⁸ He sees this “textual imprecision” as a failure by the Appellate Body to perform one of its fundamental roles: “develop over the years a

⁵⁶ Appellate Body Report, *United States – Continued Dumping and Subsidy Offset Act of 2000*, WT/DS217/AB/R, WT/DS234/AB/R, adopted 27 January 2003 (*Byrd Amendment*) para 271.

⁵⁷ Diebold is one of the rare scholars who has proactively sought to bring some clarity to key terms and concepts within the necessity test. For example, he critiques the fact that the terms “policy objective” and “public interest” are often used interchangeably even though they are substantively different. For a further discussion, see Nicolas Diebold, ‘The Morals and Order Exceptions in WTO Law: Balancing the Toothless Tiger and the Undermining Mole’ (2007) 11(1) Journal of International Economic Law 43.

⁵⁸ Weiler (n 33) 138.

coherent jurisprudence with each new case offering a chance of honing, self-correcting, and at times even changing course".⁵⁹

WTO legal tests and standards should gradually become clearer over time, but Weiler is doubtful that the Appellate Body is achieving this. Doyle is equally critical:

"By the time *Brazil Tyres* was decided, the "necessary" analysis had shown its truly convoluted nature. With vague, interrelated, and overlapping elements, the analysis had become almost free-form. For the Panel and Appellate Body, there exists a fundamental problem of establishing a clear and consistent standard for use in future cases".⁶⁰

This chapter has sought to illustrate the main doctrinal controversies within the W&B and AM tests. The next chapter will focus on the theoretical rationale for each test and the debate about which approach strikes the right balance between regulatory autonomy and substantive obligations.

⁵⁹ ibid.

⁶⁰ Christopher Doyle, 'Gimme Shelter: The Necessary Element of GATT Article XX in the Context of the Audiovisuals Products Case (2011) 29 Boston University International Law Journal 143, 160.

CHAPTER 5 - THEORETICAL CRITIQUE

5.1 Introduction

This section will undertake a theoretical critique of the W&B and AM tests to see if they are consistent with the “spirit” of the WTO. Critiquing the theoretical foundation of a legal test (while deferring the analysis of its actual application) may seem like an artificial exercise. After all, legal tests do not exist merely as a theoretical construct; rather, the nature and meaning of a legal test is ordinarily discernible from its actual application in disputes.

However, as I will show below, many WTO scholars have used this method to critique the W&B test. Since this type of critique does not look to the application of the legal test, it relies instead on a conceptual notion of what the test (as formulated) strives to achieve. In this context, the W&B test is often compared to a search for Kaldor-Hicks efficiency or to trade-off devices from the cost-benefit analysis (CBA) family. The AM test, on the other hand, is most commonly compared to a search for Pareto efficiency gains.

5.2 W&B Test (and Kaldor-Hicks Efficiency)

Many WTO scholars treat the W&B test as a type of CBA and argue that such tests have no place in WTO law. Regan states that “judicial review by cost–benefit balancing is not in the spirit of the WTO”¹ and “a serious intrusion on the Members’ regulatory autonomy”.²

¹ Don Regan, ‘The Meaning of ‘Necessary’ in GATT Article XX and GATS Article XIV: The Myth of Cost-Benefit Balancing’ (2007) 6 World Trade Review 347, 366.

² ibid 349.

Weiler is equally critical, noting “the Appellate Body would seem to be violating that very cardinal principle, regulatory autonomy, on which it repeatedly insists”.³ Ortino concurs that the W&B test is “more intrusive in WTO Members' regulatory autonomy” than the AM test.⁴

Fontanelli considers that CBA devices “are a prerogative of constitutional adjudication and are at variance with the negative integration paradigm” of the WTO.⁵ Bown and Trachtman consider a CBA approach “anathema to judicial restraint and national sovereignty”.⁶ Scholars are particularly critical of those legal elements which require value judgments about Members' regulatory choices, such as importance and the balancing step.

On “importance”, Regan states that “there is nothing to suggest that it is appropriate for the Appellate Body to rank Members' regulatory purposes according to the Appellate Body's intuitions about their value.”⁷ Ming Du concurs that “there is no textual warrant for such judgments of importance by the AB” and that “it is a serious intrusion on the Members' regulatory autonomy.”⁸

Howse suggests the WTO “is not well equipped to make substantive judgments about the moral choices” of its Members.⁹ Regan suggests that “the real question should not be about the importance of the purpose in the abstract, or by the

³ JHH Weiler, 'Comment on Brazil - Measures Affecting Imports of Retreaded Tyres' (2009) 8 World Trade Review 137, 141.

⁴ Federico Ortino, 'GATT' in Daniel Bethlehem, Isabelle Van Damme, Donald McRae, and Rodney Neufeld (eds), *The Oxford Handbook of International Trade Law* (OUP 2009) 142.

⁵ Fontanelli F, 'Necessity Killed the GATT - Art XX GATT and the Misleading Rhetoric about Weighing and Balancing' (2012) 5 European Journal of Legal Studies 39, 58.

⁶ Chad Bown and Joel Trachtman, 'Brazil – Measures Affecting the Import of Retreaded Tyres: A Balancing Act' (2009) 8 World Trade Review 85, 119.

⁷ Regan (n 1) 356.

⁸ Michael Ming Du, 'Autonomy in Setting Appropriate Level of Protection under the WTO Law: Rhetoric or Reality?' (2010) 13 Journal of International Economic Law 1077, 1101.

⁹ Robert Howse, Joanna Langille and Katie Sykes, 'Pluralism in Practice: Moral Legislation and the Law of the WTO after Seal Products' 48 George Washington International Law Review (2015) 81, 86.

Appellate Body's standards; it should be about the importance of the purpose to the Member whose regulation is in question.”¹⁰

The balancing step is equally problematic. If an adjudicator rejects a measure under the W&B test, it would essentially be “substituting its value judgment” for that of the regulator.¹¹ As Charnovitz notes, “the WTO has no institutional competence for weighing incommensurate values, such as the exporting interests of one country against the environmental interests of another”.¹²

In his detailed comparative study of “trade-off devices” (in the US, EU and WTO), Trachtman identified at least three devices which fall within the CBA family:¹³

CBA “juxtaposes the regulatory benefits and the trade costs of regulation, as well as other costs involved, and would strike down regulation where the costs exceed the benefits.”¹⁴ This approach “necessarily involves the comparison of differently denominated values, such as free trade versus environmental protection.”¹⁵

Proportionality “might be viewed as cost-benefit analysis with a ‘margin of appreciation’, as it does not require that the costs be less than the benefits.”¹⁶ In this sense, it provides greater deference to local regulation”.¹⁷ In EU law, this type of mechanism is known as *proportionality stricto sensu*.¹⁸

Balancing can be understood as an “amorphous or imprecise cost-benefit analysis... that recognizes the difficulty of formalizing the analysis, and seeks to achieve similar results informally”.¹⁹ In this way, it “avoids the claims to precision” and “the apparent arrogance” of seeking to quantify different social values. This is consistent with the Appellate Body’s “qualitative” approach of

¹⁰ Regan (n 1) 352.

¹¹ Weiler (n 3) 140.

¹² Charnovitz S, ‘The Law of Environmental “PPMs” in the WTO: Debunking the Myth of Illegality’ (2002) 27 Yale Journal of International Law 59, 101.

¹³ Trachtman J, ‘Trade and... Problems, Cost-Benefit Analysis and Subsidiarity’ (1998) 9 European Journal of International Law 32. Trachtman also discusses “comparative CBA” which would be the most intrusive (and efficient) trade-off device. Since this device is not used in any of the systems he studies (and serves as a theoretical point of comparison), I will only discuss it in passing.

¹⁴ ibid 36.

¹⁵ ibid 56.

¹⁶ ibid 35.

¹⁷ ibid 76.

¹⁸ For a discussion of proportionality in EU law, see Benedikt Pirker, *Proportionality Analysis and Models of Judicial Review: A Theoretical and Comparative Study* (Europa Law Publishing 2013).

¹⁹ Trachtman (n 13) 36.

describing, rather than quantifying, factors such as importance, contribution and trade restrictiveness.

Prior to *Korea Beef*, Trachtman stated that the WTO did not use any CBA devices (a claim which was true at the time).²⁰ His approach has nonetheless provided a conceptual framework for subsequent authors to analyse the W&B test.

Trachtman recognises that CBA devices have pros and cons. If a legal system is purely interested in efficiency, Trachtman notes that “comparative cost-benefit analysis unqualifiedly maximizes the net sum of gains from trade and from regulation”.²¹ However, he recognises that this device has other weaknesses in terms of “administrability” as well as “distributive, moral and theoretical concerns”.²² He notes that “comparative cost-benefit analysis is inevitably political, and is never neutral”.²³

The choice of trade-off device therefore seems directly related to the foundational question about the underlying rationale for a given legal system. If the WTO sought to maximise global efficiency, CBA would be the most appropriate trade-off device. However, many scholars describe it as an overly-intrusive device for the WTO system which reserves policy space for Members to pursue national welfare at the expense of global efficiency.

I will not make claims about which CBA device best describes the W&B test. In fact, Trachtman notes that they “are often used interchangeably”.²⁴ For the purpose of my analysis, each of these devices share the key features which make W&B testing so controversial, as they are all based on “commensuration” (of matters such as

²⁰ ibid 58.

²¹ ibid 80.

²² ibid.

²³ ibid 84.

²⁴ ibid 77.

importance) and “interpersonal comparison of utility” between different social values and stakeholders (such as the balancing step).²⁵

Bown and Trachtman recognise that the W&B test “does refer to most of the factors that would be relevant in cost–benefit analysis”²⁶, including the controversial aspect of commensuration.²⁷ On paper, the W&B test seems like it belongs within the CBA family of trade-off devices and I’ll therefore treat it as a type of CBA device.

In Trachtman’s view, issues which require commensuration across different social values should ideally be dealt with through the political process as it is “best able to engage in subtle balancing and weighing of competing social interests”.²⁸

Ehlermann agrees that “unclear, controversial and politically delicate” issues require political solutions through “the negotiation and conclusion of agreements”.²⁹

However, for a variety of reasons, this task has regularly fallen to adjudicators in controversial disputes. Ehlermann describes the current situation afflicting the political (negotiation) function as a “stalemate”,³⁰ while Trachtman argues that the WTO “is totally lacking in conventional legislative capacity”.³¹ A recent EU concept paper notes that since its inception in 1995:

“The WTO’s negotiating function has not been able to deliver any significant improvements in the trade rulebook apart from the agreements reached on Trade Facilitation and Export Competition”.³²

²⁵ ibid 84.

²⁶ Bown and Trachtman (n 6) 117.

²⁷ Bown and Trachtman (ibid 117) note that, as a technical matter, the W&B test is not identical to a CBA as some factors are missing, such as implementation costs for the regulator or costs/benefits in the exporting country.

²⁸ Trachtman (n 13) 47.

²⁹ Claus-Dieter Ehlermann, ‘Tensions between the dispute settlement process and the diplomatic and treaty-making activities of the WTO’ (2002) 1(3) World Trade Review 301, 307.

³⁰ ibid 305.

³¹ Trachtman (n 13) 51.

³² EU Concept Paper and Press Release (18 September 2018) <https://europa.eu/rapid/press-release_IP-18-5786_en.htm> (consulted 5 August 2019) 1-2.

In summary, a wide range of scholars agree that the W&B test strongly resembles CBA testing and that this method is overly-intrusive and inappropriate as the standard of review for Article XX.

5.3 AM Test (and Pareto Efficiency)

At first glance, the AM test appears far less controversial. Ortino suggests the AM test is less intrusive because “the policy objective pursued by the Member and the level of protection determined by the Member itself is not called into question”.³³ This is consistent with Howse’s view that WTO rules “should not constrain the moral ends sought by its Member states, but only the means used to achieve those ends.”³⁴ Regan describes it as the test which is “suggested most naturally” by the term “necessary”.³⁵

Fontanelli describes the AM test as “narrowly devoted to ensure Pareto optimization”³⁶ or “a formula of (Pareto) efficiency”.³⁷ However, it is clear that he also understands the test’s potential to be applied in intrusive ways which go beyond Pareto efficiency.

Sykes acknowledges that there could be certain non-controversial cases where “an alternative regulation unquestionably achieves a clearly stipulated regulatory objective at equal or lower cost to regulators”.³⁸ These cases could be resolved, in a non-intrusive manner, under the AM test.

However, Sykes notes that these “easy” cases are rare in practice. In reality, adjudicators must deal with more difficult scenarios where proposed alternatives are

³³ Ortino (n 4) 143.

³⁴ Howse et al (n 9) 87.

³⁵ Regan (n 1) 348.

³⁶ Fontanelli (n 5) 68.

³⁷ ibid 45.

³⁸ Alan Sykes, ‘The Least Restrictive Means’ (2003) 70 University of Chicago Law Review 403, 403.

“somewhat more costly to implement” or “slightly less effective at achieving the stated regulatory objective”.³⁹ In these cases, adjudicators must engage in commensuration to determine whether a proposed alternative produces enough utility (in terms of increased trade) to justify this burden on the regulator. Due to this commensuration, Sykes argues that the AM test is a “crude” CBA and that “cost-benefit logic lies at the center of analysis”.⁴⁰

Bown and Trachtman appear to agree. In their view, the “reasonably available” requirement makes it “difficult to describe this language as anything but the announcement of a balancing test.”⁴¹ Fontanelli suggests that the AM test is “less value-neutral than the quasi-judicial bodies would claim it to be”.⁴²

On paper, the AM test appears like a non-intrusive approach which can identify Pareto gains without treading on regulatory autonomy. However, scholars have highlighted a number of ways in which the AM test is equipped to undertake intrusive analysis, especially under the “reasonably available” requirement. This intrusiveness will be examined comprehensively in the *Korea Beef* and *Brazil Tyres* case studies.

5.4 Interaction between the Tests

The theory suggests that the W&B and AM operate in significantly different ways. CBA devices are value-laden and intrusive on regulatory autonomy. They allow adjudicators to commensurate between the positive and negative consequences of a

³⁹ ibid.

⁴⁰ ibid 404-407.

⁴¹ Bown and Trachtman (n 6) 121.

⁴² Fontanelli (n 5) 39.

measure. By contrast, AM testing appears technical and deferential. It respects the regulator's goals and only seeks to review the means employed to achieve them.⁴³

The Appellate Body insists that both tests are part of a "single, integrated, yet multifaceted inquiry"⁴⁴ and that the AM test merely "confirms" the result of the W&B test. This implies that the W&B and AM tests are functionally identical (or at least highly similar), but the theory suggests they use different methods to answer different questions.

This section will seek to demonstrate some key theoretical differences between the two tests. The two tests have different views on regulatory autonomy, are capable of producing different results and (at least in theory) offer different remedies. I will further highlight some logical contradictions which flow from the Appellate Body's failure to acknowledge the theoretical differences between the tests.

5.4.1 Regulatory Autonomy and the Object of Review

Since regulatory autonomy plays a critical role under Article XX, I will seek to define what it means in practical terms. In *Brazil Tyres*, the Appellate Body stated that:

"it is within the authority of a WTO Member to set the public health or environmental objectives it seeks to achieve, as well as the level of protection that it wants to obtain, through the measure or the policy it chooses to adopt".⁴⁵

In my view, this quote highlights three key elements of regulatory autonomy:

- (i) Policy objective
- (ii) Level of protection
- (iii) Means employed

⁴³ This is the stereotype of the AM test. As already discussed, Sykes and other scholars see through this stereotype and describe the AM test as a "crude CBA".

⁴⁴ Appellate Body Report, *China – Measures Affecting Trading Rights and Distribution Services for Certain Publications and Audiovisual Entertainment Products*, WT/DS363/AB/R, adopted 19 January 2010 (*China Audiovisuals*) para 237.

⁴⁵ Appellate Body Report, *Brazil – Measures Affecting Imports of Retreaded Tyres*, WT/DS332/AB/R, adopted 17 December 2007 (*Brazil Tyres*) para 140.

My analysis will use these three elements as a framework for discussing regulatory autonomy. A key question for WTO adjudicators is which of these elements to treat as a legitimate object of review (thus potentially intruding on regulatory autonomy) and which elements to exclude from judicial review (thus deferring to Members).

In *Gasoline*, under the AM test, the Appellate Body stated that the US “remained free to regulate in order to obtain whatever air quality it wished”.⁴⁶ This statement, made prior to the introduction of the W&B test, appeared to exclude from judicial review both the objective and level of protection. The US possessed a sovereign right to improve air quality (its objective) and to choose its desired level of air quality protection.⁴⁷ Adjudicators limited their review to the means employed: could the US have achieved its objective and level of protection through an alternative measure which was less treaty-inconsistent?

The W&B test appeared to broaden the scope of which elements of regulatory autonomy could legitimately be reviewed under the necessity test. Adjudicators could (and indeed had to) review the “importance of the *objective*” and the “contribution of the measure to the objective”⁴⁸ which seems to render the objective and level of protection subject to judicial review. This section will treat the three elements of regulatory autonomy as a framework for illustrating some highly-significant differences between the W&B and AM tests.

Policy Objective (Importance)

⁴⁶ Appellate Body Report, *United States – Standards for Reformulated and Conventional Gasoline*, WT/DS2/AB/R, adopted 20 May 1996 (*Gasoline*) 15 (citing *Gasoline* panel).

⁴⁷ The Appellate Body did not explicitly recognise the US’ objective as, prior to the W&B test, it tended to describe regulatory autonomy solely by reference to the “level of protection”. As previously discussed, the US’ objective was implicit in the level of protection.

⁴⁸ In the doctrinal analysis, I showed that the “contribution” and “level of protection” appear to be analogous concepts. This idea will be developed further in the *Brazil Tyres* case study.

The *Brazil Tyres* panel confirmed that Members retain regulatory autonomy to choose their policy objectives: “we are not, in our view, required to examine the desirability of the declared policy goal as such”.⁴⁹ The Appellate Body further confirmed that “it is within the authority of a WTO Member to set the public health or environmental objectives it seeks to achieve”.⁵⁰ These statements are consistent with the traditional AM test approach which treats the policy objective as non-reviewable.

However, the *Brazil Tyres* adjudicators contradicted this view when they reviewed the importance of Brazil’s policy objective. They found that “few interests are more ‘vital’ and ‘important’ than protecting human beings from health risks”.⁵¹ On one hand, they may not “examine the desirability of the declared policy goal” but, on the other hand, they can assess the “importance of the objective”. Regan astutely describes this situation as a “logical contradiction”.⁵²

Level of Protection (Contribution)

The level of protection follows a very similar pattern to the policy objective.⁵³ The fact that it is addressed under both the W&B and AM tests might not be immediately apparent because the Appellate Body uses different terminology. The AM test refers to the “level of protection”, while the W&B test refers to the “contribution”, however they are substantively the same concept.⁵⁴ Sometimes adjudicators interchange the

⁴⁹ Panel Report, *Brazil – Measures Affecting Imports of Retreaded Tyres*, WT/DS332/R, adopted 17 December 2007 (*Brazil Tyres*) para 7.97.

⁵⁰ Appellate Body Report, *Brazil Tyres* (n 45) para 140.

⁵¹ ibid para 144 (confirming the panel’s finding).

⁵² Regan (n 1) 361. I will argue in the next chapter that adjudicators “solve” this problem by always finding that the objective is highly important, as they did in *Brazil Tyres*.

⁵³ I will explain below that the policy objective and level of protection are closely-related concepts and that the principles which apply to them are highly similar. In my view, it would make sense to combine them under a single banner of social utility.

⁵⁴ See footnote 48 (above)

terminology and refer to an “equivalent contribution” (instead of the level of protection) under the AM test.⁵⁵

The AM test treats the level of protection as non-reviewable. Rather than judging a Member’s chosen level of protection, adjudicators merely measure it in order to establish a benchmark for proposed alternatives. This approach is highly-deferential as it entails no value judgment about the chosen level of protection and ensures that proposed alternatives must respect the regulator’s choice.

For example, in *Gasoline*, the US sought a “15 per cent reduction in the emissions” of certain pollutants.⁵⁶ The AM test barred adjudicators from considering alternative measures with a lower emissions reduction benchmark (let’s say 10%) as this would lower the US’ level of protection.⁵⁷ Adjudicators could merely engage in a technical exercise to see if the US could achieve its 15% reduction target in ways which were less treaty-inconsistent.

Under a W&B analysis, adjudicators consider the “contribution of the measure to the objective” and may review if the contribution is sufficiently high to justify the measure’s adverse trade effects. The *Brazil Tyres* report suggested that some contributions might be too low:

we disagree with Brazil's suggestion that... an import ban that brings a marginal or insignificant contribution can nevertheless be considered necessary.⁵⁸

⁵⁵ See eg Appellate Body Report, *Brazil Tyres* (n 45) para 178.

⁵⁶ Appellate Body Report, *Gasoline* (n 46) 4.

⁵⁷ It was also barred from considering alternative measures with a higher level of protection, such as a 25% emissions reduction, even if it considered that such an approach would achieve the best balance between air quality and trade restrictiveness.

⁵⁸ Appellate Body Report, *Brazil Tyres* (n 45) para 150.

Does this mean that some levels of protection are too low to justify a trade-restrictive measure? Adjudicators seem to suggest that, at least for highly trade-restrictive measures like import bans, the level of protection must meet a minimum threshold (relative to the other W&B factors):

a panel might conclude that an import ban is necessary on the basis of a demonstration that the import ban at issue is apt to produce a material contribution to the achievement of its objective.⁵⁹

This entails a review of the regulator's level of protection. The Appellate Body claims that "the fundamental principle is the right that WTO Members have to determine the level of protection that they consider appropriate".⁶⁰ However, under the W&B test, the level of protection is subject to review and capable of being deemed insufficient. It seems that different principles apply to the level of protection when adjudicators refer to it as the "contribution".⁶¹

Means Employed (Trade-Restrictiveness)

Under both the W&B and AM tests, the Appellate Body must assess the trade restrictiveness of the means employed. However, its purpose for looking at trade restrictiveness will vary under each test.

The AM test looks at trade restrictiveness to see if the same objective could be achieved in a less restrictive way. In *Korea Beef*, the regulator's dual retail system failed the necessity test as Korea supposedly could have achieved its objective (and level of protection) by employing less restrictive means. Adjudicators found that a

⁵⁹ ibid para 151.

⁶⁰ ibid para 210

⁶¹ Adjudicators apply different rules to the terms "contribution" and "level of protection" even where they appear to be substantively the same concept. Is this an example of a "framing effect"?

proposed alternative was reasonably available: “traditional enforcement measures” including “record-keeping, investigations, policing and fines”⁶². For the AM test, the Appellate Body compares the trade restrictiveness of one (actual) measure with the trade restrictiveness of another (hypothetical) measure.⁶³

Under the W&B test, the AB assesses the measure’s trade-restrictiveness to see if it can justify the social utility. It compares the trade restrictiveness of one (actual) measure with the importance/contribution of the same (actual) measure. Unlike the AM test, this comparative analysis requires commensuration between trade and other social values.

Summary of Key Differences

By treating the policy objective and level of protection as non-reviewable, the AM test seems highly deferential. However, regarding the means employed, its approach can be highly-intrusive. In *Shrimp*, the US highlighted this regulatory intrusiveness when it suggested the AM test:

“require[s] dispute settlement panels to dictate the specific measure to be adopted by a WTO Member, since presumably there was only one measure among all the alternatives that was the ‘least inconsistent’ with the GATT 1994”⁶⁴.

If adjudicators may reject “necessity” by identifying a superior alternative, is this tantamount to telling regulators which means to employ to achieve their objective? In theory, the proposed alternative is merely a conceptual tool to demonstrate that a

⁶² Appellate Body Report, *Korea – Measures Affecting Imports of Fresh, Chilled and Frozen Beef*, WT/DS161/AB/R, WT/DS169/AB/R, adopted 10 January 2001 (*Korea Beef*) para 153. In the *Korea Beef* case study, I will show compelling arguments from some scholars that this proposed alternative actually lowered Korea’s level of protection.

⁶³ Adjudicators may choose to consider a number of different alternative measures (rather than a single proposed alternative).

⁶⁴ Panel Report, *United States – Import Prohibition of Certain Shrimp and Shrimp Products*, WT/DS58/R and Corr.1, adopted 6 November 1998, as modified by Appellate Body (*Shrimp*) 3.228.

measure is excessively trade-restrictive. However, in practice, it might be interpreted as a prescriptive statement of the regulator's compliance requirements.

When scholars invoke "regulatory autonomy" to express a preference for the less-intrusive AM test, they are implicitly suggesting that some elements of regulatory autonomy (the objective and level of protection) are more important than others (means employed).

In fact, adjudicators openly assert their right to question the means employed. In *Sardines*, the panel offered a "degree of deference" to regulators regarding their "domestic policy objectives".⁶⁵ It recognised that:

"It is up to the Members to decide which policy objectives they wish to pursue and the levels at which they wish to pursue them".⁶⁶

However, this deference regarding the policy objective and level of protection dissipated for the means employed:

"the TBT Agreement, like the GATT 1994, shows less deference to the means which Members choose to employ to achieve their domestic policy goals".⁶⁷

In summary, there appear to be some major differences between the W&B and AM tests with respect to regulatory autonomy and the object of review. The AM test is deferential regarding the policy objective and level of protection and treats both matters as non-reviewable.

The W&B test, on the other hand, treats each of these matters as subject to judicial review (albeit using different terminology). This is at odds with the traditional view at

⁶⁵ Panel Report, *European Communities – Trade Description of Sardines*, WT/DS231/R and Corr.1, adopted 23 October 2002, as modified by Appellate Body Report WT/DS231/AB/R (*Sardines*) para 7.120. The panel's analysis occurred under the TBT Agreement but was considered equally relevant to GATT Article XX.

⁶⁶ *ibid* 7.120.

⁶⁷ *ibid*.

the WTO that “ends are not justiciable, means are”.⁶⁸ There are significant consequences which flow from these differences. Are the tests capable of producing different results and compliance obligations? The analysis will now turn to this issue.

Different Results and Remedies

Scholars rightly highlight that the W&B and AM tests can produce different results. Regan suggests it is “easy to imagine that some measure passes the less-restrictive alternative test but fails the cost–benefit balancing test”.⁶⁹ The practice in disputes suggests his claim is accurate.

In *Audiovisuals*, the US challenged three Chinese measures which operated as a package to restrict trade (and trading rights) for certain audiovisual products. The panel determined that two of the measures (the State-Ownership Requirement and the Exclusion of Foreign-Invested Enterprises) failed the W&B test as they made “no material contribution to the protection of public morals in China”.⁷⁰ When adjudicators find that a measure did not make a sufficiently “material contribution” to justify its trade-restrictiveness, is this not tantamount to a finding that the level of protection was too low?⁷¹

The panel deemed China’s measures unnecessary under the W&B test and found no need to perform an AM analysis to “confirm” this finding.⁷² It saw no value in

⁶⁸ Fontanelli (n 5) 48 (citing Mavroidis).

⁶⁹ Regan (n 1) 351.

⁷⁰ Panel Report, *China – Measures Affecting Trading Rights and Distribution Services for Certain Publications and Audiovisual Entertainment Products*, WT/DS363/R and Corr.1, adopted 19 January 2010, as modified by Appellate Body Report WT/DS363/AB/R (*China Audiovisuals*) para 234.

⁷¹ In principle, there could be rare cases where a measure makes no contribution whatsoever and an adjudicator dismisses the necessity of the measure on this basis alone. In my view, “no material contribution” represents a low (but higher than zero) contribution.

⁷² This suggests that the W&B test can be decisive (at least in some cases where a measure is deemed unnecessary). It further suggests that the AM test cannot overrule a W&B finding that a measure is unnecessary. Presumably this is why the AM test was skipped altogether and adjudicators did not consider it useful to see if it would produce a different result.

trying to identify a less trade-restrictive way of achieving the contested measure's contribution (which presumably fell somewhere between no contribution and a material contribution). This approach seems to imply that China must remove its measures in their entirety (on the basis that they are unnecessary). There was no suggestion by adjudicators that China could merely reform the means employed (to make the measures less trade-restrictive and therefore WTO-compliant).

This is very different to a loss under the AM test, where regulators may retain their policy objective and level of protection, but merely reform the means employed to eliminate the excessively trade-restrictive aspects of the measure. For example, in *Gasoline*, the US was allowed to retain its benchmark of a 15% emissions reduction, but was required to remove certain aspects of the benchmark establishment rules which adjudicators deemed more treaty-inconsistent than necessary.

Should Adjudicators Review the Measure or the Breach?

It is not surprising that the W&B and AM tests produce different results as they use different objects of review to probe different aspects of regulatory autonomy. We can interrogate these differences by reference to another longstanding debate at the WTO: does an Article XX defence apply to a measure in its entirety or just the WTO-inconsistent aspects?

In *Gasoline*, the Appellate Body stated that "the chapeau of Article XX makes it clear that it is the 'measures' which are to be examined under Article XX(g), and not the legal finding of 'less favourable treatment.'"⁷³ The Appellate Body has reinforced this message in subsequent disputes, such as *Seals*, where it stated that "what must be

⁷³ Appellate Body Report, *Gasoline* (n 46) 16.

justified is, as we have said, both the prohibitive and permissive components of the EU Seal Regime, taken together".⁷⁴

Some adjudicators, however, apply a different approach. The Appellate Body has stated that "what must be shown to be 'necessary' is the treatment giving rise to the finding of less favourable treatment".⁷⁵ In *Brazil Taxation*, for example, the panel concluded that "Brazil has not demonstrated that the discriminatory aspects of the [measure] have led, will lead, or are apt to lead, to an increase in vehicle safety or energy efficiency."⁷⁶ The focus of this panel's analysis was on the "discriminatory aspects" rather than the measure in its entirety.

Adjudicators appear confused about whether it is the measure or the violation which must be defended under Article XX? In my view, any confusion is fully justified. Under the W&B test, it appears to be the measure which is on trial. An adverse finding – that the objective is insufficiently important or that the contribution is too low - will logically require the regulator to revoke the measure. This occurred in *Audiovisuals* and it is coherent with Fontanelli's view that the Appellate Body is "looking into the merits of the measures under review (not simply into their efficiency and their functional design)".⁷⁷

Under the AM test, it appears that only the treaty-inconsistent (or excessively trade-restrictive) aspects of the measure are on trial. An adverse finding under the AM test may require a regulator to reform the WTO-inconsistent aspects of the measure, but

⁷⁴ Appellate Body Reports, *European Communities – Measures Prohibiting the Importation and Marketing of Seal Products*, WT/DS400/AB/R / WT/DS401/AB/R, adopted 18 June 2014 (*Seals*) para 5.193.

⁷⁵ Appellate Body Report, *Thailand – Customs and Fiscal Measures on Cigarettes from the Philippines*, WT/DS371/AB/R, adopted 15 July 2011 (*Thailand – Cigarettes (Philippines)*) para 177.

⁷⁶ Panel Reports, *Brazil – Certain Measures Concerning Taxation and Charges*, WT/DS472/R, Add.1 and Corr.1, WT/DS497/R, Add.1 and Corr.1, adopted 11 January 2019, as modified by Appellate Body Reports WT/DS472/AB/R, WT/DS497/AB/R (*Brazil Taxation*) para 7.921.

⁷⁷ Fontanelli (n 5) 53.

it does not require the regulator to revoke the measure in its entirety. The policy objective and level of protection are excluded from review and can be retained regardless of the dispute outcome.

Unfortunately, the Appellate Body often describes the necessity test with broad aphorisms (about respect for regulatory autonomy) which appear true for the AM test but are highly questionable for the W&B test. For example, when the Appellate Body affirms that the policy objective or level of protection is inviolate, this statement is only valid under the AM test.⁷⁸ Rather than using broad aphorisms, the Appellate Body should establish separate and precise methods for describing each of the necessity tests.

5.4.2 Logical Contradictions

Many WTO scholars criticise the Appellate Body for developing “a contradictory doctrinal basis and a deeply flawed methodology”.⁷⁹ Regan astutely notes that “a standard cost–benefit balancing test is inconsistent with allowing the Member to choose its own level of protection.”⁸⁰ Weiler agrees that “adopting a balancing test has to mean that, at least in some circumstances, the supposed hallowed right of a State to set its own level of acceptable risk is not absolute”.⁸¹ In his view, the principle of regulatory autonomy appears to “actually prohibit balancing”.⁸²

Bown and Trachtman agree that “balancing tests seem to some to intervene too greatly in national regulatory autonomy”.⁸³ Ming Du suggests that the Appellate Body “is pragmatic in the sense that it both retains de jure regulatory autonomy, but

⁷⁸ By contrast, when adjudicators state that Article XX reviews the measure (rather than the treaty-inconsistent aspects), this statement would be valid under the W&B test, but not the AM test.

⁷⁹ Weiler (n 3) 142.

⁸⁰ Regan (n 1) 348.

⁸¹ Weiler (n 3) 140.

⁸² ibid 139.

⁸³ Bown and Trachtman (n 6) 119.

de facto allows balancing scrutiny to root out indefensible, haphazardly set risk levels".⁸⁴ He wonders whether the principle of regulatory autonomy is "real or rhetorical".⁸⁵

Ortino recognises the Appellate Body's use of "contradictory language" in *Korea Beef* and "all the subsequent reports" and further suggests that the use of CBA devices entails "a review of the given objective itself".⁸⁶ Kapterian suggests that the Appellate Body has failed "to develop a coherent legal doctrine on the degree of deference to be afforded to Members".⁸⁷

Regan has some sympathy for the Appellate Body when he states that:

"I think the Appellate Body is right not to engage in cost–benefit balancing... But the doctrinal state of affairs – involving repeated statements of contradictory tests side-by-side – is unfortunate and dangerous".⁸⁸

The Appellate Body is faced with a no win situation on regulatory autonomy. If it seeks to weed out protectionism by using the W&B test as a CBA, it will be accused of regulatory intrusiveness. However, if it seeks to be deferential to regulators by applying a narrow (Pareto) version of the AM test, it may open the door to protectionism. Trachtman highlights that a narrow approach to AM testing "would allow states 'incidentally' to confer grave detriments on other states, in pursuit of even the smallest benefit at home".⁸⁹

Weiler further notes the weakness of a narrow AM approach, asking:

⁸⁴ Ming Du (n 8) 1101.

⁸⁵ ibid 1097.

⁸⁶ Ortino (n 4) 143-144.

⁸⁷ Gisele Kapterian, 'A Critique of the WTO Jurisprudence on 'Necessity' (2010) 59 International & Comparative Law Quarterly 89, 118.

⁸⁸ Regan (n 1) 350.

⁸⁹ Trachtman (n 13) 48.

“Does it really make sense to stipulate zero or a very low level of tolerance to any consumer confusion if the cost is excluding totally or seriously impeding an imported product from the market place?”⁹⁰

It is inevitable that adjudicators will be heavily scrutinised (and even criticised) for their judgments about where to draw the line on retained sovereignty, however this does not justify their “internally incoherent” approach.⁹¹

Other jurisdictions are transparent about their use of a composite approach such as German proportionality testing which requires “legitimacy of purpose”, an AM test and “cost–benefit balancing”.⁹² It is reasonable (at least in the German context) to have multiple methodologies (based on different premises) which operate cumulatively to test the same measure. Perhaps the Appellate Body struggles to admit that it uses two different necessity tests for two different purposes because it has not been given a mandate to engage in constitutional adjudication.⁹³

Chapters 6 and 7 will deconstruct how the necessity tests are applied in practice. I will argue that the Appellate Body fails to apply the tests as the theory would predict. It has taken a highly-deferential approach to the W&B test (using it more as a conceptual framework than a true CBA), but it has applied the AM test in ways which are surprisingly intrusive and ultimately amount to “crude” CBA testing.⁹⁴

The application of the tests leads to a strange inversion where the true threat to regulatory autonomy comes from hidden value judgments within the AM test.

Fontanelli argues that the W&B test is functionally “irrelevant” but that this has not

⁹⁰ Weiler (n 3) 144. These comments were made as part of a critique of *Korea Beef*.

⁹¹ Bown and Trachtman (n 6) 86.

⁹² Weiler (n 3) 139. This approach seems to closely resemble Article XX’s design step, AM test and W&B test.

⁹³ This is particularly relevant for those aspects of the test which rely on balancing and commensuration.

⁹⁴ Sykes (n 38) 415. Sykes further states: “By ‘crude’ I mean that the WTO decisionmaker does not actually quantify the costs and benefits of alternative regulatory policies in dollars or some other metric”.

prevented adjudicators from “an unavowed pattern of judicial interference into States’ policies” under the AM test.⁹⁵ He states that:

“an element of stricto sensu proportionality (or cost-benefit analysis) guides at times the necessity test performed by Panels and AB, but this exercise of appreciation is not embedded in the balancing effort (as it would be normal to assume), but in the loose application of the [AM test] (which would, in principle, bar discretionary evaluation)⁹⁶”.

This inversion of the tests renders the assessment of necessity non-transparent, unpredictable and potentially arbitrary. As Bown and Trachtman note, “in order to rationalize deference, the decisions have done much violence to text, to precedent, and to legal logic”.⁹⁷

5.5 Summary of Theoretical Critique

At a theoretical level, the W&B test has been highly criticised for introducing CBA into the WTO and undermining regulatory autonomy. Many scholars view the AM test as more consistent with the WTO’s negative integration model, though a handful of scholars astutely point out that the “reasonably available” requirement is essentially a CBA in disguise.

The Appellate Body has stubbornly held its line that the W&B and AM test are merely two different ways of asking the same question, even though this view leads to logical contradictions. It would be *inutile* to have two ways of addressing the same question and it is obvious to many WTO observers that the two necessity tests differ in critical ways. They each have a different philosophical basis, probe different aspects of regulatory autonomy, have different objects of review, produce different results (at least in some disputes), offer different remedies and interrogate different

⁹⁵ Fontanelli (n 5) 44.

⁹⁶ ibid 43.

⁹⁷ Bown and Trachtman (n 6) 131.

aspects of the measure. How has the Appellate Body developed a sophisticated jurisprudence, and adjudicated over a dozen disputes, without acknowledging these differences?

The next chapter will look at how the Appellate Body applies its “contradictory doctrinal basis” and “deeply flawed methodology” in practice. I will argue that the Appellate Body has been highly deferential under the W&B test, but that it has found techniques to embed its intrusive value judgments within the AM test.

CHAPTER 6 - DECONSTRUCTING THE NECESSITY TEST

This chapter will deconstruct the W&B and AM tests to see how they are applied in practice. I will argue that the W&B test has been applied in a highly-deferential way which does not truly resemble CBA. The *Korea Beef* case study will show how adjudicators have found ways to embed intrusive value judgments into the AM test.¹

6.1 W&B Test

The previous chapter showed that leading WTO scholars criticise the W&B test as a type of CBA which has the potential to intrude excessively into the regulatory autonomy of Members. In practice, adjudicators have found techniques to apply the W&B test non-intrusively (contrary to its plain meaning) in ways which respect the pluralist “spirit” of the WTO. This is particularly noticeable for the normative elements of the test: “importance” and “balancing”.

This has led to criticism by scholars due to divergence between the description of the test and its application. Pirker describes the W&B analysis as “a virtually incomprehensible legal test, which causes great costs in terms of predictability and legal certainty.”² In his view, the “true content” of the legal test “is virtually impossible to discern”.³

6.1.1 Importance

Some scholars argue (convincingly) that adjudicators do not really undertake an importance analysis. Following the early W&B cases, Regan concluded that “the

¹ The next chapter will undertake a similar case study for *Brazil Tyres* to show how the W&B test is applied in practice in ways which can favour the regulator.

² Benedikt Pirker, *Proportionality Analysis and Models of Judicial Review: A Theoretical and Comparative Study* (Europa Law Publishing 2013) 299.

³ ibid 298.

Appellate Body has yet to say of any legitimate purpose that it is not important".⁴ In some cases, such as *Asbestos*, this conclusion was "easy", but Regan seems surprised at the Panel's view in *DR - Cigarettes*⁵ (endorsed by the Appellate Body) that collecting tax revenue on cigarettes is "vitally important."⁶

This trend continued in subsequent disputes. In my assessment, the lowest level of "importance" on record is "reasonably important".⁷ There have yet to be any findings that a policy objective was "unimportant" or "of minor importance". Sometimes, adjudicators describe importance in comparative terms, such as the "commercial" objectives in *Canada Wheat* which were considered less important than the health objectives in *Asbestos*.⁸

Fontanelli agrees that adjudicators are extremely unlikely to deem any objective unimportant and describes this legal element as "virtually untouchable".⁹ Could WTO adjudicators really reject a defence on the basis that a policy objective was not sufficiently important?

These scholars are correct that importance is scrutinised less rigorously than the non-normative factors (contribution and trade-restrictiveness) where adjudicators

⁴ Don Regan, 'The Meaning of 'Necessary' in GATT Article XX and GATS Article XIV: The Myth of Cost-Benefit Balancing' (2007) 6 *World Trade Review* 347, 360.

⁵ Panel Report, *Dominican Republic – Measures Affecting the Importation and Internal Sale of Cigarettes*, WT/DS302/R, adopted 19 May 2005, as modified by Appellate Body Report WT/DS302/AB/R (*DR – Cigarettes*).

⁶ Regan (n 4) 362.

⁷ Panel Reports, *Brazil – Certain Measures Concerning Taxation and Charges*, WT/DS472/R, Add.1 and Corr.1, WT/DS497/R, Add.1 and Corr.1, adopted 11 January 2019, as modified by Appellate Body Reports WT/DS472/AB/R, WT/DS497/AB/R (*Brazil Taxation*) para 7.592.

⁸ Panel Report, *Canada – Measures Relating to Exports of Wheat and Treatment of Imported Grain*, WT/DS276/R, adopted 27 September 2004, upheld by Appellate Body Report WT/DS276/AB/R (*Canada Wheat*) para 6.224. The panel stated: "It is clear that these interests, which appear to be essentially commercial in nature, are important. It seems equally clear, however, that these interests are not as important as, for instance, the protection of human life and health against a life-threatening health risk".

⁹ Filippo Fontanelli, 'Necessity Killed the GATT - Art XX GATT and the Misleading Rhetoric about Weighing and Balancing' (2012) 5 *European Journal of Legal Studies* 39, 56.

have developed prescriptive methodologies.¹⁰ Why have adjudicators failed to develop similar rules and methodologies for “importance”? In practice, adjudicators avoid substantive analysis of importance through vague reasoning and a practice of always determining that the objective is important.

6.1.2 Balancing Step

WTO observers note that the Appellate Body ultimately avoids any balancing step as this would require commensuration between competing values. Fontanelli claims that “no real balancing is ever performed - or in any event, relied on - to determine the outcome of a dispute”.¹¹ Ortino goes further and calls on the Appellate Body to confirm that “there is no real weighing and balancing at play in applying the necessity test”.¹²

Bown and Trachtman fail to find any “compelling” evidence of “actual balancing”.¹³ They describe the analysis of the W&B factors (such as importance, contribution and trade-restrictiveness) as “imprecise and ultimately malleable”.¹⁴ They argue that the Appellate Body has “shown itself unwilling to evaluate for itself, or to require a Panel to evaluate, in any but the most gross categories, any of these four factors.”¹⁵

Members seem to be aware of the disconnect between the formulation and application of the W&B test. In *Brazil Tyres*, the EU argued that the panel had failed

¹⁰ As discussed in the previous chapter, for the non-normative factors, adjudicators must use qualitative or quantitative reasoning to identify a “degree” or “extent” of contribution/trade-restrictiveness and they should be as precise as possible.

¹¹ Fontanelli (n 9) 39.

¹² Federico Ortino, ‘GATT’ in Daniel Bethlehem, Isabelle Van Damme, Donald McRae, and Rodney Neufeld (eds), *The Oxford Handbook of International Trade Law* (OUP 2009) 145.

¹³ Chad Bown and Joel Trachtman, ‘Brazil – Measures Affecting the Import of Retreaded Tyres: A Balancing Act’ (2009) 8 World Trade Review 85, 122.

¹⁴ ibid 89.

¹⁵ ibid 88. The four factors being referred to are: importance, contribution, trade-restrictiveness and balancing.

to undertake the balancing analysis “as a separate step” and described the overall approach as a “superficial analysis”.¹⁶

The Appellate Body rejected this claim, stating:

“The weighing and balancing is a holistic operation that involves putting all the variables of the equation together and evaluating them in relation to each other after having examined them individually, in order to reach an overall judgement”.¹⁷

Apart from asserting it as a fact, the Appellate Body provided no evidence or other support for its view that the panel had undertaken a balancing analysis. The EU’s claim was ultimately ignored and unrefuted.

6.1.3 Purpose of W&B Test

If adjudicators refrain from meaningful analysis of importance and the balancing step, what purpose does the W&B test serve? Scholars propose two explanations.

First, the W&B test serves as “a gateway filter for unacceptable measures”.¹⁸ It can filter out measures which are “not trade-restrictive” or which make “no contribution”.¹⁹ This approach was used in *Colombia – Ports of Entry*²⁰ where the measure made an “insignificant contribution” and thus failed the W&B test.²¹ Pirker notes that this filtering function is only effective for “rather obvious cases”.²² The most complex and controversial cases will pass the W&B gateway filter and be addressed under the AM test.

¹⁶ Appellate Body Report, *Brazil – Measures Affecting Imports of Retreaded Tyres*, WT/DS332/AB/R, adopted 17 December 2007 (*Brazil Tyres*) para 176.

¹⁷ ibid para 182.

¹⁸ Fontanelli (n 9) 43.

¹⁹ Pirker (n 2) 322.

²⁰ Panel Report, *Colombia – Indicative Prices and Restrictions on Ports of Entry*, WT/DS366/R and Corr.1, adopted 20 May 2009 (*Colombia – Ports of Entry*).

²¹ Fontanelli (n 9) 58.

²² Pirker (n 2) 322.

Perhaps the hardest case which a panel has dismissed at the W&B stage was *Audiovisuals*, where the panel found that certain contested measures made “no material contribution to the protection of public morals in China”.²³ Since this was based on “no material contribution” (rather than zero contribution), it arguably could have required some balancing (albeit with a very low value for contribution).²⁴

The second view is that adjudicators ultimately decide difficult cases under the AM test, but that the value judgment which underpins the decision can be gleaned from the W&B analysis. This suggests that the real reasons behind the decision (the persuasive moral arguments) appear in the W&B analysis, but that the decisive judgment is made under the AM test.

Fontanelli describes the W&B test as a “preparatory exercise” and a “warm-up test”.²⁵ Pirker agrees that “the first stage [W&B test] is undertaken with such light scrutiny that in most cases the comparison with alternatives in fact constitutes the decisive part of the analysis of measures”.²⁶

It is concerning that the Appellate Body has invented a new (W&B) test, but failed to respect the plain language meaning of its own jurisprudence. While Regan appears to be happy that the Appellate Body does not actually conduct CBA analysis, he expresses concern about the evolution of the jurisprudence:

²³ Panel Report, *China – Measures Affecting Trading Rights and Distribution Services for Certain Publications and Audiovisual Entertainment Products*, WT/DS363/R and Corr.1, adopted 19 January 2010, as modified by Appellate Body Report WT/DS363/AB/R (*China Audiovisuals*) para 234.

²⁴ The Appellate Body clarified that “since a measure's contribution is only one component of the necessity calculus under Article XX, the assessment of whether a measure is 'necessary' cannot be determined by the degree of contribution alone” and that there is no “generally applicable standard requiring the use of a pre-determined threshold of contribution”; Appellate Body Report, *Colombia – Measures Relating to the Importation of Textiles, Apparel and Footwear*, WT/DS461/AB/R and Add.1, adopted 22 June 2016 (*Colombia Textiles*) para 5.72.

²⁵ Fontanelli (n 9) 57.

²⁶ Pirker (n 2) 306.

“there is always the possibility that the Appellate Body will be taken in by its own misdescription of what it has been doing and will start actually trying to balance the domestic benefits of a measure against the cost in reduced trade. That would be at odds with the whole spirit of the WTO agreements”.²⁷

If the AM test is being used to resolve “hard” cases, this suggests that it involves value judgments (in disguise) rather than being a mere technical exercise.²⁸

6.1.4 Summary

By introducing the W&B test in *Korea Beef*, the Appellate Body asserted a right to engage in balancing between different values and to second-guess the assessment of regulators. This was a highly controversial move which has been roundly criticised by WTO Members and observers. Adjudicators have wisely avoided controversy by refraining from resolving hard cases under the W&B test. As Regan notes, the law states that they have a right to commensurate, but they have so far refrained from exercising it.

Or have they? Rather than renouncing the right to commensurate, adjudicators have found a way to engage in it surreptitiously under the AM test. The rest of this chapter entails a short case study (*Korea Beef*) to show how the AM test can incorporate value judgments against the regulator.²⁹

6.2 AM Test: Korea Beef Case Study

6.2.1 Introduction

The theoretical analysis in chapter 5 identified the possibility of “easy” cases where the AM test can operate as a technical search for Pareto gains. Such cases are “easy” as adjudicators can find ways to make foreign producers better off, without

²⁷ Regan (n 4) 350.

²⁸ Several leading scholars have made this claim, including Sykes, Bown and Trachtman. Their views were described in the theoretical critique in the previous chapter.

²⁹ The next chapter will consider *Brazil Tyres* as a case study to show how the AM test can facilitate value judgments in favour of the regulator.

making the regulator worse off. Sykes has observed that such measures are rare, so adjudicators are often faced with “hard” cases requiring mediation between competing social values. Where and how does that mediation occur under the AM test?

Korea Beef demonstrates the complexities of a “hard” case where the regulator engaged in “balancing” (not just a narrow search for Pareto improvements) under the AM test.³⁰ My analysis will focus on three key techniques which the Appellate Body used in *Korea Beef* to make value judgments: (i) re-characterising the regulator’s chosen level of protection, (ii) allowing an alternative measure which achieved a lower level of protection; and (iii) allowing an alternative measure which created a significant financial or administrative burden.³¹

6.2.2 Critique 1 – Level of Protection

6.2.2.1 Re-characterising the Level of Protection

In *Korea Beef*, the Appellate Body re-characterised the level of protection which Korea claimed to pursue:

“We think it unlikely that Korea intended to establish a level of protection that *totally eliminates* fraud with respect to the origin of beef (domestic or foreign) sold by retailers. The total elimination of fraud would probably require a total ban of imports. Consequently, we assume that in effect Korea intended to *reduce considerably* the number of cases of fraud”.³²

This line of reasoning raises several problems. First, since the level of protection is a question of fact, the Appellate Body is not authorised to overrule the panel’s

³⁰ Interestingly, it is rare to see criticism of the outcome in *Korea Beef* (the dual retail system was widely viewed as unjustified protectionism), but there is much criticism of the legal reasoning.

³¹ The first two techniques will be discussed as a package for reasons which will become apparent.

³² Appellate Body Report, *Korea – Measures Affecting Imports of Fresh, Chilled and Frozen Beef*, WT/DS161/AB/R, WT/DS169/AB/R, adopted 10 January 2001 (*Korea Beef*) para 178.

characterisation in its role as the “trier of facts”.³³ The Appellate Body should work with the factual record as provided by the panel.³⁴

Second, it is strange that the Appellate Body assesses the level of protection “intended” by Korea. Is it seeking to determine Korea’s subjective motives? As discussed in the doctrinal section, the AM test should assess whether the proposed alternative would make an “equivalent contribution” to the regulator’s objective. Adjudicators should therefore use “actual contribution” as the benchmark (to address an empirical question) rather than “desired” or “intended” level of protection.

In *Korea Beef*, there was no contribution analysis to serve as the benchmark. Even though the Appellate Body introduced the W&B test in that dispute, it undertook no assessment of importance, contribution or trade-restrictiveness.³⁵ This did not prevent adjudicators from establishing an AM test benchmark, but they did so by constructing an “intended” level of protection rather than determining the “actual contribution”. This was a misstep.³⁶

Third, the Appellate Body fails to cite any meaningful evidentiary basis for its benchmark of “reduce considerably”. The finding is based on nothing more than a glib (inductive) statement that it would be impossible to eliminate all fraud without banning all imports. A determination of the contested measure’s contribution should be based on evidence and a justifiable (qualitative or quantitative) methodology. It is

³³ For a critique of the Appellate Body’s re-characterisation of the policy objective in *Brazil Tyres*, see Benn McGrady, ‘Necessity Exceptions in WTO Law: Retreaded Tyres, Regulatory Purpose and Cumulative Regulatory Measures’ (2008) 12 Journal of International Economic Law 153, 159. In principle, McGrady’s critiques could apply equally to the re-characterisation of the level of protection in *Korea Beef*.

³⁴ This criticism is not central to my overall claim. Even if the panel had characterised the level of protection as “reduce considerably”, the main criticisms of the Appellate Body (described below) would still apply.

³⁵ For example, Regan notes that the Appellate Body “never actually tells us anywhere in the opinion how important it thinks Korea’s goal of fraud-prevention is.” Regan (n 4) 360.

³⁶ As with the first criticism, this one is not central to my overall claim. Even if adjudicators had assessed the “actual contribution” (rather than the “intended” level of protection), my final two criticisms would remain equally applicable to the Appellate Body’s reasoning.

clear the Appellate Body sought to lower the panel's characterisation of the level of protection (from "elimination" to "considerable reduction"), but the Appellate Body failed to undertake any empirical work to justify this decision. The Appellate Body failed to determine how much the contested measure contributed to reducing fraud.³⁷

Fourth, the Appellate Body's approach is highly inconsistent with analogous cases. In *Asbestos*, Canada argued that France's level of protection was not really zero risk as there were some exceptions to the ban³⁸ and because the substitute products which would replace asbestos also carried certain health risks. Neither the panel nor Appellate Body was prepared to re-characterise the level of protection based on this argument.³⁹

Kapterian rightly notes that "if the logic of the Appellate Body in *Korea Beef* is applied" to *Asbestos*, it would have to change France's level of protection against Asbestos-related health risks from "totally eliminate" to "reduce it considerably".⁴⁰ France obviously could not eliminate such risks completely unless it implemented a blanket ban with no exceptions.⁴¹

The treatment of "level of protection" in *Korea Beef* is based on four key errors and contradictions. First, the Appellate Body changed a factual finding by the panel. Second, it used an aspirational ("intended") level of protection as the benchmark for proposed alternatives. Third, it lacked any evidentiary basis for the new characterisation of the level of protection. Fourth, it took a contradictory approach

³⁷ Perhaps the panel had failed to obtain meaningful evidence regarding the actual contribution (to serve as the benchmark) and the Appellate Body had insufficient facts to work with.

³⁸ Kapterian describes "a number of small exceptions to the ban"; Gisele Kapterian, 'A Critique of the WTO Jurisprudence on 'Necessity' (2010) 59 International & Comparative Law Quarterly 89, 109.

³⁹ See Alan Sykes, 'The Least Restrictive Means' (2003) 70 University of Chicago Law Review 403, 417.

⁴⁰ Kapterian (n 38) 109.

⁴¹ In *DR – Cigarettes* (n 5) para 7.229, the panel took a similar approach to *Korea Beef* when it questioned the regulator's actual level of enforcement. It found there was "no evidence to conclude that the tax stamp requirement secures a zero tolerance level of enforcement".

across different disputes (*Korea Beef* and *Asbestos*). Why did adjudicators make these elementary errors?

6.2.2.2 Lowering the Level of Protection

The previous section discussed the Appellate Body's decision to lower the factual finding about the level of protection produced by the contested measure. This section focusses on its comparison between the contested measure and the proposed alternative to see if the latter preserves the regulator's level of protection.

In *Korea Beef*, adjudicators found that

"Korea has not shown to the satisfaction of the Panel that measures, other than a dual retail system, compatible with the WTO Agreement, are not sufficient to deal with cases of misrepresentation of origin involving imported beef".⁴²

On what basis did the Appellate Body find that the proposed alternative would be "sufficient" to address Korea's policy concern regarding fraudulent meat marketing? There does not appear to be any evidence on the record supporting this conclusion. Interestingly, the Appellate Body appeared to put the onus on Korea to show that the proposed alternative would be less effective.⁴³ Ming Du rejects the notion that the alternative measure would make an "equivalent contribution":

"There is every reason to believe that the traditional measures that the AB suggested will not be as effective as the dual-retail system to better realize the Korean regulatory objective, even if more governmental resources are devoted to prevent passing off".⁴⁴

⁴² Panel Report, *Korea – Measures Affecting Imports of Fresh, Chilled and Frozen Beef*, WT/DS161/R, WT/DS169/R, adopted 10 January 2001, as modified by Appellate Body Report WT/DS161/AB/R, WT/DS169/AB/R (*Korea Beef*) para 659.

⁴³ In my view, the onus should be on the complainant to put forward proposed alternatives which respect the regulator's level of protection. However, if a complainant puts forward a proposed alternative which lowers the level of protection, WTO law appears to be vague on who has the onus of showing that the proposed alternative lowers the level of protection. Normally, the onus of proof should shift to the respondent if it wishes to argue that the proposed alternative is not "reasonably available".

⁴⁴ Michael Ming Du, 'Autonomy in Setting Appropriate Level of Protection under the WTO Law: Rhetoric or Reality?' (2010) 13 Journal of International Economic Law 1077, 1098. He further states: "the AB did not

He is supported by Weiler who states explicitly that the proposed alternative “would not achieve the same result”.⁴⁵ The Appellate Body’s finding that the proposed alternative would be equally effective appears speculative. There is no compelling evidence to support it. If the proposed alternative is a less effective measure, this suggests the Appellate Body lowered Korea’s level of protection under the AM test. This would violate a “fundamental principle” of WTO law: “the right that WTO Members have to determine the level of protection that they consider appropriate”⁴⁶

Bown and Trachtman propose two possible explanations for the Appellate Body’s approach in *Korea Beef*:

“First, it could be understood from an evidentiary standpoint, by which the Appellate Body is declining to accept at face value the chosen level of protection expressed by Korea, but is substituting a “real” chosen level of protection. Second, it could be understood as an instance in which the Appellate Body has found the national chosen level of protection to be unreasonable, and has judicially ‘reduced’ it for purposes of further analysis”.⁴⁷

In fact, the Appellate Body used both techniques cumulatively. First, it re-characterised, as a factual matter, the level of protection that was achieved by the contested measure. Second, it accepted a proposed alternative which achieved a lower level of protection.

compare the effectiveness of its suggested traditional measures with the ‘dual retail system’ in achieving Korea’s regulatory purpose” (at 1097).

⁴⁵ JHH Weiler, ‘Comment on Brazil - Measures Affecting Imports of Retreaded Tyres’ (2009) 8 World Trade Review 137, 140.

⁴⁶ Appellate Body Report, *Brazil Tyres* (n 16) para 210.

⁴⁷ Bown and Trachtman (n 13) 123.

6.2.2.3 The Margin of Appreciation

Regan defends the inconsistent approaches in *Korea Beef* and *Asbestos* by suggesting the Appellate Body applies a “margin of appreciation”.⁴⁸ He states:

“we might say the Appellate Body should be more deferential to the national regulator’s finding that some alternative is inadequate when the regulation aims at an important goal”.⁴⁹

Other authors agree that adjudicators apply a margin of appreciation. Kapterian suggests the AM test contains “two possible versions”: “a strict analysis” where adjudicators are highly-deferential and “a more lax test” where they are prepared to lower the level of protection.⁵⁰ Pirker argues that adjudicators are “more deferent” for “more important” values.⁵¹

Fontanelli agrees that adjudicators are highly-deferential for “human health” but that for other (seemingly less important) objectives:

“it is not unheard of that adjudicators, when ascertaining whether the less-restrictive alternative can meet the level of protection of the original measure, lower the ‘appropriate level of protection’ predetermined by the State, so as to make the alternative eligible”.⁵²

Ming Du interprets the jurisprudence as applying different levels of deference to different measures based on whether the “value at stake is high” (such as *Asbestos*, *Gambling* and *Brazil Tyres*) or whether the measure merely relates to “some less important interests” (such as *Korea Beef*).⁵³

⁴⁸ Regan (n 4) 352. The concept of “margin of appreciation” comes from the proportionality test in EU law. It is worth noting that in the EU, the margin of appreciation applies across the board as a way of deferring to regulators. Regan’s margin of appreciation at the WTO appears to vary depending on how important the policy objective is.

⁴⁹ *ibid.*

⁵⁰ Kapterian (n 38) 104.

⁵¹ Pirker (n 2) 290.

⁵² Fontanelli (n 9) 65-66.

⁵³ Ming Du (n 44) 1100-01.

However, there are two major problems with the “margin of appreciation” explanation. First, the Appellate Body has created no jurisprudence suggesting that there are varying degrees of deference under the AM test. WTO law does not consider “importance” a relevant factor under the AM test. In fact, the AM test should avoid making the regulator worse off (through a lower “contribution”) regardless of the objective.

Ming Du notes that “there is no textual warrant for such judgments” and that it leaves the regulatory autonomy of Members “in the discretionary hands of the WTO adjudicating bodies.”⁵⁴ Pirker agrees that “judicial reasoning on this tendency remains scarce”.⁵⁵ While it might be intuitively appealing that human health is more important than fraudulent meat marketing, this is not relevant to the AM analysis.

Second, even under the W&B test where “importance” is considered a relevant factor, Regan has argued that the Appellate Body (i) has no right to undertake such an analysis and (ii) does not really undertake this analysis in practice. If the Appellate Body offers a higher margin of appreciation for more important objectives, it does so on the basis of an unspoken (and unwritten) value judgment about the importance of different objectives.

There’s an inherent contradiction in Regan’s approach to the “margin of appreciation”. If the Appellate Body has no right to assess importance (and fails to do so), how can it rely on “importance” to grant a margin of appreciation under the AM test? Apart from intuitive feelings, how does the Appellate Body know (and justify) that asbestos-related health concerns are more important than fraudulent beef marketing concerns?

⁵⁴ ibid 1101.

⁵⁵ Pirker (n 2) 291.

Regan's descriptive claim - that adjudicators apply a margin of appreciation - may well be correct. However, his normative claim – that they “should” do so – has no basis in the law (or even the “spirit”) of the WTO.

6.2.3 Critique 2 - Financial and Administrative Burdens

WTO jurisprudence requires that any proposed alternative must be “reasonably available”. This means that the proposed alternative can entail additional financial or administrative burdens for the regulator, as long as it does not create an “undue burden”.

In *Korea Beef*, the Appellate Body determined that the proposed alternative could be equally effective if Korea “would devote more resources to its enforcement efforts on the beef sector”.⁵⁶ This shows that *Korea Beef* was a “hard” case (based on more than Pareto efficiency) as the judgment could make Korea worse off (by requiring it to reallocate financial and human resources). The key question was how much resource allocation would it be reasonable to impose before this would amount to an “undue burden”?

If the AM test can make Korea’s regulators worse off, is this not tantamount to a type of CBA which seeks to balance the interests of all stakeholders?⁵⁷ In his early writings, Trachtman argued that where an AM test includes a “reasonable availability” qualification, this “involves significant interpersonal comparison of utility, albeit avoiding the most difficult evaluation of the domestic regulatory goal per se.”⁵⁸ He seemed to suggest that the AM test shared some features of CBA (including

⁵⁶ Appellate Body Report, *Korea Beef* (n 32) para 180.

⁵⁷ This view is argued by Sykes, Bown and Trachtman (discussed below).

⁵⁸ Trachtman J, ‘Trade and... Problems, Cost-Benefit Analysis and Subsidiarity’ (1998) 9 European Journal of International Law 32, 68.

commensuration) but was less intrusive as the regulatory goal was exempted from review.

Regan built on Trachtman's approach, describing the AM test as "a limited sort of balancing":

"At this point the comparison that is called for between the actual measure and the alternative measure requires us to balance the extra administrative/enforcement cost of the alternative measure against the saving in trade cost from the alternative measure, but this is still not standard cost-benefit balancing".⁵⁹

Regan suggested this was less intrusive than CBA as it left the regulatory goal untouched:

"But if the only balancing they engage in does not involve balancing trade costs against the achievement of the underlying goal, but only against administrative/enforcement costs, then there is never any need for the Appellate Body to judge the importance of the underlying goal. This is a great advantage of the less-restrictive alternative approach".⁶⁰

Regan's view allows proposed alternatives to increase the resource burden, but not to lower the level of protection, which is consistent with the classic view that the policy objective and level of protection are not subject to review. This approach seemingly inquires whether the proposed alternative produces more benefit (in terms of trade openness) than additional costs for the regulator. If so, the proposed alternative is more efficient (and superior), even though the cost for the regulator goes up.

This seems like a technical exercise where the adjudicator merely compares two matters which are capable of being expressed in dollar terms: enforcement costs and trade costs. While this argument is potentially appealing, it contains some

⁵⁹ Regan (n 4) 354.

⁶⁰ ibid 349-350.

fundamental weaknesses. Even though enforcement and trade costs sound comparable (when discussed in abstract terms), they are not truly comparable.

No methodology on trade costs

The first problem is that WTO law contains no methodology for quantifying trade-restrictiveness costs. The notion of trade-restrictiveness remains poorly-developed in WTO law, even at a conceptual level.⁶¹ As shown in chapter 5, adjudicators are not required to quantify it in numerical terms.

In practice, adjudicators generally assess trade restrictiveness in qualitative terms. They reach vague conclusions that a measure has “a material impact on imports”⁶² or is “a trade restriction to the highest degree”.⁶³ Referring to the trade-restrictiveness finding as the calculation of a “trade cost” (as Regan does) is misleading, as adjudicators undertake no such calculation.

Enforcement costs: practice and jurisprudence

Further, adjudicators make no effort to calculate the additional enforcement costs imposed by proposed alternatives. In *Korea Beef*, the Appellate Body determined that the additional enforcement costs would not be unduly burdensome, but there is no evidence or reasoning to underpin this claim.

Weiler contests the Appellate Body’s view, and notes that “achieving the same result by other enforcement mechanisms will entail a far heavier administrative and

⁶¹ For a detailed discussion, see Tania Voon, ‘Exploring the Meaning of Trade-Restrictiveness in the WTO’ (2015) 14(3) *World Trade Review* 451.

⁶² Panel Report, *Brazil Taxation* (n 7) para 7.607.

⁶³ Panel Report, *Indonesia – Measures Concerning the Importation of Chicken Meat and Chicken Products*, WT/DS484/R and Add.1, adopted 22 November 2017 (*Indonesia – Chicken*) para 7.227.

financial burden".⁶⁴ Korea argued that it lacked the resources to increase enforcement in the beef sector, but adjudicators appeared unconvinced by this argument (without explaining why).⁶⁵ If the AM test compares trade and enforcement costs, why do adjudicators fail to quantify either of these factors in a way which could support meaningful comparison?

The jurisprudence does not support Regan's argument that the AM test balances trade costs against enforcement costs. The AM test merely requires that the proposed alternative may not create an "undue burden". This appears to resemble a *de minimis* requirement rather than a comparative requirement.

Regan's views about comparing costs could be based on the Appellate Body's reasoning in *Korea Beef* that the dual retail system involved "onerous shifting of enforcement costs which ordinarily are borne by the Member's public purse". This statement contains two elements which merit critical analysis. Was the dual retail system an "enforcement cost" and was it appropriate for regulators to pass that cost onto the private sector (including foreign producers)?

It is doubtful that Korea engaged in an "onerous shifting of enforcement costs". Its dual retail system, which required imported and domestic beef to be retailed separately, introduced a regulatory requirement rather than a "cost".⁶⁶ This regulatory requirement was imposed on Korean retailers who bore the regulatory burden in the first instance. It affected the competitive position of foreign beef by

⁶⁴ Weiler (n 45) 140.

⁶⁵ See eg Kapterian (n 38) 108.

⁶⁶ A conceptual tool which can help clarify this distinction comes from WTO rules on subsidies. We can invert the question and ask whether the dual retail system was a subsidy to Korean beef producers. Under WTO law, the answer would be no as it was a regulatory requirement which offered no "financial contribution". By the same token, the dual retail system did not impose financial costs on foreign beef producers, even though the regulatory requirement may have adversely affected their competitive position.

making it less attractive to Korean retailers. However, the Appellate Body stretches reality when it reframes this regulatory requirement as an “enforcement cost”.

Further, the Appellate Body’s reasoning contains a normative judgment. However, rather than making judgments about how much fraud a society should tolerate (which would impinge on Korea’s regulatory autonomy), it makes a normative judgment about how societies should allocate enforcement costs (when it states that these “ordinarily are borne by the Member’s public purse”).

Does such a norm about the allocation of enforcement costs truly exist? WTO case law suggests not. In *EC – Trademarks*, the panel explicitly affirmed the EU’s right to “continue to require producers to bear the [inspection] costs.”⁶⁷ This suggests that regulators retain the right to decide when it is appropriate to pass enforcement costs onto the private sector.

In the policy sphere where the WTO has a norm-making function (trade/customs matters), it explicitly recognises the right of governments to pass on enforcement costs to business. For example, it is legitimate for government to impose a customs service charge on private actors.⁶⁸

The Appellate Body’s normative claim that regulators should bear enforcement costs is highly implausible. Perhaps it felt more comfortable making a normative claim about enforcement costs than about the underlying policy objective. Nevertheless,

⁶⁷ Panel Report, *European Communities – Protection of Trademarks and Geographical Indications for Agricultural Products and Foodstuffs, Complaint by the United States*, WT/DS174/R, adopted 20 April 2005 (*EC – Trademarks and Geographical Indications (US)*) para 7.460.

⁶⁸ For example, the Protocol amending the Marrakesh Agreement establishing the World Trade Organization (adopted 27 November 2014, entered into force 22 February 2017) WTO Treaty Series No. 47, WT/Let/1030, WT/L/940 (Trade Facilitation Agreement) Art 6.2 allows governments to impose a customs service charge on importers, provided it is “limited in amount to the approximate cost of the services rendered”. There’s no norm that such enforcement costs be borne by governments.

the normative claim about the allocation of enforcement costs was equally intrusive of Korea's regulatory autonomy.

Does the AM test undermine policy objectives?

Regan suggests it is possible for adjudicators to impose regulatory costs without undermining regulatory goals. This is consistent with Trachtman's early writings which implied the policy objective could be ring-fenced from review under the AM test. However, in his later writings (alongside Bown), Trachtman's views evolved and he concluded that the AM test was tantamount to a CBA:

“Once “reasonable availability” is included, at least in connection with the assessment of alternative measures, it would be difficult to describe this language as anything but the announcement of a balancing test”.⁶⁹

Sykes strongly supports Bown and Trachtman's view of the AM test, noting that “WTO decisions to date strongly suggest that cost-benefit logic lies at the center of analysis.”⁷⁰

Regan suggests that the “reasonable availability” requirement is not as intrusive as a CBA. This is based on the notion that you can increase a regulator's financial and administrative burden, without undermining its policy objective or level of protection. In my view, this argument does not hold up to scrutiny. As Sykes and Trachtman strongly imply, increasing regulatory costs cannot be separated from the achievement of regulatory goals.

Regan's explanation treats the policy objective as sacred: the AM test focusses purely on enforcement and trade costs. However, imposing enforcement costs on a government can ultimately undermine its policy objectives or level of protection.

⁶⁹ Bown and Trachtman (n 13) 121.

⁷⁰ Sykes (n 39) 407.

When governments make political decisions to mediate between different social values, they generally do so by allocating (financial and human) resources. When Korea chose to use a (cheaper) regulatory requirement to manage beef fraud rather than a (more expensive) enforcement approach, it made a sovereign decision that its resources could be better allocated elsewhere. Perhaps it used those additional financial and human resources to subsidise healthcare or recruit food safety inspectors. Those were political decisions which required the Korean government to prioritise its resources among different social values.

The AM test allows additional burdens which are “reasonable” but is silent on where those resources should be reallocated from. If Korea must allocate more resources to beef fraud enforcement, it must reduce resources for other policy objectives. Assuming a government has finite resources, how is it possible that additional resource burdens will not undermine any of its objectives?

In order to allocate increased resources to beef fraud enforcement, Korea would have to divert resources from other policy objectives and ultimately lower its level of protection. Alternatively, Korea could choose not to “devote more resources to its enforcement efforts”, but this would lower its level of protection regarding beef fraud. Requiring a government to reallocate resources from one priority to another is highly-intrusive and strikes at the very heart of regulatory autonomy.⁷¹

6.2.4 Summary

Under the AM test, adjudicators do not acknowledge any mediation between different social values. Instead, they ask two questions: is the proposed alternative

⁷¹ In WTO disputes, these additional resources are discussed in conceptual terms only. The intrusiveness of this method is not fully explored because adjudicators are not required to consider where those resources will be reallocated from and what policy compromises they would therefore require.

as effective as the contested measure and is it reasonably available? These questions seem to be technical and innocuous, but they are actually loaded with value judgments, as was shown in the analysis above.

I argue that the *Korea Beef* adjudicators used a number of techniques to disguise their value judgment that the dual retail system was excessively protectionist. Adjudicators were able to rely on a (questionable) factual finding that the proposed alternative would be as effective as the contested measure.

Scholars have been critical of *Korea Beef*, suggesting that adjudicators underestimated the effectiveness of the contested measure and overestimated the effectiveness of the proposed alternative. Disagreements over factual findings are not uncommon although, in *Korea Beef*, leading scholars imply that certain critical findings had no evidentiary basis whatsoever. This is a deeply concerning accusation.

Adjudicators also managed the “optics” of their decision by incorporating their value judgment within the AM test, rather than the seemingly more intrusive W&B test. Adjudicators imply that the use of “traditional enforcement measures” instead of the dual retail system is a mere Pareto improvement which does not make the regulator worse off. In reality, there is a compelling argument that the “reasonably available” criteria essentially imports a “crude CBA” requirement into the AM test. Fontanelli notes that the AM test is “less value-neutral than the quasi-judicial bodies would claim it to be”, but that WTO Members do not appreciate this fact.⁷²

Regan seeks out a middle ground by suggesting the AM test contains “a limited sort of balancing”, but that it is less intrusive than CBA. In his view, adjudicators engage

⁷² Fontanelli (n 9) 39.

in a mere technical exercise where they compare the regulator's enforcement costs with the exporter's trade costs to find the most efficient measure. The Appellate Body has never enunciated its AM test in such terms nor does this model accurately describe what happens under the AM test.

Regan offers a second justification for the AM test's intrusiveness. He suggests that adjudicators consider the importance of a given objective leading to a "margin of appreciation" for more important objectives (such as health) or a lowering of the level of protection for less important objectives. If this is true, the AM test incorporates the most intrusive feature of CBA (a normative determination of importance) without any of the transparency or analytical rigour. Adjudicators are not required to articulate or justify their "importance" determination under the AM test.

The alternative view, advocated by Sykes, Bown and Trachtman, appears more compelling. An AM test with a "reasonable availability" qualification is tantamount to a CBA which requires adjudicators to make value judgments and mediate between competing values. This approach mediates by imposing costs, rather than lowering social utility, but the net effect is the same. After all, financial and human resources are the (finite and fungible) resources which adjudicators use to pursue utility.

The Appellate Body seemed to inadvertently admit the CBA logic behind the *Korea Beef* outcome when it concluded that the dual retail system was "a disproportionate measure not necessary to secure compliance with the Korean law against deceptive practices".⁷³ The reference to "disproportionate" strongly hints that the adjudicators were engaged in commensuration and balancing, rather than a technical search for Pareto improvements.

⁷³ Appellate Body Report, *Korea Beef* (n 32) para 179.

CHAPTER 7 – THE BRAZIL TYRES CASE STUDY

7.1 Introduction to *Brazil Tyres*

This chapter will focus on the techniques adjudicators use to make value judgments in favour of the regulator. In *Brazil Tyres*, adjudicators were required to determine whether an import ban on retreaded tyres was legitimate in order to prevent the negative health and environmental consequences of waste tyre disposal. The Appellate Body recognised that this dispute “illustrates the tensions that may exist between, on the one hand, international trade and, on the other hand, public health and environmental concerns”.¹

This was a hard case requiring mediation between different social values. In their landmark analysis,² Bown and Trachtman highlighted this tension between competing values:

“Member states of the WTO were serious both about allowing greater flexibility for national environmental measures, and about establishing some conditions so that this flexibility is neither unlimited nor abused”.³

Adjudicators had to “decide the scope of Brazil’s retained flexibility under WTO law to maintain an import ban on certain retreaded tyres.”⁴ Ultimately, that decision was made under the AM test, following a detailed W&B analysis.

This section will introduce a new technique which adjudicators can use to disguise their value judgments in the necessity test. It focusses on the use of double

¹ Appellate Body Report, *Brazil – Measures Affecting Imports of Retreaded Tyres*, WT/DS332/AB/R, adopted 17 December 2007 (*Brazil Tyres*) para 210.

² Weiler opens his analysis by stating: “The importance of the Bown and Trachtman analysis of *Tyres* transcends the specifics of the case itself. It is one of the most powerful and explicit indictments of the very competence of the Appellate Body to perform a task at the core of its judicial mission”; JHH Weiler, ‘Comment on *Brazil - Measures Affecting Imports of Retreaded Tyres*’ (2009) 8 World Trade Review 137, 137.

³ Chad Bown and Joel Trachtman, ‘Brazil – Measures Affecting the Import of Retreaded Tyres: A Balancing Act’ (2009) 8 World Trade Review 85, 86.

⁴ ibid 86.

standards, especially regarding the characterisation of the objective, and how this can significantly influence the outcome of a dispute in non-transparent ways.

This section will critique the adjudicators analysis under the AM test, including (i) the use of three different standards for both the policy objective (and “degree of contribution”); (ii) accusations that the objective was characterised too narrowly; (iii) the failure to establish an objective “magnitude” for the “degree of contribution”; and (iv) suggestions that adjudicators took an “impermissible” approach to characterising the policy objective.

I will further use methods from behavioural economics to draw conclusions about the adjudicators’ use of non-rational double standards. I will seek to identify whether the flawed reasoning of adjudicators was random or whether it reflected systematic bias. Finally, I’ll make some doctrinal suggestions designed to prevent confusion about how the AM test operates and how it interacts with the W&B test.

7.2 Critique of *Brazil Tyres*

7.2.1 Critique 1: Three Different Benchmarks

In chapter 5, I argued that the legal standard for the AM test had not been fully settled (in several key ways which will be further analysed below). I further argued that the AM test could be made more rigorous and predictable by importing its benchmarks from the W&B test. Any proposed alternative would have to match the “actual contribution” of the contested measure (as determined under the W&B test) to the regulator’s policy objective (as characterised under the W&B test).⁵

⁵ Complainants would also have to show that the proposed alternative was less trade-restrictive than the contested measure. The analysis of trade-restrictiveness (under the W&B test) could be used as the benchmark.

This section will show that the *Brazil Tyres* adjudicators failed to clarify a clear legal standard for the AM test. At different stages, they describe the legal standard as either an “equivalent contribution” or a “desired level of protection”. This lack of clarity creates dissonance between the W&B and AM tests and ultimately leads to the use of double standards by adjudicators.

Further, adjudicators fail to consistently apply the same benchmarks for the “contribution” and “objective”. In fact, I will show that they use three different standards for each of these questions.

The first standard

Under the “equivalent contribution” legal standard, adjudicators calculate the “actual contribution” under the W&B test. In *Brazil Tyres*, the panel took a qualitative approach to contribution and found that “the Import Ban is capable of making a contribution and can result in a reduction of exposure to the targeted risks” (the “first standard”).⁶ This determination should have established the benchmark for the AM test.

The first standard suffers from vagueness regarding both aspects of the AM benchmark (contribution and objective). It is vague about what the contested measure actually achieves (the contribution). It suggests that adjudicators should accept any proposed alternative which is “capable of making a contribution” or “can result in a reduction of exposure to the targeted risks”. If applied to the AM test, this would represent a low threshold and easy benchmark for proposed alternatives to meet.

⁶ Appellate Body Report, *Brazil Tyres* (n 1) para 149.

Further, the first standard is vague about what needs to be achieved by the proposed alternative (the policy objective). It refers to “targeted risks” but does not clarify what those risks are. Is the Appellate Body referring to the first order risks (accumulation of waste tyres) or the second order risks (mosquito-borne diseases)?

This vagueness makes it difficult to identify the benchmark for the AM test. Must the proposed alternative make an “equivalent contribution” to reducing waste tyres or to fighting mosquito-borne diseases? Under the latter approach, the range of potential alternatives would be much broader.

The panel’s conclusion on contribution contained unanswered questions: What was the contested measure’s degree of contribution?⁷ What policy objective did the contested measure pursue? Without these answers, it is not possible to meaningfully assess alternative measures.

The second standard

Strangely, the panel engages in further analysis to establish a second benchmark. It determined that “Brazil’s chosen level of protection is the reduction of the risks of waste tyre accumulation to the maximum extent possible” (the “second standard”).⁸ This second standard is based on “chosen” or “desired” level of protection rather than “equivalent contribution”.⁹

While the language is similar under both standards, there are also some important differences. In terms of the contribution, the first standard merely requires the proposed alternative to be “capable of making a contribution” while the second standard requires it to contribute “to the maximum extent possible”.

⁷ In the analysis below, Bown and Trachtman refer to this as “the magnitude”.

⁸ Appellate Body Report, *Brazil Tyres* (n 1) para 144.

⁹ I have argued above that the “equivalent contribution” approach is superior.

The second standard appears more ambitious (“maximum extent”) but it also contains a qualifier which makes it hard to assess the magnitude of the benchmark. This is because the term “possible” renders the benchmark value-laden rather than objective. The contribution which the proposed alternative must make depends on our interpretation of the highly-malleable term “possible”.¹⁰

There is also an important difference between the two standards in terms of the objective. The first standard vaguely refers to “targeted risks” which leaves discretion to adjudicators to assess either the first order (tyre) or second order (health) risks. The second standard mentions “risks” in general terms, but it also explicitly refers to “waste tyre accumulation”. This would suggest that any proposed alternative must address waste tyre accumulation and cannot operate to reduce the incidence of mosquito-borne diseases in ways unrelated to tyre reduction.

Under the contribution analysis, the panel therefore appears to establish two different standards which are capable of serving as the benchmark for the AM test. One standard is based on “actual contribution” of the contested measure, while the other is based on a constructed “desired level of protection”.

The third standard

To make matters more complicated, rather than relying on either of these benchmarks under the AM test, adjudicators actually established a third standard. The Appellate Body defined the AM benchmark as “reducing unnecessary generation of tyre waste”.¹¹ This means there are three standards in total.

¹⁰ In an extreme case, we could assume that banning all tyres would be impossible as this would be tantamount to banning the use of motor vehicles. But how do know which tyres it is “possible” to ban and which ones it is impossible to ban? This would seem to require a value judgment about the social utility of different tyres and their uses.

¹¹ Appellate Body Report, *Brazil Tyres* (n 1) para 67.

First Standard	“capable of making a contribution and can result in a reduction of exposure to the targeted risks”
Second Standard	“the reduction of the risks of waste tyre accumulation to the maximum extent possible”
Third Standard	“reducing unnecessary generation of tyre waste”

Regarding the degree of contribution, the third standard resembles the second one as they both take a value-laden approach. The regulator must either reduce tyres when “possible” (second standard) or prevent tyres which are “unnecessary” (third standard). Both approaches are vague and rely on value judgments (about what is possible/necessary) rather than objective magnitudes.

However, with respect to the objective, the third standard makes some significant changes. It retains the focus on waste tyres, but removes any reference to “risks” in general. This indicates that the first order (tyre) effects of the measure are the sole objective. The second order (health) effects cease to be directly relevant to the analysis.

The first two standards arguably left open the possibility of proposed alternatives which addressed the risk of mosquito-borne diseases without necessarily using waste tyres as the means of achieving this. The third standard, on the other hand, is crystal clear: any proposed alternative must contribute directly to waste tyre reduction.¹²

The second significant change is that even within the narrow focus on waste tyres, the standard shifts from reducing “accumulation” to reducing “generation”. This has the effect of further narrowing down the objective being pursued. The practical consequence of these different standards is clearly pointed out by McGrady:

¹² Especially by “reducing unnecessary generation of tyre waste”.

“The panel could have characterized Brazil’s regulatory goal in more general terms, for example, as being ‘to protect health by reducing the harmful impact of waste tyres’. Had the panel taken this approach, any number of measures including the clean-up of existing tyres, could have been construed as reasonably available alternatives to the ban”.¹³

McGrady’s analysis includes a practical example of the difference between the second and third standards: the “clean-up of existing tyres” could be a valid alternative under the second standard, but not under the third. This narrowing of the standard ultimately makes it harder for the EU to identify potential alternative measures.

The use of multiple standards creates two major problems. First, it is unclear what the objective is. It can be defined anywhere along a spectrum from the broad goal (human health and mosquito-borne diseases) to the narrow goal (unnecessary generation of waste tyres). Second, there is little clarity on the “degree of contribution” required by the proposed alternative as adjudicators failed to quantify what the contested measure achieved. This problem is exacerbated by the use of value-laden (rather than objective) benchmarks such as “possible” and “necessary”. These problems will be analysed further in critique 2 (policy objective) and critique 3 (“degree of contribution”).

7.2.2 Critique 2: The Objective is Too Narrow

Adjudicators established three different standards for characterising the objective.

Under the AM test, the Appellate Body ultimately used “reducing unnecessary generation of tyre waste”, the narrowest of the three standards. This standard is narrow for two main reasons. First, it limited the objective to the first order (tyre) objectives of the measure. Second, reducing “generation” is a narrower formulation

¹³ Benn McGrady, ‘Necessity Exceptions in WTO Law: Retreaded Tyres, Regulatory Purpose and Cumulative Regulatory Measures’ (2008) 12 Journal of International Economic Law 153, 157-8.

of the tyre-related objective than reducing “accumulation” (which would allow tyres to be generated, provided they are disposed of properly).

In his early writings, Trachtman foresaw that the characterisation of the objective could be highly significant:

“The more precisely defined the regulatory goal, and the more it refers both to maximizing regulatory benefits and minimising regulatory costs, the more difficult it will be to find a less trade restrictive alternative that will achieve the same complex goal”.¹⁴

This suggests that a narrow characterisation of the objective significantly favours the regulator under the AM test. In *Brazil Tyres*, Bown and Trachtman criticise the panel’s narrow focus on “waste tyre accumulation”, as it therefore “rejected alternatives that might contribute to health without reducing the number of waste tyres.”¹⁵

Fontanelli agrees that adjudicators should have taken a broad (health) focus rather than a narrow (tyre-related) focus:

“If Brazil had declared that health protection was the [objective], it would have been easier for the Panel to point at alternative less-restrictive measures that could ensure a similar or better result, and had nothing to do with disposed tyres. By focusing on tyre disposal as the ultimate objective, instead, Brazil managed to limit the Panel’s review to the tyre-reduction effect of the measure, drastically narrowing down the Panel’s margin of discretion in looking for alternative measures”.¹⁶

McGrady agrees that “the narrower the characterization of a regulatory goal, the less likely it is that a panel will find reasonably available alternatives to the impugned

¹⁴ Trachtman J, ‘Trade and... Problems, Cost-Benefit Analysis and Subsidiarity’ (1998) 9 European Journal of International Law 32, 73.

¹⁵ Bown and Trachtman (n 4) 127. Their critique would have been even stronger if they had chosen the narrowest formulation (“reducing unnecessary generation of tyre waste”) rather than “waste tyre accumulation”.

¹⁶ Filippo Fontanelli, ‘Necessity Killed the GATT - Art XX GATT and the Misleading Rhetoric about Weighing and Balancing’ (2012) 5 European Journal of Legal Studies 39, 60.

measure.”¹⁷ He suggests the objective was characterised too narrowly in *Brazil Tyres* and expresses a preference for a composite approach combining both health and tyres:

“The panel could have characterized Brazil’s regulatory goal in more general terms, for example, as being ‘to protect health by reducing the harmful impact of waste tyres’. Had the panel taken this approach, any number of measures including the clean-up of existing tyres, could have been construed as reasonably available alternatives to the ban”.¹⁸

Bown and Trachtman highlight that “there may be other, less trade restrictive, ways to achieve the same health effects without reducing the number of waste tyres”.¹⁹

These arguments suggest that the adjudicators’ choice about how to characterise the objective impact the dispute in critical ways. If adjudicators had adopted a broader characterisation of the objective, the EU likely would have been able to demonstrate that a proposed alternative was reasonably available. Is it really possible for this seemingly innocuous question of fact (the characterisation of the objective) to be the decisive factor which determines the outcome of a dispute?

Even though these authors all criticise the narrow characterisation of the objective, they do not necessarily agree on what the correct characterisation should have been. There does not appear to be any clear protocol which meaningfully constrains how adjudicators exercise their discretion to characterise the objective. On what grounds can you accuse an adjudicator of characterising the objective too narrowly (or incorrectly)?²⁰

¹⁷ McGrady (n 13) 157.

¹⁸ McGrady (n 13) 157-158.

¹⁹ Bown and Trachtman (n 4) 124.

²⁰ I will return to this matter in critique 4.

7.2.3 Critique 3: The Benchmark Contained No “Magnitude”

This section critiques the adjudicators consideration of the “degree of contribution”, that is “how much” the contested measure contributed to the objective. Adjudicators used three different standards for contribution, yet failed to establish a magnitude which was capable of serving as an objective benchmark.

In *Brazil Tyres*, the Appellate Body confirmed the discretion of Panels to determine their preferred methodology and to favour a qualitative approach over precise quantification:

“we do not accept the European Communities' contention that the Panel was under an obligation to quantify the contribution of the Import Ban to the reduction in the number of waste tyres and to determine the number of waste tyres that would be reduced as a result of the Import Ban”.²¹

While Bown and Trachtman accept the use of a qualitative approach, they argue that the benchmark in *Brazil Tyres* is deficient nonetheless:

“the simple dichotomy between quantitative and qualitative evaluation fails to recognize that there is something in between: a non-quantitative assessment of magnitude. It is impossible to determine whether an alternative measure could contribute as much as the existing measure without some assessment of magnitude”.²²

Bown and Trachtman insist that the absence of a “magnitude”, even an imprecise one,²³ renders the AM test unworkable:

“What can possibly be meant by an ‘equivalent contribution’ when no assessment of the magnitude of the Brazilian measure’s contribution has been made?”²⁴

²¹ Appellate Body Report, *Brazil Tyres* (n 1) para 147.

²² Bown and Trachtman (n 4) 125.

²³ To understand how an imprecise “qualitative assessment” can still produce some sort of “magnitude”, *Korea Beef* provides some guidance. In that case, the regulator sought to “reduce considerably” certain types of fraud (this was analysed in section 6.2.2.1). This is expressed in qualitative terms, but it still identifies a magnitude which is capable of serving as a benchmark (albeit not a precise quantitative one). It allowed adjudicators to inquire whether the proposed alternative would also “reduce considerably” the incidence of fraud in the beef sector.

²⁴ Bown and Trachtman (n 4) 126.

This problem can be resolved by looking to the W&B jurisprudence. After all, the W&B test requires the contested measure to be analysed in order to determine the “extent” or “degree” of contribution. This establishes a “magnitude” which can be used for analytical purposes under the AM test.

In *Brazil Tyres*, adjudicators adopted three different standards for the degree of contribution, but it is doubtful that any contain a “magnitude” which is capable of serving as an objective benchmark:

Legal Element	Contested Measure’s Contribution to Policy Objective	Is There a Magnitude?
Contribution (first standard)	“capable of making a contribution”, “can result in a reduction”	Doubtful
Chosen Level of Protection (second standard)	“to the maximum extent possible” ²⁵	No
AM Test (third standard)	“unnecessary tyre waste”	No

Under the first standard, the import ban is “capable of making a contribution” and “can result in a reduction of exposure to the targeted risks”. This standard is equivocal on whether the measure will make a contribution or reduce risk. Even if the measure does make some contribution, there is little indication of how much that contribution is (the magnitude). It remains possible that the measure’s contribution might be insignificant or marginal. This standard appears incapable of serving as a benchmark.

As a side note, it is worth recalling that adjudicators in *Brazil Tyres* established that due to the highly-restrictive nature of the measure (an import ban), it was required to make a “material contribution”. It is unclear how “capable of making a contribution” could possibly satisfy a “material contribution” threshold.

²⁵ Appellate Body Report, *Brazil Tyres* (n 1) para 144.

The second standard requires non-accumulation of waste tyres “to the maximum extent possible” while the third standard requires non-generation of “unnecessary” waste tyres. The quintessential problem with these standards, as Bown and Trachtman explained, is that they do not contain a magnitude. Rather, they use terms such as “possible” and “unnecessary” which require us to make a value judgment about what magnitude is appropriate.

If the standard was “reducing generation of waste tyres”, an adjudicator could certainly use this to compare which measures are effective and which ones are not. But by adding the qualifier “unnecessary”, this seemingly objective standard becomes dependent on our subjective perceptions of necessity. Ironically, the “necessity test” ultimately turns on a judgment about whether certain waste tyres are “necessary”. This is the epitome of circular logic.

Neither the “possible” nor “unnecessary” standard is capable of operating as an objective benchmark. They are both subjective standards which are capable of being manipulated. While there might not be a requirement for the magnitude/degree to be precise, or converted into numerical terms, there must be some standard of an objective nature which can be used for analytical purposes. Neither the second nor third standard is capable of satisfying this requirement.

Did Brazil have an estimate of the likely impact of the measure in terms of fewer cases of mosquito-borne diseases or the reduction of waste tyres? Adjudicators should have probed this question more deeply.

7.2.4 Critique 4: The Benchmark was “Impermissible”

Several scholars suggest that Brazil may have succeeded under the AM test, purely because the objective was defined narrowly.²⁶ A broad characterisation might have seen the EU win the dispute under the AM test.

Bown and Trachtman subscribed to this view and made the additional claim that “non-generation of waste tyres alone is an impermissible level of protection”.²⁷ This is a strong accusation, especially considering that the WTO has so few protocols regarding how adjudicators should characterise the objective or level of protection. This seems to be a matter where adjudicators have much discretion.

Bown and Trachtman further clarify that “reducing the number of waste tyres cannot be understood as a level of protection, but only as a means of protection.”²⁸ They seem to argue that adjudicators cannot characterise the objective purely by reference to the first order (tyre-related) goal, but must do so by reference to the health-related goal. Perhaps this is because the legitimate public interest in Article XX(b) refers to health but not to tyres. If this is their argument, it proposes a potential solution by creating a protocol for the characterisation of objectives which links them more closely to the language in Article XX.

In the next section, I will propose an alternative method which draws from behavioural economics. I will argue that the *Brazil Tyres* adjudicators should have refrained from using double standards (and “anomalies”²⁹) in their characterisation of the objective.

²⁶ This matter was discussed under critique 2.

²⁷ Bown and Trachtman (n 4) 129.

²⁸ ibid 124.

²⁹ This term comes from behavioural economics and will be defined in the next section.

7.3 Behavioural Economics, “Anomalies” and Double Standards

7.3.1 Thaler, Anomalies and the Law

When Richard Thaler challenged standard economic theory, one of the techniques he used was to identify “anomalies”.³⁰ An anomaly “is a fact or observation that is inconsistent with the theory”.³¹ Notably, Thaler used anomalies to challenge the assumption that economic agents are rational.³² A well-known example is the following:

“A wine-loving economist we know purchased some nice Bordeaux wines years ago at low prices. The wines have greatly appreciated in value, so that a bottle that cost only \$10 when purchased would now fetch \$200 at auction. This economist now drinks some of this wine occasionally, but would neither be willing to sell the wine at the auction price nor buy an additional bottle at that price”.³³

The “rationality” assumption in economics implies “the near equality of buying and selling prices”.³⁴ In this scenario, *homo economicus* should determine an objective value for the Bordeaux wine to serve as both the ceiling for wine purchases and the floor for wine sales.³⁵ The variance between the “buy” and “sell” prices suggests that the wine-loving economist did not have an objective view of value and was influenced by non-rational considerations.

Judicial decisions also contain “anomalies” which cannot be explained rationally, though lawyers tend to refer to them as “double standards”. I relied on this tool in the *Brazil Tyres* analysis when I found that adjudicators did not have a fixed view on

³⁰ Thaler collated these anomalies in Richard Thaler, *The Winner's Curse: Paradoxes and Anomalies of Economic Life* (Simon & Schuster 1992).

³¹ ibid 3-4.

³² The two key assumptions Thaler sought to challenge were “rationality” and “self-interest”. While the self-interest assumption is not necessarily relevant to the decisions of judges, the rationality assumption certainly is.

³³ Daniel Kahneman, Jack Knetsch and Richard Thaler, ‘Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias’ (1991) 5(1) *The Journal of Economic Perspectives* 193, 194.

³⁴ Thaler (n 30) 6.

³⁵ Behavioural economists refer to this as the “willingness to pay” (WTP) and “willingness to accept” (WTA). As Thaler notes, there should be “near equality” between them. See Kahneman et al (n 33) 194.

the policy objective.³⁶ Rather, the adjudicators' characterisation of the objective seemed to fluctuate between a number of different standards.³⁷ Does this suggest that the *Brazil Tyres* adjudicators are as irrational as the wine-loving economist?

To find anomalies, all Thaler required was the curiosity to observe non-rational behaviours other economists ignored or considered uninteresting.³⁸ This section will show that "double standards" in law have many similarities with Thaler's "anomalies". Interestingly, to demonstrate that the economist is non-rational, Thaler does not seek to establish the true value of the wine. Instead, he simply observes that it should have a "fixed" objective value regardless of whether it is being bought or sold.

Should the same principle apply to legal double standards?

In my analysis of *Brazil Tyres*, I did not make claims about which characterisation of the objective was correct.³⁹ Rather, I argued that there should be a single "fixed" definition which applies for all legal elements.⁴⁰ When legal scholars identify a "double standard" in a judgment, it may hint that the adjudicator is influenced by non-rational factors.

Uncovering judicial "anomalies" can be pain-staking work. In the brief paragraph (cited above), Thaler effectively demonstrated the wine "anomaly" in a way which

³⁶ Or the degree of contribution, though I will focus on the policy objective in this section.

³⁷ See McGrady (n 13) 158 where he used a similar technique to conclude that the Appellate Body applied a double standard regarding the objective. He stated that the "non-generation" standard "is used only in analysis of reasonably available alternatives. The phrase is not used earlier in passages that seek to characterize the regulatory goal for purposes of determining whether the measure fell within the ambit of Article XX(b)".

³⁸ Thaler invited readers of the *Journal of Economic Perspectives* to share their anomalies with him. He eventually found so many that he was no longer able to publish them all and had to deal with the "the dilemma of choosing which juicy anomaly to discuss"; Kahneman et al (n 33) 193.

³⁹ An alternative method of critique would be to suggest that the economist valued the wine incorrectly or that the *Brazil Tyres* adjudicators characterised the objective incorrectly. Bown and Trachtman use this alternative method when they suggest adjudicators erred by characterising the objective based on the intermediate "means of protection" rather than the ultimate goal. This type of normative critique may be useful, but my analysis focusses on Thaler's method of finding "anomalies" and "double standards".

⁴⁰ I believe McGrady is making a similar "double standards" argument.

should be intuitively-obvious to most readers. By contrast, the double standard in *Brazil Tyres* is much harder to identify or render intuitively-obvious to readers.

It is understandable why judicial “anomalies” are hard to identify. Many adjudicators are adept at using equivocal language or rhetorical devices to disguise any anomalies in their reasoning. Even if judges are influenced by non-rational or unconscious considerations, they are trained to present their conclusions as flowing logically from rational arguments. They are unlikely to explicitly acknowledge that their decision was based on non-rational factors (and they might not even consciously realise this fact).

If the *Brazil Tyres* adjudicators openly acknowledged their use of double standards, the job of legal scholars would be easy. However, the use of disguised double standards creates an important role for legal scholars to uncover them. As an intermediate step, scholars may need to “deconstruct” or “reframe” the adjudicator’s reasoning to make visible the underlying double standard.⁴¹

A further barrier to identifying legal “anomalies” is that it requires deep expertise of the relevant area of law. If a lawyer with little WTO expertise (or a non-lawyer) read *Brazil Tyres*, it is far from certain they would identify the double standards at play.

Even experts in the field may fail to uncover these double standards. In *Brazil Tyres*, Bown and Trachtman claimed that adjudicators “failed to clarify the regulatory goal”,⁴² while Fontanelli accused them of “interlacing health and tyre-waste reduction within one policy objective”.⁴³ These leading scholars (rightly) criticised the

⁴¹ In chapter 1, I provided an explanation of the critical tool of “deconstruction”.

⁴² Bown and Trachtman (n 4) 120.

⁴³ Fontanelli (n 16) 60.

treatment of the policy objective, but stopped short of accusing adjudicators of double standards.

In my analysis, in order to identify double standards, I had to struggle with the fact that WTO law remains unclear about which legal elements treat the policy objective as a relevant object of review. For example, the “contribution” analysis clearly treats the policy objective as relevant, but the AM test is based on competing jurisprudence. Under the “equivalent contribution” strand of jurisprudence, the policy objective would be an object of review, but this would not be the case under the “desired level of protection” jurisprudence.

Bown, Trachtman and Fontanelli produce rigorous and highly-insightful critiques of *Brazil* *Tyres*, including the weaknesses in the policy objective, but they refrain from exposing “anomalies”. Is this because they failed to observe the double standard or because they would contest my interpretation that a double standard exists?

It would not be surprising if scholars disagree about the presence of legal “anomalies” as much depends on individual interpretation. In the rest of this section, I will seek to “replicate” McGrady’s view that the *Brazil* *Tyres* adjudicators relied on a double standard. I will seek to show that the double standard runs even deeper than he suggests.

Further, I will present these double standards in a (hopefully) easy-to-understand table. While my presentation of the *Brazil* *Tyres* double standard will not be as intuitively-obvious as Thaler’s “wine-loving economist”, I hope to present it in a way which is sufficiently clear for other WTO experts to either validate the “anomaly” or explain why they contest its existence. Once I demonstrate the existence of this

“anomaly” in *Brazil Tyres*, I will make further comments on how Thaler’s method can assist the study of “behavioural adjudication”.⁴⁴

7.3.2 Replicating the *Brazil Tyres* Anomaly

Under the necessity test, the policy objective should arise in at least four places:

design step, importance, contribution and the AM test. In *Brazil Tyres*, it arguably arose a fifth time (unnecessarily) under the “chosen level of protection” analysis.⁴⁵

For some of these elements, the object of review is contested.⁴⁶

In theory, the characterisation of the policy objective is a question of fact which should remain fixed but, in practice, it sometimes fluctuates for different legal elements. In section 7.2.1, I identified three characterisations of the policy objective (for “contribution”, “chosen level of protection” and the AM test). I will now describe the adjudicators’ characterisation of the policy objective for two remaining legal elements: design step and importance. This will produce a total of five “data points” to assess the anomaly.

Design Step

Under the design step, the Panel concluded that:

“the policy objective behind the import ban is the protection of human life and health and the environment and that it is designed to prevent the generation of additional amounts of waste tyres in Brazil and, by so doing, to reduce the incidence of cancer, dengue, yellow fever, respiratory diseases, reproductive problems, environmental contamination, and other risks associated with waste tyres”.⁴⁷

⁴⁴ In Chapter 1, I introduced the notion of “behavioural adjudication” and what it entails.

⁴⁵ In section 7.4.3, I will explain why this element should be redundant.

⁴⁶ This was addressed in chapter 2, but I will briefly recap the possible objects of review here; (i) Design step: “measure” or “policy objective”; (ii) Importance: “policy objective”, “societal value” or even “public interest”; (iii) AM test: “equivalent contribution to policy objective” or “desired level of protection”.

⁴⁷ Panel Report, *Brazil – Measures Affecting Imports of Retreaded Tyres*, WT/DS332/R, adopted 17 December 2007 (*Brazil Tyres*) para 7.99.

For this legal element, the panel defines the objective broadly: “the protection of human life and health and the environment”.⁴⁸ It also mentions other intermediate goals (“reduce the incidence of cancer, dengue, yellow fever”), while waste tyre management appears to be a mere “means of protection” to achieve these goals.⁴⁹

Importance

For importance, adjudicators also adopt a broad characterisation of the objective. The Appellate Body confirmed the Panel’s importance conclusion that “few interests are more ‘vital’ and ‘important’ than protecting human beings from health risks”.⁵⁰ By referring to the health (rather than tyre) objective, the adjudicators enabled Brazil to demonstrate a high level of importance.⁵¹

The adjudicators’ five different characterisations of the objective can be summarised as follows:

Legal Element	Characterisation of Objective
Design step	“the protection of human life and health and the environment” and “reduce the incidence of cancer, dengue, yellow fever, respiratory diseases, reproductive problems, environmental contamination, and other risks associated with waste tyres”
Importance	“protecting human beings from health risks”
Actual Contribution	“a reduction of exposure to the targeted risks”
Chosen Level of Protection	“reduction of the risks of waste tyre accumulation”
AM Test	“reducing unnecessary generation of tyre waste”

⁴⁸ If adjudicators had characterised the objective narrowly as “reducing unnecessary generation of tyre waste” (the AM test definition), how would this have affected the examination under the design step? When Bown and Trachtman treat the narrow (tyre-related) objective as “impermissible”, this could be their way of suggesting it would fail the design step.

⁴⁹ This draws from Bown and Trachtman’s view that waste tyre management is a mere “means of protection”.

⁵⁰ Appellate Body Report, *Brazil Tyres* (n 1), para 144.

⁵¹ Presumably it would have been harder to demonstrate importance if “reducing unnecessary generation of tyre waste” had been the objective (as was the case under the AM test).

This analysis shows that adjudicators used five different characterisations of the policy objective. There is some overlap between objectives, and they all have a link to health and/or tyres, but there are also some important differences. There is a big gap in the spectrum from “protecting human beings from health risks” (very broad) to “reducing unnecessary generation of tyre waste” (very narrow).

This suggests that, like the “wine-loving economist”, the *Brazil Tyres* adjudicators were subject to non-rational influences. Was the fluctuating definition of the objective due to randomness (such as judicial sloppiness or incompetence) or was there a causal factor which motivated the adjudicators? Let’s turn to behavioural economics to see if it can help answer this question.

7.3.3 Anomalies plus Psychology

Observing that economic agents (or judges) take into account non-rational considerations is an interesting, but far from revolutionary, discovery. Thaler did not win his Nobel Prize for showing that humans sometimes behave irrationally.⁵² However, when he drew on insights and techniques from cognitive psychology (especially from Kahneman and Tversky), he was able to test if seemingly irrational behaviour was random or if it reflected systematic cognitive biases. This gave birth to the behavioural economics revolution.

Behavioural economists were able to demonstrate that the behaviour of the “wine-loving economist” was not “random”. They used empirical methods to demonstrate the existence of an “endowment effect” whereby “people often demand much more

⁵² Richard Thaler won the Nobel Prize in 2017. According to the Nobel Committee: “Richard Thaler has analyzed economic decision-making with the aid of insights from psychology. He has paid special attention to three psychological factors: the tendency to not behave completely rationally, notions of fairness and reasonableness, and lack of self-control. His findings have had a profound influence on many areas of economic research and policy.”

to give up an object than they would be willing to pay to acquire it".⁵³ This is consistent with the cognitive bias of "loss aversion" which finds that "changes that make things worse (losses) loom larger than improvements or gains".⁵⁴

Thaler extracted two useful insights from his observation (and follow up study) of the wine-loving economist. First, anomalies can be used to identify non-rational behaviour. Second, non-rational behaviour might be systematic (and therefore not random) when it can be explained by reference to an underlying bias.

Up to this point, my analysis has drawn from the first insight: the fluctuating policy objective in *Brazil Tyres* illustrates non-rational behaviour. However, is it also possible for legal scholars to demonstrate that non-rational behaviour by judges was driven by an underlying bias?

In *Brazil Tyres*, the leading scholars hint at bias. McGrady speculates that the "analysis may have been directed improperly so as to achieve a desired outcome".⁵⁵ Fontanelli, Bown and Trachtman also observed that the policy objective was characterised narrowly (under the AM test) in a way which favoured Brazil. Is there a rigorous method that legal scholars can use to test for systematic bias?

In *Brazil Tyres*, we can test to see if adjudicators were biased in favour of Brazil by asking whether the changing characterisation of the objective systematically favoured it. The table below reproduces the five characterisations of the objective with two additional questions: (i) whether the objective is characterised in a narrow or broad manner; and (ii) which party was favoured by the chosen approach.

⁵³ Kahneman et al (n 33) 194.

⁵⁴ ibid 199. This leads economic agents to systematically value a given object more highly when they possess it than when they are seeking to acquire it.

⁵⁵ McGrady (n 13) 159. He further criticises the "absence of a methodology underlying the analysis" (at 163) and the "unpredictable" application of the law (at 158).

Legal Element	Characterisation of Objective	Narrow or Broad	Which Party is Favoured?
Design Step	“the protection of human life and health and the environment” and “reduce the incidence of cancer, dengue, yellow fever, respiratory diseases, reproductive problems, environmental contamination, and other risks associated with waste tyres” ⁵⁶	Broad	Brazil
Importance	“protecting human beings from health risks” ⁵⁷	Very Broad	Brazil
Actual Contribution	“a reduction of exposure to the targeted risks” ⁵⁸	(Vague)	(Unclear)
Chosen Level of Protection	“reduction of the risks of waste tyre accumulation”	Narrow	Brazil
AM Test	“reducing unnecessary generation of tyre waste”	Very Narrow	Brazil

This analysis may suggest that the fluctuating characterisation of the policy objective reflected systematic bias in favour of Brazil. Adjudicators characterised the objective in a broad manner when this was regulator-friendly and in a narrow (or very narrow) manner when this was regulator-unfriendly. It seems unlikely this was motivated by pure randomness.

7.4 Lessons from *Brazil Tyres*

7.4.1 The Policy Objective in WTO Law

As a matter of legal (and logical) reasoning, adjudicators should avoid double standards and “anomalies”. Where they resort to double standards, this may well reveal the presence of non-rational considerations. It also provides raw data for further analysis.

⁵⁶ Panel Report, *Brazil Tyres* (n 47) para 7.99.

⁵⁷ Appellate Body Report, *Brazil Tyres* (n 1) para 144.

⁵⁸ *ibid* para 149.

I have identified a type of double standard, regarding the policy objective under the necessity test, which can be applied to any Article XX dispute to see if one party was systematically favoured by a fluctuating policy objective. This is because a narrow policy objective favours the regulator for some elements (contribution and the AM test), while a broad policy objective favours it for others (the design step and importance).

If the policy objective oscillates from broad to narrow according to this pattern, it may well indicate that (i) there are non-rational factors at play and (ii) those factors systematically favour the regulator. Adjudicators can tip the scales in favour of one party by making the policy objective narrower or broader (like an accordion) depending on which legal element is being assessed.

The “anomaly” tool allows us to test the impartiality of a necessity test analysis, without having to take a view on what the policy objective should be.⁵⁹ As long as adjudicators apply a consistent definition of the policy objective, the necessity test appears to contain an internal check which can prevent abuse.⁶⁰ WTO adjudicators should be required to avoid anomalies by using the same characterisation of the objective across all legal elements: design step, importance, contribution and the AM test.

This is particularly important for the benchmark in the AM test which is often decisive in necessity cases. In *Brazil Tyres*, the AM analysis relied on a characterisation of the objective which was different from (and far narrower than) the characterisation

⁵⁹ Just like Thaler had no view on the value of the Bordeaux wine, but could still find that the “wine-loving economist” had an irrational view of its value.

⁶⁰ Adjudicators have an important role to ensure that the policy objective is characterised in a reasonable manner which enables them to analyse the issues under dispute. Excessively broad or narrow objectives could make analysis difficult. However, as long as adjudicators choose a reasonable objective, their most important responsibility is to apply it consistently across the different legal elements.

for other legal elements. Rather than inventing a new benchmark solely for the AM test (an approach which is susceptible to manipulation), the benchmark should be based on earlier factual findings by adjudicators. The AM test should import its benchmarks from the W&B analysis (for the policy objective and degree of contribution) to bring consistency and predictability to the analysis.

7.4.2 WTO Adjudication and Value Judgments

Certain elements of the necessity test are highly controversial and are viewed as a threat to regulatory autonomy. Many scholars denounce the right of adjudicators to assess importance or engage in either overt balancing (under the W&B test) or covert balancing (under the AM test). The pragmatic benefit of the fluctuating policy objective was that it enabled adjudicators to address value-laden questions in a cursory manner which appeared formalist rather than intrusive.

On importance, adjudicators found that “few interests are more ‘vital’ and ‘important’ than protecting human beings from health risks”.⁶¹ By adopting an extremely broad objective, adjudicators were able to find in favour of the regulator with virtually no analysis. Their normative finding, that human health is important, was non-controversial (and intuitively-obvious).

If adjudicators had considered the importance of the narrow objective (“reducing unnecessary generation of tyre waste”), their “importance” analysis would likely have been more detailed, intrusive and controversial. Perhaps, they would have needed to make some borderline normative judgments calls. How would Members and scholars have reacted if adjudicators had determined that this objective was insufficiently important to justify an import ban?

⁶¹ Appellate Body Report, *Brazil Tyres* (n 1) para 144.

Under the AM test, adjudicators found that there were no alternatives which could make an equivalent contribution to “reducing unnecessary generation of tyre waste”. Considering how narrow the objective was, this conclusion seems reasonable. But what if adjudicators considered proposed alternatives which could make an equivalent contribution to “protecting human beings from health risks”?

This would have opened up an expansive universe of proposed alternatives. Bown and Trachtman considered whether “a production subsidy on retreads of once-used Brazilian tyres” or “a tariff on imports of retreads” would be more effective,⁶² but why stop there? Under such a broad characterisation, perhaps Brazil could have achieved an equivalent contribution to “protecting human beings from health risks” by distributing mosquito nets, raising tobacco taxes or subsidising cancer research. Under the AM test, a broad characterisation of the objective would have opened the door to highly-intrusive analysis.

It is highly-implausible to believe that adjudicators confirmed the legitimacy of Brazil’s import ban without engaging in normative analysis and balancing. However, WTO law remains based on a myth that such normative analysis is not permitted. This is why the W&B test, which explicitly requires balancing (and an importance analysis), has been so controversial. Some scholars applaud the fact that this test has not been operationalised, but this does not eliminate normative value judgments by adjudicators; it just sends them underground.

Rather than justifying Brazil’s measure pursuant to a meaningful analysis of importance and balancing, adjudicators reached this conclusion by manipulating the definition of the policy objective. Adjudicators did not renounce their right to make

⁶² Bown and Trachtman (n 4) 90.

normative value judgments, but simply made those judgments surreptitiously. They have now developed techniques to embed value judgments in their decisions through flawed legal reasoning (such as double standards on the policy objective), rather than through explicit statements about their true reasons.

When we identify “anomalies” in the legal reasoning of WTO adjudicators, the job of legal scholars is not complete. We should further inquire what this non-rationalist legal reasoning reveals about the disguised value judgments (and real reasons) of adjudicators. In *Brazil Tyres*, the use of double standards does not reveal a random lack of rationality; rather, it hints at a deeper problem which merits further investigation.

7.4.3 Brazil Tyres and Doctrinal Law

There are a number of doctrinal matters where the “necessity” jurisprudence requires clarification and I will make some proposals. I do not necessarily seek changes to the law - in some cases I merely call on adjudicators to explicitly state which of the current competing strands of jurisprudence is correct. After a quarter of a century, the Appellate Body should cease treating WTO law as a menu and start to settle on some established tests. Here are my recommendations.

The AM test should use the “equivalent contribution” benchmark and explicitly disavow the “desired level of protection”

There are two main benefits to this approach. First, as Regan highlights, it ensures that the test looks at the actual contribution of the contested measure rather than a hypothetical, aspirational level of protection. Second, it would require adjudicators to use consistent reasoning across both the W&B and AM tests on critical matters such

as the characterisation of the objective and the degree of contribution. This would prevent anomalies (or at least make them easier to spot).

The “desired level of protection” jurisprudence opens the door to adjudicators constructing a new benchmark for the AM test. This divorces the AM benchmark from the “actual contribution” and allows for the manipulation of results. If adjudicators wish to favour the regulator, they can simply invent a new, narrow policy objective under the AM test.

In *Colombia Textiles*, the Appellate Body stated that a panel must “compare the challenged measure and possible alternative measures that achieve the same level of protection while being less trade restrictive.”⁶³ While retaining the reference to “level of protection”, this represents an evolution in the language from aspirational terms (“chosen” or “desired”) to the language of actual contribution (“achieved”). If adjudicators consistently apply this standard in future, it would produce coherence between the W&B and AM tests. That said, the continued use of the term “level of protection” is undesirable. The “equivalent contribution” approach makes the link to the W&B test much clearer.

Contribution must include a “magnitude” (even if the analysis is qualitative)

As Bown and Trachtman suggest, the contribution analysis must include a “magnitude” to serve as the benchmark for the “equivalent contribution” analysis. This does not mean that the contribution must be expressed in quantitative or numerical terms. The “considerably reduce” magnitude from *Korea Beef* is expressed in qualitative terms, but remains capable of serving as a benchmark. By

⁶³ Appellate Body Report, *Colombia – Measures Relating to the Importation of Textiles, Apparel and Footwear*, WT/DS461/AB/R and Add.1, adopted 22 June 2016 (*Colombia Textiles*) para 5.74.

contrast, the “reduction of exposure to the targeted risks” and “reducing unnecessary generation of tyre waste” benchmarks are too vague and value-laden.

Adjudicators must apply the characterisation of the objective consistently across all legal elements

While I do not seek to establish a protocol around how the policy objective is defined, it is imperative that adjudicators characterise the objective consistently across all relevant legal elements. If a broader objective like human health is used for the importance analysis (thus favouring the regulator), it must also be available under the AM test (thus favouring the complainant). This would preserve a fair balance between the parties.⁶⁴

Adjudicators must clarify the object of review

For both elements where a broad characterisation is regulator-friendly (threshold examination and importance), jurisprudence remains unclear about the object of review. Does the design step test the measure or the policy objective? Does importance test the policy objective or the “societal value” (or even the “public interest”)? Adjudicators must clarify the object of review. In both cases, it should be the policy objective.

7.5 Brazil Tyres Summary

In *Brazil Tyres*, adjudicators relied on an anomaly to conceal their underlying value judgments. The characterisation of the policy objective, which should have been a

⁶⁴ That said, the appropriate protocol for characterising the objective is an important issue which merits further analysis (and could build upon Bown and Trachtman’s approach). It is clear that some characterisations would be too broad or too narrow to permit an appropriate analysis of all the relevant legal elements.

“fixed” factual matter, varied across five different legal elements.⁶⁵ The policy objective went from extremely broad (“protecting human beings from health risks”) in the importance analysis to extremely narrow (“reducing unnecessary generation of tyre waste”) under the AM test. A “rationalist” approach to legal reasoning would require the policy objective to remain fixed regardless of which legal element is being assessed.

I further sought to demonstrate that the adjudicators’ non-rational approach may reveal a “systematic bias” in favour of Brazil’s measure. It also enabled adjudicators to appear value-neutral even though their decision was presumably based on some compelling value judgments about the importance of reducing waste tyres.

If we recognise the fundamental “anomaly” which underpins the analysis, it becomes clear that the adjudicators’ decision cannot be justified rationally.⁶⁶ Adjudicators rely on normative value judgments, but they manage to disguise this fact through the use of flawed “reasoning”. This non-rationalist approach is antithetical to law and leaves the real reasons for the decision unstated. However, based on the current mandate of WTO adjudicators, it is hard to see what other options they have.

7.6 Reflections: The Necessity Test and Regulatory Autonomy

This analysis seeks to highlight the difficulty of clearly identifying how much regulatory autonomy Members retained under GATT/WTO and how much they ceded. For regulatory domestic measures, adjudicators have decided to address

⁶⁵ The “degree of contribution” analysis suffers from similar problems, though not quite as blatantly.

⁶⁶ In this sense, it cannot be justified by legal formalism.

this question under Article XX.⁶⁷ Oftentimes, it is addressed most substantively under the necessity test.⁶⁸

For GATT panels, and early WTO disputes, there was a single approach to necessity (the AM test) which seemingly sought out Pareto improvements to a regulator's measure. In *Korea Beef*, adjudicators introduced the W&B test which resembled (at least on paper) a more intrusive search for Kaldor-Hicks efficiency gains.

The W&B test was highly controversial as it seemed to introduce "balancing" into WTO law, including trade-offs which could make regulators worse off where this was justified by global efficiency gains. Scholars were particularly scathing of the normative elements of the W&B test (importance and balancing), though some have taken solace from the fact that adjudicators do not truly operationalise these legal elements in practice.

The *Korea Beef* case study demonstrated that the AM test is not as value-neutral as it may seem due to the use of disguised "balancing". Adjudicators can reduce the regulator's utility by lowering the level of protection or by accepting a "reasonably available" alternative which imposes significant financial and human resource costs.

In *Korea Beef*, adjudicators used both techniques. While the dispute outcome was not particularly controversial,⁶⁹ the reasoning under both "steps" of the necessity test has been extremely controversial. The Appellate Body's assertion of a right to engage in balancing, via the W&B test, was widely-criticised.⁷⁰ The disguised use of

⁶⁷ Despite fierce opposition from many scholars who suggest it should be addressed within Article III.

⁶⁸ Though it may also arise under the chapeau, a matter which falls outside the scope of this thesis.

⁶⁹ Despite the widespread discussion (and criticism) of *Korea Beef*, there have been no major critiques suggesting that the ultimate finding was wrong or that the dual retail system was legitimate under WTO law.

⁷⁰ Even though the Appellate Body did not actually apply the W&B test in practice.

balancing in the AM test (to reduce utility for the regulator) has also been criticised, but only by a handful of scholars.⁷¹

Despite the controversies about the W&B test and disguised balancing, Weiler notes that “Members have largely reacted with indifference” and “simply do not seem to care.”⁷² Members appear more concerned by pragmatic considerations (the right result) than legal niceties. Perhaps, at its core, WTO dispute settlement remains essentially “diplomatic” even if it has become increasingly encased in sophisticated legal trappings.

Weiler cautions, however, that “in law, a right result might be a necessary but not a sufficient condition”.⁷³ Adjudicators should not be content with producing adequate results on a case-by-case basis. They must also develop a coherent and consistent jurisprudence which underpins their decisions.

The *Brazil Tyres* case study considered how adjudicators make value judgments that a regulator’s health/environment objective can justify a trade restriction. The report suggests that Brazil’s import ban was justifiable, however the use of variable policy objectives implies that the necessity finding was not based on a “rationalist” approach to legal interpretation.

I sought to demonstrate that the fluctuating characterisation of the objective revealed a systematic bias in favour of Brazil. For each legal element, adjudicators chose a broad or narrow version of the objective in a way which favoured Brazil’s arguments. To avoid this bias, adjudicators should have used a single characterisation of the objective and let Brazil’s defence succeed or fail on its merits.

⁷¹ Such as Weiler and Ming Du (whose critiques were described in chapter 6).

⁷² Weiler (n 2) 141.

⁷³ ibid 139.

There is an irony in the evolution of the necessity test. By *Brazil Tyres*, the necessity test had theoretically become more intrusive. GATT-era regulators only had to pass the (Pareto) AM test, but modern regulators face a second hurdle under the (Kaldor-Hicks) W&B test. This seems to allow modern adjudicators to balance conflicting interests in a way which makes regulators worse off. Such “balancing” was not available to their GATT-era predecessors.

However, the culture of adjudication has seemingly become increasingly deferential towards regulators as the WTO moves away from the “pro-trade bias” of GATT panels.⁷⁴ If *Brazil Tyres* had arisen during the GATT era, would Brazil have lost (due to the culture of adjudication at the time)⁷⁵ even though the legal test was theoretically more lenient? Theoretically, the necessity test has become more intrusive over time, but perhaps Members are satisfied that the “reduced trade bias” has led to regulator-friendly (and deferential) outcomes.⁷⁶

In my view, the controversy about the necessity test is a natural by-product of an underlying tension which remains unresolved at the WTO: what is the balance between disciplining protectionism and deferring to regulatory autonomy? *Korea Beef* revealed that a narrow AM test (based on Pareto efficiency) might be insufficient to curb some measures which are widely viewed as protectionist.

⁷⁴ Don Regan, ‘International Adjudication: A Response to Paulus – Courts, Custom, Treaties, Regimes, and the WTO’ in Samantha Besson and John Tasioulas (eds), *The Philosophy of International Law* (Oxford University Press New York 2010) 238-9.

⁷⁵ Regan’s views about the culture of GATT adjudication is supported by Pirker who observes that GATT panels suffered from a “strong bias in favour of trade” and Kapterian who notes that seemingly legitimate non-trade measures often failed on the basis of proposed alternatives which were only “hypothetically available”; Benedikt Pirker, *Proportionality Analysis and Models of Judicial Review: A Theoretical and Comparative Study* (Europa Law Publishing 2013) 303-304, Gisele Kapterian, ‘A Critique of the WTO Jurisprudence on ‘Necessity’’ (2010) 59 International & Comparative Law Quarterly 89, 103.

⁷⁶ Regan (n 74), p238-9.

However, Members refuse to recognise the right of adjudicators to engage in balancing and trade-offs. In *Brazil Tyres*, the outcome was regulator-friendly and seemed to rely on a value-neutral application of the necessity test. However, adjudicators used non-rationalist techniques which strongly suggest that their value-neutral approach was a mere façade.

There is still much resistance to the notion that adjudicators should engage in balancing or make normative value judgments about non-trade policy objectives. Nevertheless, scholars recognise that adjudicators do make such value judgments (even if they are not supposed to) and speculate about where and how those value judgments are made. I have sought to contribute to this discussion by mapping some of the techniques adjudicators use to make value judgments, both in support of regulatory autonomy (*Brazil Tyres*) or when striking down protectionist measures (*Korea Beef*).

After a quarter of a century of WTO law (following decades of GATT panels), the necessity test remains vague and poorly-defined: it is too malleable to constrain the discretion of adjudicators or produce determinate (predictable) results. Ming Du believes the Appellate Body has reserved for itself “maximum adjudicatory flexibility”.⁷⁷ Adjudicators appear to be influenced by their “own substantive views”⁷⁸ and their “own value system”.⁷⁹ McGrady implies that adjudicators might direct their analysis “to achieve a desired outcome”.⁸⁰

⁷⁷ Michael Ming Du, ‘Autonomy in Setting Appropriate Level of Protection under the WTO Law: Rhetoric or Reality?’ (2010) 13 Journal of International Economic Law 1077, 1096.

⁷⁸ Don Regan, ‘The Meaning of ‘Necessary’ in GATT Article XX and GATS Article XIV: The Myth of Cost-Benefit Balancing’ (2007) 6 World Trade Review 347, 352.

⁷⁹ Kapterian (n 75) 91.

⁸⁰ McGrady (n 13) 159.

This situation renders the Appellate Body’s value judgments “opaque”.⁸¹ How did adjudicators determine that the dual retail system was protectionist but that the used tyre import ban was legitimate? The true reasons for these decisions remain a mystery to Members and scholars.

The Appellate Body is in a difficult position. There is strident criticism any time it appears to intrude on regulatory autonomy. However, its approach of unspoken (yet clearly present) value judgments, especially under the AM test, detracts from the WTO’s credentials as a system based on the rule of law.

How can the optimal solution be for adjudicators to make determinative value judgments surreptitiously and then violate basic principles of legal reasoning to keep their true reasons hidden? Fontanelli suggests that maybe the WTO should “start embracing, very cautiously, a bit of proportionality *proprement dite*”.⁸² His proposal may well represent the lesser of two evils.

In the remainder of this thesis, I will undertake a detailed analysis of the *Seals* dispute. As in *Brazil Tyres*, I will identify an “anomaly” regarding the characterisation of the objective, however the *Seals* double standard is more extreme as there are two substantively different policy objectives (rather than different points on a spectrum).

Further, the *Seals* double standard is harder to spot as adjudicators rely on equivocal language. They use the term “seal welfare” to describe the policy objective, but this term has different meanings in different contexts. Notwithstanding these difficulties, I will develop techniques to render the “anomaly” visible.

⁸¹ Kapterian (n 75) 91.

⁸² Fontanelli (n 16) 68.

Finally, in one further respect, my analysis of *Seals* will be more ambitious. In *Brazil Tyres* (and *Korea Beef*), I refrained from trying to explain the real reasons for the adjudicators' decision. Regan claims that adjudicators' value judgments are made "unconsciously" or "*sub rosa*"⁸³ which suggests we would require psychological expertise to access them. In *Seals*, I will seek to enter the psyche of adjudicators to see if I can explain what unconscious processes motivated their systematic bias in favour of "seal welfare".

⁸³ Regan (n 78) 365.

CHAPTER 8 – THE PUBLIC MORALS JURISPRUDENCE

8.1 Introduction

This chapter considers which policy objectives fall within the ambit of “public morals” and can therefore be used to justify trade-restrictive measures under Article XX(a). It will describe the legal test and its application using different approaches including unilateralism, originalism, transnationalism and internationalism.

It will further address several overarching questions which arise in the academic debate: (i) whether public morals are different from the other public interest exceptions; (ii) if so, what makes them different; and (iii) whether their uniqueness calls for greater deference towards moral regulation or stricter safeguards against abuse.

This chapter will show that public morals disputes have generally been adjudicated with the same techniques used for other public interests. Even the *Seals* dispute, which seemed to raise sensitive moral and cultural issues, was ultimately adjudicated using the usual methods (with some minor changes).

Section 8.2 will consider whether public morals are truly different and, if so, what makes them unique. Section 8.3 will summarise the jurisprudence and scholarly debate about which moral objectives are protected by Article XX(a). Section 8.4 summarises some key observations from the jurisprudence prior to *Seals*. Sections 8.5 and 8.6 look at how the jurisprudence evolved in *Seals*, including the use of evidence from moral “experts” (philosophers) and some minor changes to the formulation of the legal test.

8.2 Are Public Morals Different?

The public morals exception is “hotly contested”¹ and many scholars view it as different from the other Article XX exceptions. My analysis will focus on the moral exception in the context of the “design step” where adjudicators determine if a measure has “a relationship to” the protection of public morals.² In *Seals*, adjudicators addressed this question under the TBT Agreement where they found that “seal welfare” was a “legitimate objective”.³ The panel considered this finding to be “equally applicable” to Article XX(a), a view endorsed by the Appellate Body.⁴

The public morals debate inspires two conflicting fears. Wu and Diebold argue that the vagueness of the exception will encourage protectionism and abuse.⁵ Marwell promotes the opposite view that WTO law might overly restrict policy space for culturally-specific moral objectives.⁶

The first fear demands prescriptive rules, which would allow WTO adjudicators to clamp down on abuse. The second fear requires greater deference to regulators to

¹ Tamara Nachmani, ‘To Each His Own: The Case for Unilateral Determination of Public Morality under Article XX(a) of the GATT’ (2013) 71 University of Toronto Faculty of Law Review 31, 34.

² Appellate Body Report, *Colombia – Measures Relating to the Importation of Textiles, Apparel and Footwear*, WT/DS461/AB/R and Add.1, adopted 22 June 2016 (*Colombia Textiles*) paras 5.67-5.68. I will not focus on the necessity test in this chapter, though I note that many issues discussed in this chapter would be equally relevant to the “importance” analysis under the W&B test.

³ The full analysis appears in the Panel Reports, *European Communities – Measures Prohibiting the Importation and Marketing of Seal Products*, WT/DS400/R and Add.1 / WT/DS401/R and Add.1, adopted 18 June 2014 (*Seals*) paras 7.372-7.411.

⁴ Panel, ibid, para 7.382; Appellate Body Reports, *European Communities – Measures Prohibiting the Importation and Marketing of Seal Products*, WT/DS400/AB/R / WT/DS401/AB/R, adopted 18 June 2014 (*Seals*) paras 5.134-5.140.

⁵ See eg Mark Wu, ‘Free Trade and the Protection of Public Morals: An Analysis of the Newly Emerging Public Moral Clause Doctrine’ (2008) 33 The Yale Journal of International Law 215; and Nicolas Diebold, ‘The Morals and Order Exceptions in WTO Law: Balancing the Toothless Tiger and the Undermining Mole’ (2007) 11(1) Journal of International Economic Law 43.

⁶ See eg Jeremy Marwell, ‘Trade and Morality: the WTO Public Morals Exception after Gambling’ (2006) 81 New York University Law Review 816.

make judgments about the moral values within their society. Adjudicators must find the “precarious balance” between these views.⁷

Nothing in the text of Article XX suggests that public morals are different from the other exceptions or require special rules.⁸ However, many leading commentators subscribe to the view that Article XX(a) is different and requires a unique approach.⁹ What reasons do scholars offer for the special status of public morals?

Public morals are ambiguous and dynamic

Scholars have argued that the concept of public morals is more “amorphous”¹⁰, “vague”¹¹ and “ambiguous”¹² than other public interests. Diebold suggests that public morals lack a “fixed scope and content” and may therefore evolve over time.¹³

Is this claim true and can it justify treating public morals differently?

In reality, public morals are not alone (among Article XX exceptions) in being ambiguous or dynamic. In *Shrimp*, adjudicators had to deal with the thorny question of whether the term “exhaustible natural resources” allowed for the protection of living resources (such as endangered sea turtles) or whether it was limited to mineral resources.¹⁴

⁷ Wu (n 5) 215. This replicates the broader tension in WTO law between retained and ceded regulatory autonomy.

⁸ In fact, public morals rely on the same “necessary to protect” standard as Articles XX(b) and XX(d).

⁹ For example, Marwell (n 5, 823) and Feddersen suggest that public morals should not apply where another Article XX general exception is available; Christoph Feddersen, ‘Focusing on Substantive Law in International Economic Relations: The Public Morals of GATT’s Article XX(a) and “Conventional” Rules of Interpretation’, (1998) 7 Minnesota Journal of Global Trade 75, 107.

¹⁰ Steve Charnovitz, ‘The Moral Exception in Trade Policy’ (1998) 38(4) Virginia Journal of International Law 689.

¹¹ Wu, above n 5, at 219.

¹² Diebold (n 5) 46.

¹³ ibid 53.

¹⁴ Appellate Body Report, *United States – Import Prohibition of Certain Shrimp and Shrimp Products*, WT/DS58/AB/R, adopted 6 November 1998 (*Shrimp*).

The Appellate Body determined that Article XX(g) "must be read by a treaty interpreter in the light of contemporary concerns of the community of nations".¹⁵ This established the idea that the substantive meaning of exceptions could evolve based on dynamic interpretation.¹⁶

In the first public morals dispute, adjudicators appeared to endorse flexibility and dynamic interpretation. In response to "sensitivities" about policy space for moral matters, adjudicators noted that "the content of these concepts for Members can vary in time and space".¹⁷

Even if the term "public morals" is vague, and the concept is dynamic, this would not necessarily justify special treatment under Article XX. These features are shared by other public interests.

Public morals are likely to be abused

Building on the notion that public morals are ambiguous, several commentators suggest that Article XX(a) is more likely to be abused by regulators. Wu proposes special (and stricter) rules for public morals as a means of "preventing states from abusing that power to enact protectionist measures in disguise".¹⁸ Diebold opposes "individual interpretation by Members" (unilateralism) as this would encourage "excessive application or even abuse".¹⁹ Gonzalez states that the lack of a "fixed

¹⁵ Ibid, para 129.

¹⁶ Article XX(d) is a further example of a vague exception with an extremely broad coverage as it excepts any measures "necessary to secure compliance with laws or regulations which are not inconsistent with the provisions of this Agreement".

¹⁷ Panel Report, *United States – Measures Affecting the Cross-Border Supply of Gambling and Betting Services*, WT/DS285/R, adopted 20 April 2005 (US – Gambling) para 6.461.

¹⁸ Wu (n 5) 216.

¹⁹ Diebold (n 5) 51.

meaning" of "public morals" renders it "subject to abuse and may result in the frustration of the objectives of the WTO".²⁰

These scholars rely on the twin notions that (i) public morals are vaguer than other exceptions and (ii) vagueness facilitates abuse. I have already argued that Article XX(a) is not the only vague exception. Even if it was, does vagueness necessarily create a higher risk of abuse?

Adjudicators are conscious of their role in identifying protectionism and preventing abuse. In *Gasoline*, with respect to "exhaustible natural resources", the Appellate Body noted that the concept "may not be read so expansively as seriously to subvert the purpose and object of Article III:4".²¹

Despite this note of caution, *Shrimp* adjudicators subsequently gave this term an expansive definition which allowed for the protection of living natural resources, not just mineral resources. In *Shrimp*, the textual ambiguity of Article XX(g) led adjudicators to increase policy space for regulators.

Nuzzo argues that adjudicators should (and do) take a similarly open-minded approach to the ambiguity of public morals:

"WTO adjudicators undoubtedly aimed to grant a high degree of deference to national authorities in such a sensitive and uncertain matter like 'public morals'."²²

Proponents of the view that vagueness leads to abuse have yet to produce compelling empirical evidence. Despite the widespread use of public moral

²⁰ Miguel Gonzalez, 'Trade and Morality: Preserving "Public Morals" Without Sacrificing the Global Economy' (2006) 39 Vanderbilt Journal of International Law 939, 970.

²¹ Appellate Body Report, *United States – Standards for Reformulated and Conventional Gasoline*, WT/DS2/AB/R, adopted 20 May 1996 (*Gasoline*) 18.

²² Silvia Nuzzo, 'Tackling Diversity inside WTO: GATT Moral Clause after Colombia-Textiles' (2017) 10 European Journal of Legal Studies 267, 275.

measures, the first fifty years of GATT/WTO did not see a single public morals measure challenged in dispute settlement.²³ Further, in WTO disputes, the design step has rarely been controversial. Nuzzo notes:

“In the four disputes which have occurred hitherto, the objective pursued by the measures at stake was always recognized as a matter of ‘public morals’.”²⁴

Notwithstanding the legitimacy of the objective, the regulator’s defence ultimately failed in these four disputes.²⁵ This shows that just because an *objective* is legitimate, this does not guarantee that the *measure* will be found WTO-consistent. Indeed, even for non-moral exceptions, regulators have sometimes failed to defend measures despite the fact the objective was clearly legitimate such as endangered species protection,²⁶ air quality control²⁷ and tobacco control.²⁸

In these failed defences, adjudicators occasionally highlight explicitly that the legitimacy of the objective was not the problem. In *Shrimp*, the Appellate Body stated:

“We have not decided that the sovereign nations that are Members of the WTO cannot adopt effective measures to protect endangered species, such as sea turtles. Clearly, they can and should”.²⁹

The link between vagueness and abuse is not automatic. If a government wishes to protect its domestic industry, it is not clear why a supposedly vague public morals exception is a better place to hide protection than a health or environmental measure. There is no compelling evidence that protectionist governments seek to

²³ For a mapping of the types of public moral measures in force (based on WTO Trade Policy Reviews), see Marwell (n 5) 818.

²⁴ Nuzzo (n 22) 278. Nuzzo’s analysis only focusses on the (four) disputes which had occurred at that time.

²⁵ Nuzzo (n 22) 278. Nachmani reaches a similar conclusion.

²⁶ Appellate Body Report, *Shrimp* (n 14).

²⁷ Appellate Body Report, *Gasoline* (n 21).

²⁸ Appellate Body Report, *United States – Measures Affecting the Production and Sale of Clove Cigarettes*, WT/DS406/AB/R, adopted 24 April 2012 (*Clove Cigarettes*). Strictly speaking, *Clove Cigarettes* was adjudicated under the TBT Agreement using its equivalent to an Article XX defence.

²⁹ Appellate Body Report, *Shrimp* (n 14) para 185.

target and exploit the public morals exception. The “vagueness” and “abuse” explanations for singling out the public morals exception as a special case appear empirically and conceptually questionable.

Public morals are subjective and culturally-specific

Some authors argue that public morals are different because they are culturally-specific. Diebold suggests that if WTO law fails to protect “subjective values”, this would render the public morals exception “a toothless tiger”.³⁰ Marwell highlights that moral norms are “strongly held, geographically localized, and diverse”.³¹ He further argues:

“The most significant contrast between public morals and natural resources or health is the existence or absence of internationally accepted objective evidence as to the nature of the exception itself”.³²

In *Seals*, the *amici*³³ distinguish between “instrumental regulation” which aims at “a certain empirical consequence” and “non-instrumental regulation” which “expresses moral convictions”.³⁴ This distinction is useful as it suggests that some Article XX(a) defences are analogous to health/environmental objectives, but that there is also a subset of “non-instrumental” moral objectives which may raise unique issues.

These scholars argue that (at least some) public morals defences carry a greater risk of regulatory intrusion by adjudicators, as Members will not necessarily be able to

³⁰ Diebold (n 5) 54.

³¹ Marwell (n 6) 805.

³² *ibid* 815.

³³ Howse et al are the leading commentators on *Seals* (see footnotes 34 and 68 below). I refer to them in the text as the “*amici*” since they submitted their views as amicus briefs at both the panel and appellate stages and referred to themselves as the “Non-Party *Amici Curiae*”.

³⁴ Robert Howse, Joanna Langille and Katie Sykes, ‘Pluralism in Practice: Moral Legislation and the Law of the WTO after Seal Products’ (2015) 48 George Washington International Law Review 81, 83.

produce arguments based on instrumental reasoning, objective evidence, science or widely-accepted views on what is legitimate.³⁵

It is true that culturally-specific moral objectives may require a different approach than “instrumental” exceptions. The legitimacy of a health/environmental goal seems to be an objective matter which would not vary from one Member to another (although the perceptions of risk or pursued level of protection may still differ). For Article XX(a), however, it seems as though the ambit of the exception could potentially vary from one Member to another based on their specific moral norms? This would give rise to a thorny issue under Article XX(a) which does not arise for the other exceptions.

Nevertheless, it is important to note (as the *amici* did) that not all moral norms are culturally-specific. Many Article XX(a) defences may relate to “instrumental” or widely-recognised moral matters.³⁶ Indeed, in the early disputes, WTO adjudicators were able to rely on their usual reasoning techniques as the moral norm was instrumental in nature.³⁷

It is only for culturally-specific norms that the WTO might have to deviate from the instrumental approach it uses for health/environmental defences. *Seals* seemed to be the first dispute involving a culturally-specific norm which might require adjudicators to find non-instrumental reasoning techniques. However, my analysis in section 8.6 will show that it did not lead to special rules or significant evolution in the jurisprudence. How did adjudicators resolve this dispute while relying on their usual reasoning techniques?

³⁵ The different types of evidence will be discussed in detail in section 8.3.

³⁶ Howse et al (n 34) 83.

³⁷ For example, the effects of gambling on criminal activity and mental health.

Summary

In my view, the treatment of public morals as a special case has been exaggerated. Certain claims made about public morals (that they are vague, dynamic and capable of abuse) are also true of other general exceptions. Further, it is contested whether vagueness leads to a higher risk of protectionism or invariably requires narrow interpretative approaches. While it is true that adjudicators might face unique issues for non-instrumental and culturally-specific moral objectives, such cases have been rare in practice.

8.3 WTO Jurisprudence

This section will consider different jurisprudential and academic approaches to the perennial question: “whose moral values” are protected by Article XX(a)?³⁸ I divide the different approaches into two families: those which allow the regulator to self-determine its moral values (“unilateralism”) and those which rely on “external standards”.

Unilateral	External Standards
Evidentiary Unilateralism	Originalism Transnationalism/Universalism Internationalism Expert Evidence ³⁹

The unilateral approach is preferred by scholars seeking greater deference for regulators,⁴⁰ but would require a deviation between public morals and the health/environmental jurisprudence. The “external standards” approach is preferred

³⁸ This question has motivated scholars since Charnovitz (n 10, 4) first posed the question in his seminal work.

³⁹ Expert evidence will be analysed separately in section 8.5. *Seals* was the first public morals dispute to extensively reference expert evidence.

⁴⁰ Such as Marwell (n 6) and Nachmani (n 1).

by scholars worried about the abuse of unilateralism for protectionist purposes.⁴¹ In practice, adjudicators seem to rely on both unilateralism and external standards (where possible), even though these approaches would be difficult to reconcile for culturally-specific norms.

Some scholars suggest that different approaches may amount to decisive legal tests,⁴² though I do not believe the jurisprudence is sufficiently developed to support such a view. My analysis conceives of these approaches (especially the external standards) as different types of evidence which can be furnished by regulators to strengthen their defence. It has become common practice for regulators to offer, and for adjudicators to cite, evidence from each of these approaches (where available).

8.3.1 Unilateralism and the Legal Test

8.3.1.1 Academic Views

Unilateralism was first proposed by Charnovitz when he suggested that “a unilateral determination of morality may be appropriate for inwardly-directed concerns”.⁴³ This approach would allow regulators to self-declare a norm without needing to rely on external standards to demonstrate that it is a public moral.

Marwell further developed unilateralism by proposing an approach which allows Members “to define public morals unilaterally” but also “require[s] evidence from that

⁴¹ Such as Wu (n 5).

⁴² Such as Marwell’s endorsement of “evidentiary unilateralism” or Gonzalez’s endorsement of “originalism”. These approaches will be described in detail below.

⁴³ Charnovitz (n 10) 26. However, Charnovitz considered that unilateralism would be “too open-ended for outwardly-directed concerns”. Some later authors subscribed to his view that stricter rules were required for outwardly-directed or extra-territorial measures (including Wu and Diebold), though Nachmani questions whether the distinction between inwardly- and outwardly-directed measures is useful or practical to implement (n 1, 52-53). While extra-territorial measures may require dedicated jurisprudence, I choose not to address this as a public morals issue since I believe that extra-territoriality is a systemic issue which also affects other Article XX exceptions, such as environmental measures.

country supporting its claim that a particular issue has moral significance".⁴⁴

Marwell's major innovation was to require the regulator to provide:

"substantial evidence of its internal conditions e.g., historical practice, contemporary public opinion polls, results of political referenda, or statements of accredited religious leaders".⁴⁵

Marwell argued this would constrain abuse while still allowing regulators the "leeway to define public morals based solely on domestic circumstances".⁴⁶ His approach treats the key question as empirical ("whether the stated interest actually exists") rather than normative ("whether a particular interest is vital enough to fall under the GATT and GATS public morals exception").⁴⁷ To address this empirical question, adjudicators could use "unilateralist" evidence from within the regulator's own society.

Nachmani endorses evidentiary unilateralism and adds some procedural requirements. Based on the work of Wu and Smith⁴⁸, she argues that regulators should also be required to show that their measure was passed through a legislative procedure⁴⁹ and that there have been good faith efforts to enforce it domestically.⁵⁰ Unilateralism is further endorsed by Nuzzo, both as a description of what adjudicators actually do and as her preferred jurisprudential approach.⁵¹

⁴⁴ Marwell (n 5) 824.

⁴⁵ ibid 824-5.

⁴⁶ ibid 806.

⁴⁷ ibid 825.

⁴⁸ Tyler Smith, 'Much Needed Reform in the Realm of Public Morals: A Proposed Addition to the GATT Article XX(a) 'Public Morals' Framework, Resulting from China-Audiovisual' (2011) 19:3 Cardozo Journal of International and Comparative Law 733.

⁴⁹ This is based on Wu's argument that a legislative requirement adds a further safeguard against protectionism.

⁵⁰ Nachmani (n 1) 57-58. This is based on Smith's view that a ban should not be considered legitimate where regulators tolerate a black market for the banned product.

⁵¹ Nachmani also describes Howse and Diebold as unilateralists (ibid 47), though I am not convinced this description truly reflects their positions so I exclude them from the unilateralist community.

8.3.1.2 Jurisprudence

The first public morals dispute (under GATS Article XIV(a)) defined the term as "standards of right and wrong conduct maintained by or on behalf of a community or nation".⁵² The panel stated that the concept of public morals "can vary in time and space, depending upon a range of factors, including prevailing social, cultural, ethical and religious values."⁵³

The panel proposed a deferential approach:

"Members should be given some scope to define and apply for themselves the concepts of "public morals" and "public order" in their respective territories, according to their own systems and scales of values".⁵⁴

Nachmani suggests that these statements of law support the unilateralist view. After all, adjudicators refer to Members applying "their own systems and scales of values". However, it is important to note that adjudicators only offered "some scope" to Members, not "unlimited scope". Presumably, there are limits to pure unilateralism.

Marwell believes adjudicators rejected evidentiary unilateralism:

"Taken to an extreme, the *Gambling* doctrine might be read as implying that states cannot unilaterally define public morals".⁵⁵

While adjudicators have often made statements describing Article XX(a) as unilateralist, in practice, they tend to rely on a range of external standards in ways which would be inconsistent with pure unilateralism.

8.3.2 Evidentiary Practice and External Standards

This section will describe how adjudicators use external standards to determine which types of evidence can support public moral claims. I will show that the

⁵² Panel Report, *Gambling* (n 17) para 6.465.

⁵³ ibid 6.461.

⁵⁴ ibid. The GATS public morals exception also includes a reference to "public order" which does not appear in Article XX(a).

⁵⁵ Marwell (n 5) 817.

practice for public morals closely resembles the approach for health/environment objectives.

8.3.2.1 Originalism

The originalism approach stems from Charnovitz's seminal work on the GATT drafters' motivations for including a public morals exception.⁵⁶ In the negotiating history, he found a single reference to a moral objective: "temperance" (alcohol control).⁵⁷ He complemented this by identifying "the contemporary trade controls that could have triggered the legal need for an exception".⁵⁸ His list included "opium, pornography, liquor, slaves, firearms, blasphemous articles, products linked to animal cruelty, prize fight films, and abortion-inducing drugs".⁵⁹

Originalism has found little favour as a decisive legal test,⁶⁰ though it has been used as a persuasive argument for Article XX defences. In *Shrimp*, under Article XX(g), the complainants "referred to the drafting history of Article XX(g), and, in particular, to the mention of minerals, such as manganese" in order to narrow the definition of "exhaustible natural resources" to exclude living resources.⁶¹ The Appellate Body rejected the view that originalism arguments should be treated as an "exhaustive" list which limits the scope of an exception.

However, this does not prevent the use of originalism to make positive arguments in the regulator's favour. Wu suggests that originalist objectives should automatically

⁵⁶ I add my voice to the chorus of authors acknowledging Charnovitz's insightful analysis of public morals years before the issue was addressed by a WTO panel.

⁵⁷ Charnovitz (n 10) 8.

⁵⁸ ibid 13.

⁵⁹ ibid.

⁶⁰ With the exception of Gonzalez who argues that "the public morals exception should be limited to the traditional range of public morals issues" (n 20, 972). I should note that Charnovitz never claimed that his approach was intended to offer a decisive test.

⁶¹ Appellate Body Report, *Shrimp* (n 14) para 127.

be considered legitimate.⁶² If a Member sought to use negotiating history to demonstrate that manganese is an “exhaustible natural resource” or that alcohol regulation is a public morals issue, this would appear to be a compelling argument.

8.3.2.2 Universalism and Transnationalism

Universalism is the notion that trade-restrictive measures can only be justified by moral norms which reflect “universal or near-universal practice amongst other WTO member states”.⁶³ It has roused significant academic opposition. Wu argued:

“If one were to require that morals be near-universal before being considered a “public moral”, then the set of morals that would actually qualify might be so limited as to render the exceptions clause effectively useless”.⁶⁴

Transnationalism also looks to other Members, but does not require universal practice. Wu argues that adjudicators should permit moral norms which are “shared widely by a group of *similarly situated countries*”.⁶⁵ For example, a Muslim country could justify an alcohol ban by showing that a moral norm against alcohol consumption is “widespread among Islamic societies” even if it is not “universally shared”.⁶⁶

While it is less-demanding than universalism, transnationalism remains contentious. Marwell suggested that regulators should not be required to “present evidence of

⁶² Wu (n 5) 243. He argues that adjudicators should endorse objectives where “the measure at stake concerns a classification of morals as originally understood by the GATT’s drafters”. Therefore, originalism can be decisive in favour of the regulator, but absence from the “originalism” list cannot be decisive against the regulator. Wu’s suggestion is generally compelling but may also be slightly exaggerated: if adjudicators endorse dynamic interpretation, why would it not be possible for certain “original” issues to cease being legitimate over time?

⁶³ Marwell (n 6) 820.

⁶⁴ Wu (n 5) 232.

⁶⁵ ibid 240.

⁶⁶ ibid 232.

similar practice by other states".⁶⁷ The *amici* oppose any suggestion that a moral view must be "widely held".⁶⁸

Transnationalism has been used for other public interests. In *Shrimp*, the Appellate Body relied on transnationalism when it determined that:

"the policy of protecting and conserving the endangered sea turtles here involved is shared by all participants and third participants in this appeal, indeed, by the vast majority of the nations of the world".⁶⁹

8.3.2.3 Internationalism

Regulators often cite international instruments to support their view that their policy objective is legitimate. Wu argues that where a moral objective "falls within a *jus cogens* norm", this evidence should be dispositive in favour of the regulator.⁷⁰ The *amici* agree that international law is relevant, but use it to dismiss non-compliant measures:

"it may be impermissible to justify under Article XX(a) an otherwise WTO-inconsistent measure that is also a violation of *jus cogens* norms".⁷¹

Nachmani questions the "the amorphous and undefined nature of *jus cogens* norms" as a useful standard.⁷² In any case, much international law fails to qualify as a *jus cogens* norm and this is where harder cases are likely to arise.

In the context of *Seals*, Sykes suggests any "general principle of international law" should be considered relevant, even if only as persuasive evidence.⁷³ She describes animal welfare as a "meta-principle" under international law which "should be taken

⁶⁷ Marwell (n 6) 817.

⁶⁸ Robert Howse and Joanna Langille, 'Permitting Pluralism: The Seal Products Dispute and Why the WTO Should Accept Trade Restrictions Justified by Noninstrumental Moral Values' (2012) 37(2) Yale Journal of International Law 367, 373.

⁶⁹ Appellate Body Report, *Shrimp* (n 14) para 135. This reference arguably meets the universalism standard.

⁷⁰ Wu (n 5) 243

⁷¹ Howse et al (n 34) 100.

⁷² Nachmani (n 1) 55.

⁷³ Katie Sykes, 'Sealing Animal Welfare into the GATT Exceptions: the International Dimension of Animal Welfare in WTO disputes' (2014) 13(3) World Trade Review 471, 472.

into account in treaty interpretation, including in the WTO context".⁷⁴ Sykes cites the UN World Conservation Strategy ("people should treat all creatures decently"),⁷⁵ the draft Universal Declaration on Animal Welfare⁷⁶ and the OIE Guiding Principles on Animal Welfare.⁷⁷

Nuzzo suggests that persuasive consideration of internationalism occurred in *Colombia – Textiles* where the panel was influenced by the respondent's "international commitments related to the moral ground invoked".⁷⁸

For other general exceptions, adjudicators have relied on internationalism in different ways. In *Shrimp*, the Appellate Body cited a treaty to find that:

"The exhaustibility of sea turtles would in fact have been very difficult to controvert since all of the seven recognized species of sea turtles are today listed in Appendix 1 of the Convention on International Trade in Endangered Species of Wild Fauna and Flora ("CITES")".⁷⁹

The Appellate Body refined this view with further evidence from the Convention on Biological Diversity and UNCLOS.⁸⁰

Adjudicators have also used internationalism to support important factual findings. In *Asbestos*, the Appellate Body determined that

"the carcinogenic nature of chrysotile asbestos fibres has been acknowledged since 1977 by international bodies, such as the International Agency for Research on Cancer and the World Health Organization".⁸¹

In *Brazil Tyres*, the panel relied on an international instrument (the Basel Convention Technical Guidelines on the Identification and Management of Used Tyres) to accept

⁷⁴ ibid 485.

⁷⁵ ibid 484.

⁷⁶ ibid 483.

⁷⁷ ibid 494.

⁷⁸ Nuzzo (n 22) 283.

⁷⁹ Appellate Body Report, *Shrimp* (n 14) para 132.

⁸⁰ ibid para 130.

⁸¹ Appellate Body Report, *European Communities – Measures Affecting Asbestos and Asbestos-Containing Products*, WT/DS135/AB/R, adopted 5 April 2001 (*Asbestos*) para 162.

Brazil's argument that "waste tyre dumps or stockpiles can become the breeding grounds for insects, such as mosquitoes, which are capable of transmitting diseases to humans."⁸² The panel further cited evidence from the WHO about health risks and diseases which can be transmitted by mosquitoes.⁸³

There is strong support for the notion that adjudicators should look to international law as an external standard for moral norms, as well as health/environmental norms. For certain cases, *jus cogens* norms might provide decisive guidance (either for or against the regulator). In complex cases, adjudicators will likely only have access to "persuasive" evidence such as international law "principles".

8.4 Commentary

This section will make four observations about the public morals jurisprudence: (i) while unilateralism and external standards can operate in a complementary manner, pure unilateralism would shun any use of external standards; (ii) the different standards convert what seems to be a normative question into an empirical one; (iii) external standards do not offer determinate answers; and (iv) the public morals approach closely resembles other public interests.

Is the legal test unilateralist?

There is a potential disconnect between the adjudicators' unilateralist statements on public morals and the practice of extensively citing external standards. This has led to disagreement among unilateralists about whether their approach has actually been adopted.

⁸² Panel Report, *Brazil – Measures Affecting Imports of Retreaded Tyres*, WT/DS332/R, adopted 17 December 2007 (*Brazil Tyres*) para 7.61.

⁸³ ibid para 7.110.

Nuzzo, focussing on the legal test, considers that “in the light of the *US - Gambling* doctrine, relevant evidence may stem exclusively from WTO Members' domestic fora.”⁸⁴ She suggests that a measure can be defended based on pure unilateralism with no additional requirement to justify the measure using external standards.

Marwell, focussing on the test's application, considers that external standards (such as transnationalism) are necessary for a successful defence. He suggests that “gambling was found to constitute a legitimate issue of public morality, based primarily on evidence that many countries in addition to the United States held this view”.⁸⁵ Hypothetically, if no other Members were pursuing gambling regulation (and the US only had unilateralist arguments), Marwell implies its defence would have failed.

Adjudicators appear to have a preference for external standards (where available) possibly because this creates consistency with their approach to health/environmental disputes. However, where a public moral is truly culturally-specific (and cannot be supported with transnationalist or internationalist evidence), it remains unclear whether pure unilateralism can suffice to defend the measure.

The standards are empirical/descriptive

In order to appreciate the deferential nature of WTO practice, it is important to recognise that each standard asks a descriptive, rather than a normative question. Adjudicators inquire (empirically) whether a certain community has accepted the existence of a norm, but they do not offer their view on the moral validity of that norm. The descriptive nature of the analysis is illustrated in this table:

⁸⁴ Nuzzo (n 22) 279. Nachmani shares this view.

⁸⁵ Marwell (n 6) 806.

Approach	Community	Descriptive Question
Evidentiary Unilateralism	Respondent Member's society	Does the norm exist in the respondent country?
Originalism	GATT drafters	Did the GATT drafters recognise the norm?
Universalism/ Transnationalism	Governments (acting individually)	Does the practice of other Members recognise the norm?
International Law	Governments (acting in concert)	Is there an international treaty or instrument which recognises the norm?
Expert evidence⁸⁶	Experts	Do relevant experts recognise the norm?

By reframing the normative question in descriptive terms, adjudicators limit their inquiry to whether there is factual evidence to support the empirical claim. In the case of evidentiary unilateralism, Marwell states:

“This determination would turn on the content and credibility of documentary and other evidence as to whether a particular group held the moral belief asserted as the basis for regulation”.⁸⁷

This approach enables adjudicators to outsource the most controversial normative question about moral desirability. This is consistent with calls from scholars for the WTO to avoid normative judgments and getting “entangled in ideological battles”.⁸⁸ The *amici* agree that “the application of Article XX(a) should not entail any assessment of the validity or desirability of the reasons given”.⁸⁹

The WTO’s descriptive approach appears objective and non-obtrusive. If it is designed to protect the WTO’s institutional legitimacy, it seems like a shrewd strategy. However, the external standards lack clear rules about how and when they

⁸⁶ This will be explained in section 8.5 on expert evidence.

⁸⁷ Marwell (n 6) 825.

⁸⁸ Wu (n 5) 237.

⁸⁹ Howse et al (n 34) 104-5.

apply, thus rendering them highly-malleable. Adjudicators retain much discretion to wield these standards as they deem appropriate.

External standards do not provide determinate answers

Even though adjudicators rely on external standards, they do not articulate a precise test. When adjudicators cite evidence, we can deduce that this evidence must be relevant in some way, even if adjudicators do not offer explicit statements about the probative value (or decisiveness) of each type of external standard.⁹⁰ Adjudicators list the ingredients, but they do not provide the recipe.

For proponents of pure unilateralism, any evidence from external standards “should be only ad abundantiam”.⁹¹ However, Nuzzo recognises that, in practice, such evidence seems to have a “highly persuasive value”.⁹²

The vagueness about the probative value of external standards leads to disagreement and confusion about the reason(s) for adjudicators’ judgments. In *Gambling*, Marwell treated transnationalism as decisive and therefore concluded that *Gambling* rejected unilateralism.⁹³

Nachmani, a fellow unilateralist, had a very different view. She agreed that adjudicators referred to transnationalism in *Gambling*, but only to strengthen a conclusion they could have reached based solely on unilateral evidence. This is consistent with Nuzzo’s view that “a common understanding of a moral issue”

⁹⁰ In my view, there was therefore insufficient clarity from adjudicators for Marwell to conclude that transnationalism was decisive in *Gambling*.

⁹¹ Nuzzo (n 22) 281-2.

⁹² ibid. Nuzzo attributes this persuasive value to the choice of parties to put forward such evidence. In my view, adjudicators must also accept significant responsibility as they regularly cite such evidence.

⁹³ Marwell reads the case as “implying that states cannot unilaterally define public morals” (n 6, 817). His view is shared by Wu who states that *Gambling* “appeared to reject both the pure unilateralist and the pure universalist approach” (n 5, 233).

(transnationalism) can support a claim, but that “its absence may not lead by itself to the failure of the justification under Art XX(a) GATT”.⁹⁴

Having a list of external standards provides some shape and structure to the WTO’s public morals test. However, the test lacks clear guidance on which factors are decisive or what combination of persuasive factors is required to show that a moral objective is legitimate.

Responsibility for this uncertainty lays at the feet of adjudicators. While they extensively mention relevant evidence from external standards, they use extremely vague language when connecting that evidence to their conclusion. Adjudicators provide little guidance on the decisiveness or persuasive of different types of evidence.

Some scholars address this gap by speculating that certain external standards can be determinate, at least in some circumstances.⁹⁵ However, even if these claims are true, they would only provide clear answers in the least controversial cases.

Science appears to be an exception where adjudicators have established an explicit standard of review.⁹⁶ Scientific evidence which can meet this standard seems to be afforded a high (perhaps decisive) probative value, though science appears to have only a minor role in the early public morals disputes.

The public morals jurisprudence resembles health/environment

If we accept Marwell’s concession that adjudicators have rejected unilateralism, the remaining “external standards” are the same ones used to deal with

⁹⁴ Nuzzo (n 22) 280.

⁹⁵ See eg the discussion in section 8.3 where Wu suggests that originalism and internationalism can be decisive (in favour of the regulator) in some circumstances.

⁹⁶ This will be analysed in section 8.5.

health/environment disputes. Wu states that adjudicators “simply confirmed that the interpretative principles that had been applied to other general exceptions should also be extended to the public morals clause”.⁹⁷ He criticises adjudicators for providing “little concrete guidance” and failing to establish unique public morals jurisprudence.⁹⁸

As mentioned above, one area of divergence has been the use of expert evidence. In the early public morals disputes, adjudicators had little resort to expert evidence, though this changed in *Seals* where they relied on both scientific and moral expert evidence. This will be addressed in section 8.5.

Summary

In public morals disputes, adjudicators have a tendency to cite external standards, but it remains unclear if pure unilateralism can singlehandedly ground a claim. The current jurisprudence is characterised by strong unilateralist statements accompanied by the extensive use of external standards in practice. These external standards (originalism, transnationalism and internationalism) are strikingly similar to the approach used in health/environment disputes.

The use of external standards enables adjudicators to avoid some of the sensitivities about public morals by addressing them as an empirical question rather than a normative matter. However, adjudicators apply these external standards in a non-transparent manner which fails to clarify their probative weight. This leaves some discretion for adjudicators to make normatively-intrusive judgment calls, albeit

⁹⁷ Wu (n 5) 229.

⁹⁸ ibid 243.

indirectly, through their unspoken decisions about how to weight the relevant evidence.

8.5 Expert Evidence

This section will consider the role of expert evidence, especially in *Seals*.⁹⁹

Adjudicators have long relied on expert evidence, especially scientific evidence, to help resolve health/environment disputes. Since it is not their function “to settle a scientific debate”, adjudicator’s have established a standard of review and guiding principles for their use of scientific evidence:

“In justifying a measure under Article XX(b) of the GATT 1994, a Member may also rely, in good faith, on scientific sources which, at that time, may represent a divergent, but qualified and respected, opinion. A Member is not obliged, in setting health policy, automatically to follow what, at a given time, may constitute a majority scientific opinion”.¹⁰⁰

This standard of review requires adjudicators to defer to the regulator’s preferred approach, even if it is based on a “divergent” view, provided it comes from a “qualified and respected” source. Where scientific evidence meets this standard, adjudicators tend to give it high evidentiary weight to the point that it is arguably dispositive.

8.5.1 Science and/or Philosophy

The *Seals* adjudicators considered scientific arguments. The EU argued that “it is appropriate to take into account available relevant scientific evidence, as the EU legislators did in this case”.¹⁰¹ The *amici* suggested that “scientists can give a sense of the duration and intensity of animal suffering that likely result from a given killing

⁹⁹ While this section focusses on expert evidence, section 8.6 will consider how *Seals* adjudicators addressed the legal test more broadly.

¹⁰⁰ Appellate Body Report, *Asbestos* (n 81) para 178.

¹⁰¹ EU, ‘First Written Submission by the European Union’ (EC – *Seals*, 21 December 2012) <https://trade.ec.europa.eu/doclib/docs/2012/december/tradoc_150190.pdf> (consulted 5 April 2020) (“EU First Panel Submission”) para 82.

method".¹⁰² In this respect, *Seals* resembles non-moral disputes which use scientific and instrumental reasoning.

However, the EU and *amici* do not use science as a definitive way to demonstrate that "seal welfare" is covered by Article XX(a). Rather, science serves as a mere input to establish the fact that animal exploitation leads to suffering. In order to answer the decisive question, whether animal suffering is a moral matter, the EU and *amici* turn to a second group of experts: moral philosophers.

The *amici* suggest that adjudicators should look to "theories of morality" for guidance.¹⁰³ They argue that the EU's objective can be justified by the welfarist and utilitarian literature of Peter Singer and Jeremy Bentham.¹⁰⁴ While adjudicators have traditionally relied on scientific experts, nothing prevents them from relying on other types of expertise, including from moral philosophy.¹⁰⁵ What role should philosophy play in public morals disputes?

The *amici* recognise the existence of two major approaches to animal ethics: (i) the welfarist approach (led by Peter Singer) which has its roots in instrumental philosophy¹⁰⁶ and (ii) the animal rights approach (led by Francione and Regan) which argues that animals have inherent value.¹⁰⁷ The *amici* argue that Singer's welfarist philosophy supports the moral credentials of the seal products ban, while they are dismissive of Francione's rights-based movement, describing it as

¹⁰² Howse et al (n 68) 371.

¹⁰³ Howse et al (n 34) 96.

¹⁰⁴ Howse et al (n 68) 378. The EU's welfarist defence will be analysed in detail in chapter 9.

¹⁰⁵ The DSU states that "Panels may seek information from any relevant source and may consult experts to obtain their opinion on certain aspects of the matter"; Dispute Settlement Rules: Understanding on Rules and Procedures Governing the Settlement of Disputes, Marrakesh Agreement Establishing the World Trade Organization, Annex 2 (adopted 15 April 1994, entered into force 1 January 1995) 1869 UNTS 401 ("DSU") Art 13(2).

¹⁰⁶ Howse et al (n 68) 378.

¹⁰⁷ ibid 379-380.

“radical”.¹⁰⁸ The EU follows the *amici* by citing the welfarist movement as its preferred school of moral philosophy.¹⁰⁹ It alludes to, but dismisses, rights-based approaches.¹¹⁰

Adjudicators ultimately accept these arguments from animal ethics. The Appellate Body cites “animal welfarism” as a “moral doctrine”,¹¹¹ while the panel refers to the “philosophy of animal welfarism” as a “long-established tradition of moral thought”.¹¹²

What is the standard of review for evidence from moral philosophy? The *amici* argue that adjudicators “should consider instrumental and non-instrumental theories of morality to be equally likely sources of moral value”.¹¹³ This seems consistent with the notion that regulators may choose their preferred approach.

In *Seals*, adjudicators ultimately accepted the EU’s preferred approach (welfarism). The EU was able to rely a welfarist defence (an undoubtedly credible moral approach) without needing to make philosophical arguments showing that welfarism is superior to animal rights. Adjudicators wisely stayed out of “ideological battles”¹¹⁴ and heeded the *amici*’s advice not to “take a stance in philosophical debates”.¹¹⁵ This seems to resemble the standard of review for science where adjudicators defer to Members provided their experts are “qualified and respected”.

Moral philosophy represents a new type of expert evidence in WTO disputes. Even if it is not dispositive, and is merely another persuasive external standard, philosophy

¹⁰⁸ ibid 379.

¹⁰⁹ EU, ‘Second Written Submission by the European Union’ (*EC – Seals*, 27 March 2013) <https://trade.ec.europa.eu/doclib/docs/2013/april/tradoc_150812.pdf> (consulted 5 April 2020) (“EU Second Panel Submission”) para 54.

¹¹⁰ ibid 41 (footnote 157).

¹¹¹ Appellate Body Reports, *EC – Seals* (n 4) para 2.159.

¹¹² Panel Reports, *EC – Seals* (n 3) para 7.408.

¹¹³ Howse et al (n 34) 96.

¹¹⁴ Wu (n 5) 237.

¹¹⁵ Howse et al (n 34) 97.

still offers adjudicators a useful tool to outsource the contested normative question. Adjudicators need not make a controversial, normative assessment about the moral validity of “seal welfare”. Rather, they can ask an empirical question: would any “qualified and respected” school of animal ethics recognise “seal welfare” as moral?

8.5.2 Reconciling the Roles of Science and Philosophy

By relying on expert evidence, *Seals* may have drawn Article XX(a) even closer to the health/environment jurisprudence. However, *Seals* is also unique as the EU’s animal welfare claim relied on experts from two different camps. It relied on science for the purpose of demonstrating that animals are capable of suffering (a factual matter) and it relied on the welfarist movement in animal ethics to show that animal suffering is morally significant (a normative matter).

The relationship between the two types of evidence is apparent from the EU’s claim:

“While in selecting a level of protection of public morals it is appropriate to take into account available relevant scientific evidence, as the EU legislators did in this case, the choice of a level of protection of public morals is not a scientific judgement. It is a policy decision involving a moral judgement”.¹¹⁶

The two types of expert evidence work together as a combination: science provides a factual input, while philosophy does the heavy lifting of demonstrating that “seal welfare” is a legitimate objective.

The approach in *Seals* is significant in several respects. It is the first time adjudicators recognise moral philosophy as a type of “expert” evidence. This could signal the start of a new trend where parties invoke moral philosophy as an external standard to support Article XX(a) defences.

¹¹⁶ EU First Panel Submission (n 101) para 82.

Seals is also significant because adjudicators relied on two different types of expert evidence and relegated science to the status of a mere “input”. Unlike the dispositive approach in disputes such as *Asbestos*, science was unable to demonstrate that “seal welfare” is a legitimate objective under Article XX(a) until it was combined with arguments from animal ethics.

8.5.3 What Makes Public Morals Unique?

The dual use of science and philosophy seems to confirm that there is something different about public morals compared to health/environmental defences (which can be resolved using science alone). This section will seek to offer an explanation for this difference.

For health measures, GATT drafters explicitly allowed Members to adopt measures necessary to protect “human... life and health”. Therefore, adjudicators are not required to establish that human life/health is a legitimate objective (nor can they question its legitimacy).¹¹⁷ GATT drafters undertook the normative work by giving clear guidance that human life/health has inherent value. The adjudicator’s role is limited to the following question: does the measure at issue protect human life/health? This can be resolved using scientific evidence and instrumental reasoning.

Environmental protection is more complicated. It is unclear if GATT drafters intended Article XX(g) to allow environmental protection writ large. However, in *Shrimp*, the Appellate Body interpreted Article XX(g) broadly thanks to the Marrakesh Agreement Preamble which recognised the "objective of sustainable

¹¹⁷ Even if GATT drafters had not explicitly referred to human life/health, there is little doubt that adjudicators could have found compelling evidence to show that it is a legitimate objective.

development, seeking both to protect and preserve the environment".¹¹⁸ It further noted that signatories were "fully aware of the importance and legitimacy of environmental protection".¹¹⁹ Regardless of the GATT drafters' original intention, this interpretative act by the Appellate Body (aided by the drafters of the Marrakesh Preamble) recognised the inherent value of environmental protection.

For health/environmental objectives, it is therefore clear that adjudicators can avoid normative questions. They have been given "standing instructions" on the inherent value of these objectives¹²⁰ and must merely perform an instrumental operation (often supported by science) to see if the measure contributes to these objectives.

Article XX(a) appears different. GATT drafters acknowledged the legitimacy of moral regulation, but provided little guidance to adjudicators about which measures fall within the ambit of public morals. In order to determine if a measure is moral, must adjudicators engage in normative reasoning?

Several leading authors have proposed an interpretative approach which prevents normative intrusiveness by distinguishing between moral "categories" and moral "content". This approach allows adjudicators to determine if a measure addresses a moral matter (or category), but denies them the right to pass judgment on the desirability of the moral norm or its content. I will call this the "Category/Content Distinction".

¹¹⁸ Appellate Body Report, *Shrimp* (n 14) para 153.

¹¹⁹ ibid para 129.

¹²⁰ Thanks to the direct guidance of the GATT drafters or their own interpretative acts.

Wu suggested that public moral objectives should be accepted where they relate to “a category that is widely recognized as a moral issue”.¹²¹ In such cases, regulators would only have to show that “the category to which the issue belongs falls under the auspices of public morality”, but not that the content of the contested norm is moral or widely-accepted.¹²² Marwell agrees that adjudicators should “restrict the public morals clause to the regulation of issues broadly agreed to be matters of moral judgment”.¹²³

This approach responds to Diebold’s concern that “Israel would be unable to prohibit the importation of non-kosher meat products” under a non-unilateralist standard.¹²⁴ Under the Category/Content Distinction, Israel could produce evidence supporting the legitimacy of its “category” (religious norms), but would not be required to prove that its specific content (Kosher rules) is morally-desirable or widely-accepted.

The *amici* suggest that WTO adjudicators:

“should inquire whether the justification given is the type of justification that could legitimately be called a moral reason, under any recognizable conception of what is moral”.¹²⁵

Wu, Marwell and the *amici* inquire whether a “matter”, “category” or “reason” is moral. However, they strongly reject any normative assessment by adjudicators about the content of the moral norm. The *amici* assert that:

“the application of Article XX(a) should not entail any assessment of the validity or desirability of the reasons given”.¹²⁶

¹²¹ Wu (n 5) 243. It is important to note that Wu puts forward a composite test which looks to originalism and internationalism in the first instance. The “widely-recognised” moral matter is a tiebreaker for all remaining objectives.

¹²² *ibid* 244.

¹²³ Marwell (n 6) 820.

¹²⁴ Diebold (n 5) 54.

¹²⁵ Howse et al (n 34) 104.

¹²⁶ *ibid* 104-5.

Marwell observes that this approach is useful for issues, such as the death penalty or abortion, where “a diversity of substantive opinions exist as to the content of that judgment”.¹²⁷ The Category/Content Distinction preserves the regulatory autonomy of Members and keeps adjudicators away from ideological quagmires about controversial and sensitive moral content.

Extrapolating Marwell’s abortion example, adjudicators could find that pro-life measures are a legitimate public morals matter (in Dispute A) and that pro-choice policies are equally legitimate (in Dispute B). These findings may seem superficially contradictory, however the Category/Content Distinction offers an elegant solution. It enables adjudicators to determine that abortion is a moral matter (a non-controversial finding) while staying silent (and normatively-neutral) on the sensitive issue of “content”. Adjudicators can find that both pro-life and pro-choice views are legitimate, even if they are diametrically-opposed.

Does the Category/Content Distinction fully absolve adjudicators from engaging in normative analysis? It certainly allows adjudicators to avoid certain “hard” normative questions about how governments should regulate abortion or which religious laws to follow. However, it fails to address a deeper question: how do we determine if a “category” or “matter” is moral in nature? Is there universal agreement about which “categories” are moral or is this also a culturally-specific question?

Determining if a matter is “moral” is invariably a normative question, but the proponents of the Category/Content Distinction do not offer a test for making this determination. In the examples they offer, the answer is intuitively obvious. The death penalty and abortion are obviously moral matters, even if there are

¹²⁷ Marwell (n 6) 820.

diametrically-opposed views on the appropriate content of the moral norm. Could disputes arise where it is not obvious (or universally-agreed) that a particular “category” is moral?¹²⁸

Adjudicators appear unwilling to rely on “intuitive obviousness” as a basis to determine whether a matter is moral or not. However, as shown in section 8.4, they are adept at converting normative questions into descriptive ones by using external standards such as transnationalism and internationalism. In *Seals*, moral philosophy was adopted as an additional external standard which can establish that a given matter is “moral”.¹²⁹

8.5.4 Summary

In *Seals*, adjudicators relied on scientific evidence to show that seal exploitation produces suffering and that the seal products ban reduces cruelty. This type of instrumental assessment closely resembles the use of science in health/environmental disputes.

However, adjudicators also relied on moral philosophy, including the ethics of Singer and Bentham, to determine that animal suffering is a moral matter. Adjudicators may well have minimised normative intrusiveness by focussing on the moral “category” rather than its “content”, but they still had to address an underlying normative question. This normative assessment distinguishes public morals from health/environmental disputes where adjudicators have inherited clear normative

¹²⁸ The EU First Panel Submission (n 101, para 75) mentions horse meat bans in some US states as a moral matter. Is a horse meat ban a moral matter or a mere social convention? Would those parties on the other side of the debate (producers and consumers of horse meat) agree that banning horse meat qualifies as a moral matter? This question will be addressed in chapter 10.

¹²⁹ At its core, the Category/Content Distinction contains a normative question. In most cases, the answer is intuitively obvious. If adjudicators seek to explicitly justify their normative finding under the Category/Content Distinction, they can rely on moral philosophy. However, a key limitation with the use of moral philosophy (which I will address in subsequent chapters) is that it is only capable of justifying rationalist moral approaches, but struggles to justify moral intuitions.

guidance and can therefore limit their “design step” inquiry to scientific evidence and instrumental reasoning.

8.6 The *Seals* Dispute

The previous section considered how expert evidence, from both moral philosophy and science, played an important role in *Seals*. This section will consider the other aspects of the legal test under the “design step” including whether adjudicators relied on unilateral and/or external evidence.

8.6.1 Legal Test

In *Seals*, the EU argued for a unilateralist test:

“The European Parliament and the EU Council are best placed, and uniquely legitimised, to recognise and interpret the moral concerns of the European people that they represent. The adoption in good faith of the measure at issue by those two institutions in order to address the moral concerns invoked in this dispute is, in itself, sufficient evidence of such concerns”.¹³⁰

This resembles Charnovitz’s pure unilateralism where nothing more than a self-declaration would be required. However, the EU also provided evidence of its social norm just in case adjudicators opted for evidentiary unilateralism. It cited public opinion polls and public campaigns to end seal slaughter.¹³¹

Canada argued against unilateralism:

“[i]f it is simply left to the legislator to decide when the suffering of animal[s] is justified, the content of the so-called moral standard becomes inherently subjective and therefore arbitrary”.¹³²

The *Seals* adjudicators appeared to resolve this matter in favour of the EU. They made statements implying that public morals are to be determined unilaterally: “our

¹³⁰ The EU First Panel Submission (n 101) para 189.

¹³¹ Appellate Body Report, *EC – Seals* (n 4) paras 5.135 and 5.153.

¹³² *ibid* para 2.195.

task is confined to assessing whether the public concerns on seal welfare are anchored in the morality of European societies".¹³³

The *Seals* panel also introduced a minor innovation into the public morals “design step” by dividing it into two separate elements: (i) whether there were “public concerns regarding seal welfare” and (ii) whether those public concerns were “moral” in nature.¹³⁴ This approach added a new “public concerns” requirement to complement the pre-existing “moral standards” requirement. How did adjudicators assess these two elements in practice?

8.6.2 Application of Legal Test

This section will demonstrate that adjudicators took a unilateralist approach to the public concerns element, but that they considered both unilateral and external standards for the moral standards requirement.

8.6.2.1 Public Concerns

Adjudicators found the EU satisfied the “public concerns” requirement. The panel noted:

“During the last decade, the cruelty of seal hunting has been documented by videos from several authoritative television channels as well as by the personal observations of many members of national and European parliaments, scientists, celebrities and representatives of non-governmental organizations (NGOs). Such cruelty has generated a public morality debate in Europe”.¹³⁵

The Appellate Body further noted that:

¹³³ Panel Reports, *EC – Seals* (n 3) para 7.404.

¹³⁴ *ibid* para 7.403.

¹³⁵ *ibid* para 7.392.

“The Commission received during the last years a massive number of letters and petitions on the issue expressing citizens' deep indignation and repulsion regarding the trade in seal products”.¹³⁶

“Public survey results submitted by the European Union were also informative, although "to a limited extent", in demonstrating the EU public's concerns”.¹³⁷

This analysis is noteworthy as it relies exclusively on unilateral evidence about moral norms within the EU (from parliamentarians, media, petitions and surveys). The new “public concerns” requirement appears to address an empirical question (whether a certain norm exists in the EU). In principle, this question can be answered using pure unilateralism without needing to assess the views of other societies or moral philosophers.

8.6.2.2 Moral Standards

The panel's analysis of moral standards also included some unilateral evidence. It referred to the Treaty of Lisbon's requirement to “pay full regard to the welfare requirements of animals” since they are “sentient beings”¹³⁸ as well as certain measures and legislation in the EU and its Member States.¹³⁹ It further referred to “a comprehensive body of legislation on the welfare of farm animals within the framework of its Common Agricultural Policy”.¹⁴⁰ The panel further noted that the EU self-describes its public concerns as being ethical in nature: “we observe that the Commission Proposal describes public concerns as relating to "ethical" considerations”.¹⁴¹

¹³⁶ Appellate Body Reports, *EC – Seals* (n 4) para 5.153.

¹³⁷ ibid para 5.135.

¹³⁸ Panel Reports, *EC – Seals* (n 3) para 7.406.

¹³⁹ ibid 7.405.

¹⁴⁰ ibid 7.406.

¹⁴¹ ibid 7.395. In my view, it is poor practice for adjudicators to rely on the regulator's choice of language to make legal findings in favour of the regulator as this can easily be manipulated.

However, the moral standards analysis required adjudicators to consider external standards. The panel used transnationalism when it cited “certain other WTO Members' measures”¹⁴² and internationalism when it cited the “recommendations of the Office International des Epizooties (OIE) (Guiding Principles for Animal Welfare)”.¹⁴³ Finally, it incorporated evidence from moral philosophy and animal ethics (as analysed in section 8.5).¹⁴⁴

The Appellate Body summarised the panel's reasoning on moral standards as follows:

“The Panel considered the legislative history of the EU Seal Regime, as well as a range of other evidence, including various actions taken by the European Union as well as EU member States concerning animal welfare protection in general; various pieces of legislation and conventions on animal welfare within the European Union and other countries, including Norway and Canada; and various international instruments”.¹⁴⁵

This summary alludes to both unilateral and external evidence (transnationalism and internationalism) in order to demonstrate “moral standards”. Adjudicators further relied on the additional external standard of moral philosophy when they considered “welfarist” arguments.¹⁴⁶

At both the panel and appellate stages, adjudicators relied extensively on external standards from transnationalism, internationalism and moral philosophy. However, they also make statements which suggest that these types of evidence do not have much probative value:

“International doctrines and measures of a similar nature in other WTO Members, while not necessarily relevant to identifying the European Union's

¹⁴² ibid 7.408.

¹⁴³ ibid.

¹⁴⁴ ibid. The panel also references the *amicus*'s submission which introduces moral philosophy in more detail (footnote 672).

¹⁴⁵ Appellate Body Reports, *EC – Seals* (n 4) para 5.138.

¹⁴⁶ ibid para 2.159.

chosen objective, illustrate that animal welfare is a matter of ethical responsibility for human beings in general".¹⁴⁷

These types of statements seem to downplay the role of external standards. However, the adjudicators' practice of extensively citing external standards suggests they actually play an important role in the *Seals* dispute.

8.6.3 Summary

Seals contained some jurisprudential innovation, as adjudicators created a new element (public concerns) which appears to be based on pure unilateralism.

However, the continued presence of the moral standards requirement, based on external standards, ensures that the test also contains non-unilateralist elements.

The jurisprudence contains an unresolved tension about the relationship between unilateralism and external standards. If the public morals test is truly unilateralist, the Appellate Body should remove the "moral standards" requirement and cease relying on external evidence. Alternatively, if regulators must produce objective evidence to support their measure, the Appellate Body should tone down its language which implies that moral norms can be defined unilaterally. There are many cases where regulators will be able to rely on both unilateral and external evidence to support the same finding (as was the case in the early disputes).

However, the jurisprudence remains unclear about whether unilateralism can singlehandedly ground a defence where there is little external evidence supporting the legitimacy of a moral norm.

¹⁴⁷ Panel Reports, EC – *Seals* (n 3) para 7.409.

8.7 Conclusion

Are public morals different from other general exceptions such as health/environment? My analysis suggests that certain scholarly claims about public morals (that they are vaguer than other exceptions or more likely to be abused) are empirically and conceptually questionable. However, other claims (that some moral norms may be culturally-specific) are valid and merit deeper attention.

A key difference with public morals is that GATT drafters left open the normative question (which matters are moral) to be considered by adjudicators on a case-by-case basis. For health/environmental objectives, GATT drafters left clear guidance on the normative question so the analysis by adjudicators is purely instrumental.

The *amici's* distinction between instrumental and non-instrumental norms is useful for understanding public morals. For those moral objectives which are instrumental in nature, adjudicators can address the design step using the same standards they apply to health/environmental objectives, including transnationalism and internationalism. This was the case in early disputes, such as *Gambling*, where the US successfully used transnationalism to support its objective. Where external standards are available, it is sensible for adjudicators to cite them and take them into account.

However, there may also be non-instrumental (and culturally-specific) moral norms that cannot be adequately addressed solely by reference to science, instrumental reasoning or external standards. Scholars seek to resolve this problem with the Category/Content Distinction which allows adjudicators to determine if a regulatory category is moral (such as religion) without inquiring whether the specific content is morally desirable (such as kosher rules).

The Category/Content Distinction may help Members defend culturally specific norms which cannot be supported by transnationalism or internationalism. It would also enable adjudicators to avoid making controversial assessments on matters where Members hold diametrically-opposed views (such as abortion).

However, determining whether a “category” is moral in nature remains a normative inquiry. In many cases, the answer might be obvious, but could there be cases where it is not? In *Seals*, adjudicators introduced moral philosophy as a type of external standard which can help resolve such disputes. For future cases involving culturally-specific norms, adjudicators should continue relying on the Category/Content Distinction and drawing on moral philosophy to help determine which categories are moral.

Despite the sensitive issues raised in *Seals*, adjudicators failed to provide a clear answer on the status of pure unilateralism. Adjudicators introduced a new requirement (public concerns) which appears to be based on pure unilateralism. However, they also retained the “moral standards” requirement which is based on a still-to-be-clarified combination of unilateralism and external standards.

CHAPTER 9: THE WELFARIST DEFENCE

9.1 Moral Philosophy in WTO Law

At the core of an Article XX(a) defence, there is an underlying question about the role of morality. The legal text refers to “public morals” which the case law defines as “moral standards” or “standards of right and wrong”. Despite the language of morality in Article XX(a), early disputes did not consider arguments from moral philosophy.¹ This changed in *Seals* where adjudicators acknowledged the relevance of philosophical arguments by the EU and *amici*.²

The relevance of moral philosophy finds support from scholars. Herwig argues that:

“the interpretation of the terms “legitimate objectives” and “public morals” necessarily connects WTO law to a knowledge system beyond law, namely to moral philosophy, because it is this system of knowledge that expounds the demands of legitimacy and morality”.³

This section will elaborate on the role played by moral philosophy in the *Seals* dispute. Its main focus will be the EU’s welfarist defence, including the Suffering/Utility Standard which is the vehicle through which the EU advances welfarist arguments. Even though the EU explicitly cites welfarism and alludes to its leading authors, this chapter will show that its approach to animal exploitation is highly inconsistent with Singer’s philosophy in a number of key respects.⁴

¹ See eg Appellate Body Report, *United States – Measures Affecting the Cross-Border Supply of Gambling and Betting Services*, WT/DS285/AB/R, adopted 20 April 2005 (*US – Gambling*); Appellate Body Report, *China – Measures Affecting Trading Rights and Distribution Services for Certain Publications and Audiovisual Entertainment Products*, WT/DS363/AB/R, adopted 19 January 2010 (*China Audiovisuals*).

² This matter was analysed in the previous chapter.

³ Alexia Herwig, ‘Too much Zeal on Seals? Animal Welfare, Public Morals, and Consumer Ethics at the Bar of the WTO’ (2016) 15(1) World Trade Review 109, 132.

⁴ The detailed analysis of the EU’s welfarist defence (in this chapter) serves as a contrast to the intuitionist defence (which will be described in the next chapter).

This chapter will also briefly consider the animal rights movement which is alluded to by the EU and some scholars.⁵ Finally, the analysis will consider how moral philosophy interacts with other key notions in WTO law such as the level of protection and the standard of review.

9.2 Animal Welfarism

9.2.1 Introduction

There is a widespread belief that the *Seals* dispute is about welfarism, including many observers who describe the dispute in the language of “animal welfare”.

Writing prior to the dispute, Katie Sykes suggested that *Seals* will “probably be the first decision by an international adjudicator in which *animal welfare* is one of the central questions.”⁶ The *amici* contended that the *Seals* findings “open the door to future *animal welfare* defenses and provide an important endorsement for the protection of *animal welfare*.⁷

The EU explicitly framed its defence in “animal welfare” terms when it argued:

“*animal welfare* is recognised as a value of concern to the European Union and has been enshrined by the Treaty of Lisbon”.⁸

The EU cited moral philosophy and explicitly stated that its “moral standards” were based on a welfarist approach:

⁵ Herwig (n 3) also considers a further school of moral philosophy, “virtue ethics”. This school of thought receives little academic attention and is not taken up by the parties or adjudicators.

⁶ Katie Sykes, ‘Sealing Animal Welfare into the GATT Exceptions: the International Dimension of Animal Welfare in WTO disputes’ (2014) 13(3) *World Trade Review* 471, 472.

⁷ Robert Howse, Joanna Langille and Katie Sykes, ‘Pluralism in Practice: Moral Legislation and the Law of the WTO after Seal Products’ 48 *George Washington International Law Review* (2015) 81, 113.

⁸ EU, ‘First Written Submission by the European Union’ (*EC – Seals*, 21 December 2012) <https://trade.ec.europa.eu/doclib/docs/2012/december/tradoc_150190.pdf> (consulted 5 April 2020) (“EU First Panel Submission”) para 63. This reference to “sentience” draws directly from utilitarian philosophers, such as Bentham and Singer, who are considered the leading thinkers on welfarism.

“Humans are not free to treat and use animals as they wish, but ought instead to conform to certain moral standards of right and wrong”.⁹

“the EU Seal Regime seeks to uphold a standard of conduct according to which it is morally wrong for humans to inflict suffering upon animals without sufficient justification. This basic rule reflects a long-established tradition of moral thought, which in its modern form is usually designated as ‘animal welfarism’. This moral doctrine on the relationship between humans and animals is the most widely held in modern societies”.¹⁰

In support of its welfarist argument, the EU further cites the *amici*¹¹ who provide useful background detail about welfarism.¹² They cite Peter Singer as the leading modern proponent of welfarism and describe his book *Animal Liberation* as “perhaps the most famous exposition of the new welfarist position”.¹³ They further recognise that the welfarist movement is part of a broader utilitarian tradition dating back to Jeremy Bentham.¹⁴

Singer’s welfarism is based on two key principles which I will expound in the following sections. First, I will address the Suffering/Utility Standard which the EU heavily relies on. Second, I will consider welfarism’s consistency requirement which the EU ignores.

9.2.2 The Suffering/Utility Standard

The EU draws directly from welfarism by arguing that animal exploitation can be justified when utility exceeds the suffering produced. For ease of reference, I label

⁹ *ibid* para 61.

¹⁰ EU, ‘Second Written Submission by the European Union’ (*EC – Seals*, 27 March 2013) <https://trade.ec.europa.eu/doclib/docs/2013/april/tradoc_150812.pdf> (consulted 5 April 2020) (“EU Second Panel Submission”) para 140.

¹¹ The *amici* produced two amicus briefs (at the panel and appellate stages) and a number of scholarly articles (before and after the *Seals* dispute).

¹² EU Second Panel Submission (n 10) para 140 (footnote 157).

¹³ Robert Howse and Joanna Langille, ‘Permitting Pluralism: The Seal Products Dispute and Why the WTO Should Accept Trade Restrictions Justified by Noninstrumental Moral Values’ (2012) 37(2) *Yale Journal of International Law* 367, 378 (footnote 73).

¹⁴ *ibid* 378.

this principle the “Suffering/Utility Standard” and provide examples (below) of its use in practice.¹⁵

The EU argues that “it is morally wrong for humans to inflict suffering upon animals without sufficient justification”.¹⁶ Further, the EU’s recognition of the moral significance of animal suffering derives from the Lisbon Treaty’s respect for animals as “sentient beings”.¹⁷

The Suffering/Utility Standard is also embedded in the legislation of EU Member States. The German Welfare Act requires that “no person may cause an animal pain, suffering or harm without good reason” and it prohibits “the infliction of unjustified pain”.¹⁸ In 2002, Germany became the first EU Member to “guarantee animal rights in its constitution”.¹⁹

The *amici*’s moral arguments are based on this same principle:

“For an animal welfarist, it may be acceptable to kill animals, provided that there is some basic human need or purpose served by killing the animal and that the animal does not suffer unnecessarily”.²⁰

Katie Sykes suggests that animal welfare is enshrined in a “general principle of international law” that “inflicting unnecessary or gratuitous suffering on animals

¹⁵ Although the EU and its academic supporters use different language to express the idea, I will use deconstruction to show that they all base their arguments on this principle.

¹⁶ EU Second Panel Submission (n 10) para 140.

¹⁷ The Seal Regime was also challenged under EU law where the EU relied on the “sentient beings” provision from the Lisbon Treaty to demonstrate that it had competence to ban seal products at a region-wide level. For a detailed discussion of the legal issues raised by EU law, see Tamara Perišin, ‘Is the EU Seal Products Regulation a Sealed Deal? EU and WTO Challenges’ (2013) 62(2) International & Comparative Law Quarterly 373. Perišin highlights that many of the arguments made before the ECJ actually contradict the EU’s WTO arguments.

¹⁸ Kate Sowery, ‘Sentient Beings and Tradable Products: The Curious Constitutional Status of Animals under Union Law’ (2018) 55(1) Common Market Law Review 10.

¹⁹ Cass Sunstein, ‘Introduction: What Are Animal Rights?’ in Cass Sunstein and Martha Nussbaum (eds), *Animal Rights: Current Debates and New Directions* (OUP 2006) 4.

²⁰ Howse et al (n 13) 379.

should be proscribed.”²¹ She argues that international law condemns “unnecessary” and “avoidable” animal suffering.²²

All of these descriptions of the moral standard (from European law, international law, welfarist philosophy and legal scholars) rely on three key elements:

- (i) a determination of the negative effects of animal exploitation (suffering);
- (ii) a determination of positive effects embedded in animal products (utility);²³ and
- (iii) an assessment of whether the benefits outweigh the costs (commensuration and balancing).²⁴

The Suffering/Utility Standard is a consequentialist technique for making trade-offs between competing values. It therefore resembles cost-benefit analysis or, indeed, the WTO’s own W&B test. The subsequent analysis will demonstrate that, in practice, the EU and *amicus*’s defence of the ban is not truly consequentialist as it focusses almost exclusively on the “suffering” element of the standard to the exclusion of “utility” and “balancing”.

When an adjudicator, party or analyst refers to “sentience”, “suffering”, “cruelty” or “inhumane” acts, they are ultimately invoking welfarism’s suffering criteria. However, the EU makes no substantive claims about “utility” to defend the seal products ban.

²¹ Sykes (n 6) 480.

²² *ibid* 493.

²³ I use the neutral term “utility”, though this can also be captured in different formulations by asking if the animal suffering is “necessary”, “justifiable” or “essential”.

²⁴ This “balancing” is based on consequentialist reasoning, though proponents are vague about precisely which balancing technique they endorse (such as CBA, proportionality or balancing).

It does not seek to demonstrate that seal products have a low utility²⁵ or that other animal products are more useful. It is essentially silent on this matter.²⁶

The EU's claim that the Suffering/Utility Standard is a key pillar of welfarist philosophy is correct. Singer and Bentham would both endorse the EU's premise that animal exploitation can only be justified where the utility exceeds the suffering. However, the EU and *amici* appear to cherrypick from welfarism as they overlook other key ideas in Singer's philosophy which, if considered, would seriously undermine the EU's defence.

First, the EU applies the Suffering/Utility Standard in a way which would be anathema to Singer and welfarism. The level of protection chosen by the EU leads to the slaughter of billions of animals each year in ways which Singer would consider unethical.

Second, the EU and *amici* fail to address welfarism's consistency requirement which demands "equal consideration of interests" for all animals.²⁷ According to welfarists, all animal suffering should be given equal weight from a moral point of view. Welfarists would denounce any approach which considered the suffering of seals or dogs (or even humans) to be morally more significant than the suffering of cows and pigs. The next two sections will address each of these weaknesses in the EU claim.

²⁵ The EU makes a few passing comments which hint that seal products have low utility. For example, the EU refers to seal products as "inessential clothing items", however it fails to make any substantive arguments clarifying just how important/useful animal products are under the Suffering/Utility Standard; Panel Reports, *European Communities – Measures Prohibiting the Importation and Marketing of Seal Products*, WT/DS400/R and Add.1 / WT/DS401/R and Add.1, adopted 18 June 2014 (*Seals*) para 7.369.

²⁶ The EU only makes substantive arguments about utility when seeking to defend the Inuit Exception to the ban on the grounds that these products produce a high utility. This matter falls outside the scope of my analysis as I do not address chapeau issues. Nevertheless, it is revealing to note that the EU ban and exceptions are really based on two different standards. Superficially, they can both be accommodated under the "Suffering/Utility Standard". In reality, the ban is based on a pure "suffering" standard (with no assessment of utility) while the Inuit Exception is justified solely by a "utility" standard (with no assessment of suffering). This highlights the malleability of the Suffering/Utility Standard when it is not applied with coherent and consistent moral reasoning.

²⁷ Singer refers to this notion extensively in *Animal Liberation* (3rd edition, HarperCollins 2002).

9.2.3 The Level of Protection

Singer agrees with the EU that animal exploitation and slaughter can be moral in some circumstances. He even accepts that suffering can sometimes be morally justifiable:

“It is true that on a utilitarian view there could conceivably be circumstances in which an experiment on an animal stands to reduce suffering so much that it would be permissible to carry it out even if it involved some harm to the animal”.²⁸

Singer does not seek absolute protection for animals.²⁹ However, unlike the EU, he adopts a high level of protection which, in practice, rarely finds that the utility can justify the suffering caused. He would apply the Suffering/Utility Standard as follows:

Suffering - Singer notes that animal exploitation leads to enormous amounts of animal suffering: “Apart from taking their lives there are also many other things done to animals in order to bring them cheaply to our dinner table. Castration, the separation of mother and young, the breaking up of herds, branding, transporting, and finally the moments of slaughter - all of these are likely to involve suffering”.³⁰

Utility - While the exploitation of commodity animals leads to immense suffering, Singer assesses that it produces little utility. He suggests that “industrialized societies can easily obtain an adequate diet without the use of animal flesh” and describes it as a “luxury” rather than a necessity.³¹

Balancing - When balancing suffering and utility, Singer concludes that “appalling suffering is being inflicted on millions of animals for purposes that on any impartial view are obviously inadequate to justify the suffering.”³² He rejects any animal exploitation which is based on “factory farm methods”.³³ This is obviously a very different conclusion to the EU which allows the production and consumption of billions of factory farmed animals every year.³⁴

²⁸ Peter Singer and Tom Regan, ‘Dog in the Lifeboat: An Exchange’ *New York Review of Books* (25 April 1985).

²⁹ Such moral absolutes would be expected from deontological philosophers, not from utilitarians.

³⁰ Peter Singer, *Practical Ethics* (3rd edition, CUP 2011) 55-56. On the claim that factory farming entails immense amounts of suffering, Singer would find support from animal rights philosophers such as Francione who states that “welfare laws provide an insignificant level of protection to nonhuman animals”; Gary Francione, ‘Animal Welfare and the Moral Value of Nonhuman Animals’ (2010) 6(1) *Law, Culture and the Humanities* 24, 25.

³¹ Singer (n 30) 54.

³² Singer (n 27) 85.

³³ Singer (n 30) 55.

³⁴ Singer’s philosophy would permit animal exploitation which is truly pain-free throughout the full life cycle of the animal. Factory farming is incapable of passing this test. He would also permit exploitation which provides

This raises a key question: was it legitimate for the EU to rely on welfarism's Suffering/Utility Standard while adopting a level of protection that would be anathema to welfarists? The EU does not address this incongruity directly, but makes statements declaring its right to deviate from Singer's approach:

"While in selecting a level of protection of public morals it is appropriate to take into account available relevant scientific evidence, as the EU legislators did in this case, the choice of a level of protection of public morals is not a scientific judgement. It is a policy decision involving a moral judgement which, in the present case, was the exclusive prerogative of the EU legislators".³⁵

The *amici* elaborate on this argument:

"Although scientists can give a sense of the duration and intensity of animal suffering that likely result from a given killing method, it is ultimately the predominant moral beliefs of a particular society that will determine how much and what kinds of suffering are acceptable or unacceptable to that society, and therefore the level of protection it demands against animal suffering".³⁶

These quotes are useful for clarifying where the EU's defence overlaps with Singer and where it deviates. At the level of principle, Singer and the EU both subscribe to the Suffering/Utility Standard. Further, they both treat animal suffering as an objective and factual matter which can be determined using scientific assessments. In other words, when Singer and the EU apply the Suffering/Utility Standard, they both use the same value for "suffering".³⁷

However, there are also some significant differences between Singer and the EU. While Singer places significant emphasis on the lack of utility of animal products, the EU completely sidesteps this question. In theory, the EU's defence might be based

so much utility that it can justify the suffering involved. Singer doubts that this high utility standard could be satisfied for animal food and clothing products which can be easily replaced in modern societies with non-animal alternatives.

³⁵ EU First Panel Submission (n 8) para 82.

³⁶ Howse et al (n 13) 371

³⁷ When applying the Suffering/Utility Standard, many consumers may base their moral judgments on subjective perceptions of how much an animal suffers or how much they care about that animal's suffering. However, the EU claims to base its animal welfare settings on science and risk which means that actual suffering (not subjective perceptions of suffering) are what matters.

on the view that factory farmed products are extraordinarily useful, but if this is the site of the EU's disagreement with Singer, it articulates no arguments to this effect.

It appears likely that the difference between the EU and Singer can be explained by balancing. Singer considers animal suffering to be morally important and to be avoided at (almost) any cost. By contrast, the EU slaughters billions of animals each year to produce food and clothing items that Singer would shun in favour of alternative (like) products made without animals. The EU shares Singer's view that animal suffering is morally-relevant, but it clearly has a less animal-friendly approach to "balancing" human and animal interests.

Under Article XX(a), can the EU rely on welfarism, while implementing a level of protection which is completely inconsistent with the view of leading welfarists? It is unfortunate that adjudicators failed to engage with this question. I will address it briefly.

Under WTO law, the EU certainly has an "exclusive prerogative" to choose its preferred level of protection.³⁸ This suggests the EU was not required to follow Singer's level of protection under the Suffering/Utility Standard. However, the EU's low level of concern for animal suffering must raise doubts that its regulatory approach is truly based on the philosophy of Singer and Bentham. The "welfarist" branding provides much-needed credibility to the EU's defence, but the claim seems questionable. How can the EU claim to pursue welfarism when Singer would consider its approach abhorrent?

³⁸ The Appellate Body has stated "the fundamental principle is the right that WTO Members have to determine the level of protection that they consider appropriate"; Appellate Body Report, *Brazil – Measures Affecting Imports of Retreaded Tyres*, WT/DS332/AB/R, adopted 17 December 2007 (*Brazil Tyres*) para 210.

This incongruence does not deter the *amici*. Rather than seeing the non-compliance with Singer as a weakness (because the level of protection is too low), they go on the offensive and laud the EU for its high level of protection. They suggest that “Canada and Norway simply cannot agree with the [high] level of protection called for by European values”.³⁹ Sykes suggests that animal welfare is “a signature European value”.⁴⁰

The EU is in a difficult situation. There is no evidence that it wishes to adopt a level of animal protection which would be acceptable to Singer. However, it wishes to make rhetorical arguments which attract the moral credibility of Singer and welfarism. This approach is incoherent. The EU is not truly adopting a welfarist approach. It has designed a unique approach which has some similarities with, but also some serious deviations from, welfarism. For some reason, WTO adjudicators (and many scholars) simply ignore this incoherence and recognise the EU's objective as °welfarist° with a high level of protection.

9.2.4 Non-Discrimination and Speciesism

In addition to its non-compliance with the welfarist level of protection, the EU also ignores welfarism's consistency requirement. Singer and Bentham argue that utilitarianism endorses the “principle of equal consideration of interests” and seek to extend this principle to all sentient beings.⁴¹ In the case of welfarism, this means that all animals have an equal interest in not suffering. Moral judgments should refrain from discounting the suffering of some species (such as cows and pigs) or placing a premium on the suffering of other species (such as dogs and seals).

³⁹ Howse et al (n 13) 389.

⁴⁰ Sykes (n 6) 478.

⁴¹ Singer (n 27) 6-7.

Welfarism's consistency requirement is captured in the proscription against discrimination on the basis of species (often referred to as "speciesism"). Animal Liberation (Singer's seminal tract cited by the *amici*) contains dedicated chapters on the immorality of speciesism⁴² where Singer declares that "all animals are equal".⁴³

Since the consistency requirement is so central to welfarism's normative philosophy, the failure to acknowledge or address it leaves a significant gap in the analysis of the EU and *amici*.⁴⁴ I will address this gap with a brief discussion of "speciesism" and its opposition to (i) human supremacy and (ii) special protection for certain nonhuman animals.

Human Supremacy

Although it is not relevant to *Seals*, it is useful to discuss "human supremacy" as this is where much of the debate about "speciesism" takes place. Philosophers have inquired whether humans are morally superior to (or deserve better treatment than) other species. In modern societies, it is widely accepted (and intuitively obvious) that we should not exploit human bodies for food and clothing. Welfarists question if there is a philosophical basis for this special status of humans.

For a welfarist, humans cannot just self-declare that they are morally superior to nonhuman animals. Special treatment can only be justified if there are objective and morally-relevant criteria which distinguish humans from other species. Singer identifies some objective criteria which are frequently cited to justify human supremacy:

⁴² Singer (n 6) chapters 5 and 6.

⁴³ *Ibid* chapter 1.

⁴⁴ The failure by adjudicators, the EU and *amici* to acknowledge welfarism's consistency requirement can be contrasted with Posner's approach (discussed in the next chapter). Posner takes a more transparent by acknowledging welfarism's consistency requirement and seeking to identify arguments which discredit its validity.

“It would not be speciesist to hold that the life of a self-aware being, capable of abstract thought, of planning for the future, of complex acts of communication, and so on, is more valuable than the life of a being without these capacities”.⁴⁵

In Singer’s view, it would not be speciesist to discriminate between different animals based on criteria such as self-awareness, cognition and the capacity for future planning and complex communication. However, there are two major weaknesses which prevent these criteria from successfully justifying human supremacy.

First, these criteria fail to draw a clear line between humans and other species. Some humans fail to meet these criteria, while some animals do manage to meet them. As Singer notes, “it is clearly not the case that all humans have cognitive ability above all nonhuman animals.”⁴⁶

Singer provides examples of animals (such as primates, canines and parrots) which have superior cognitive and communication abilities than certain humans (such as babies or people with “profound mental retardation”).⁴⁷ Even if we subscribed to the use of these objective criteria, they could not justify a moral dividing line on the basis of species. Some humans would fail the test. Some animals would pass it.⁴⁸

Second, Singer highlights that these criteria are not morally relevant for justifying animal exploitation. Which criteria would be morally-relevant? Bentham addressed this matter over two centuries ago:

“But a full-grown horse or dog, is beyond comparison a more rational, as well as a more conversable animal, than an infant of a day or a week or even a

⁴⁵ Singer (n 30) 53.

⁴⁶ Peter Singer, ‘Speciesism and Moral Status’ (2009) 40(3)(4) *Metaphilosophy* 567, 570.

⁴⁷ *ibid* 567.

⁴⁸ Francione reaches a similar conclusion when he states that “the differences between humans and other animals is a matter of degree and not of kind”; Francione (n 30) 31.

month, old. But suppose the case were otherwise, what would it avail? The question is not, Can they reason? nor, Can they talk? but, Can they suffer?"⁴⁹

In Bentham's view, sentience/suffering is the morally-relevant criteria which should be used to determine which animal exploitation practices are legitimate. This is the criteria which was enshrined in the Treaty of Lisbon and it is the cornerstone of all welfarist claims made by the EU under the Suffering/Utility Standard.⁵⁰

Based on the morally-relevant criteria of suffering, welfarists reject the view that humans have a higher moral value than other species. Singer states:

"All the arguments to prove man's superiority cannot shatter this hard fact: in suffering the animals are our equals".⁵¹

Special Protection for Certain Species

In addition to rejecting human supremacy, welfarism also rejects practices which create moral hierarchies between non-human animals. They oppose special protection for certain charismatic species, such as dogs or seals. This is the type of speciesism which is most relevant to the *Seals* dispute.

This raises an interesting question for trade lawyers: under a welfarist defence, can a regulator determine that the suffering of certain species (dogs and seals) is morally more significant than the suffering of other species (cows and chickens)? The answer from welfarists is a resounding "no". Welfarism's consistency requirement denounces special protection based purely on species membership. Even if it is (descriptively) true that EU citizens care more about seals, this fact is not morally-relevant.

⁴⁹ Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation* (first published 1789, 1823 edition) chapter XVII, para 1.4 (footnote).

⁵⁰ This Standard spells out clearly the exhaustive list of morally-relevant factors which the EU can consider: (i) suffering and (ii) utility. For the purpose of defending the seal products ban, the EU focusses all of its analytical energy on the suffering criteria.

⁵¹ Peter Singer, *Animal Liberation* (1st edition, HarperCollins 1975) 12.

Adjudicators fail to acknowledge Singer's consistency requirement, but it is also embedded indirectly in the Suffering/Utility Standard. That standard provides two variables for regulators to use when determining which exploitation is moral: suffering and utility. There is no mechanism which allows other factors (such as "cuteness" or moral "intuitions") to be taken into account.⁵²

Welfarism's consistency requirement can appear counter-intuitive as it is contrary to widespread social norms. In Western societies, many people subscribe to the (incoherent) view that welfarist arguments can justify both animal exploitation on factory farms and the non-consumption of dogs.⁵³ While it should be obvious, it can be hard to realise that the special protection of dogs and seals is not welfarist in nature.

The failure of adjudicators, the EU and *amici* to acknowledge welfarism's consistency requirement could give the erroneous impression that consistency is a marginal or optional aspect of welfarist philosophy. Nothing could be further from the truth. It has been a core principle since the days of Bentham.

Speciesism is so central to the welfarist philosophy that Singer has stated:

"I use the term "speciesism" deliberately, to make a parallel with other "isms" that we are familiar with, particularly racism and sexism".⁵⁴

⁵² Such considerations would be relevant under the intuitionist approach which I analyse in the next chapter.

⁵³ In the next chapter, I will demonstrate why this view is morally-incoherent.

⁵⁴ Singer (n 46) 572. Singer elaborates that "Racists violate the principle of equality by giving greater weight to the interests of members of their own race when there is a clash between their interests and the interests of those of another race. Sexists violate the principle of equality by favoring the interests of their own sex. Similarly, speciesists allow the interests of their own species to override the greater interests of members of other species. The pattern is identical in each case". The comparison between speciesism and racism is controversial and I will avoid this debate. My key point is that welfarists view the consistency requirement as fundamental.

Singer is explicit that welfarism rejects speciesism in all its guises, including the special treatment of certain non-human animals. He highlights the moral hypocrisy of those who:

“protest about bullfighting in Spain, the eating of dogs in South Korea, or the slaughter of baby seals in Canada while continuing to eat eggs from hens who have spent their lives crammed into cages, or veal from calves who have been deprived of their mothers, their proper diet, and the freedom to lie down with their legs extended”.⁵⁵

The Suffering/Utility Standard must be applied even-handedly to all animals regardless of species membership. Singer rejects the cherrypicking of certain charismatic species (such as dogs and seals) for special protection which he describes as “moral relativism”.⁵⁶

While the term “speciesism” is not referred to in the *Seals* debate, the same concept arises in the debate on “regulatory consistency”. In chapter 12, I will discuss how adjudicators addressed the EU’s claim that it was not under an obligation to treat all animals consistently (a view which is directly contrary to welfarist philosophy).

9.2.5 Summary

The EU claims that the welfarist school of animal ethics lies at the heart of its defence. It cites welfarism directly⁵⁷ and relies on core welfarist principles indirectly when it makes arguments based on the Suffering/Utility Standard.

However, the EU’s defence deviates from welfarism in some critical ways which render its defence inconsistent with this “long-established tradition of moral thought”.

⁵⁵ Singer (n 27) 162. It is important to note that Singer is talking about the moral behaviour of individuals, but the principle is equally-applicable to governmental policy action.

⁵⁶ Like “moral aesthetics” or “moral abhorrence”, Singer does not find anything morally-valid in “moral relativism”.

⁵⁷ The *amici* are even more explicit than the EU and directly cite utilitarianism, Singer and Bentham.

The EU fails to assess “utility”, it adopts a level of animal protection which welfarists would consider too low⁵⁸ and it violates welfarism’s consistency requirement.

These glaring weaknesses are completely overlooked by adjudicators, the EU and *amici*. Is it legitimate for the EU to invent its own (non-welfarist) moral standard for animal exploitation yet still benefit from the credibility of Singer’s version? When adjudicators use the terms “welfarism”, “animal welfare” and “seal welfare”, are they referring to Singer’s or the EU’s version?

9.3 Animal Rights

9.3.1 Introduction

The EU also explicitly acknowledges animal rights as a second school of moral philosophy, but dismisses its relevance to the dispute:

“In contrast with animal welfarism, the more recent "animal rights" school of thought holds that any killing and use of animals by humans is unacceptable. This moral doctrine remains a minority view”.⁵⁹

The EU clearly has a preference for welfarism over the “more recent” and “minority” rights view. The *amici* also question the appropriateness of rights-based approaches and dismiss one of its leading proponents, Francione, as “radical”.⁶⁰

Ironically, even though the EU dismisses animal rights, this approach is better placed than welfarism to explain the abolitionist nature of the seal products ban. While species-wide bans cannot really be explained using welfarism, they might appear superficially consistent with Francione’s view that “animal exploitation should be

⁵⁸ Further, it fails to express any meaningful view on the utility of animal products (contra Posner who addresses this directly and suggests that meat provides “great utility”).

⁵⁹ EU Second Panel Submission (n 10) para 140 (footnote 157).

⁶⁰ Howse et al (n 13) 379.

abolished and not regulated”⁶¹. This might be the reason some scholars use rights-based language to describe the seal products ban. Sellheim describes the ban as granting seals a “right to life”.⁶²

Notwithstanding their abolitionist similarities, it is clear that animal rights philosophy cannot justify the seal products ban. Rights-based scholars, such as Francione and Regan, call for abolition in favour of all animals (not just certain charismatic species) and they therefore share welfarism’s opposition to “speciesism”. Rights philosophers denounce human supremacy:

“The fact that the minds of humans differ from nonhumans does not mean that the life of a human has greater moral value”.⁶³

They also oppose *a la carte* abolition (or special protection) for certain charismatic species:

“The animal rights position, as I have developed it, rejects the notion that some nonhumans, such as the nonhuman great apes, are more deserving of moral status or legal protection than are other animals because they are more “like us”.”⁶⁴

The rights position is “categorically abolitionist” and seeks to end all “animal exploitation” (not just seal exploitation).⁶⁵ Animal rights philosophers conclude that animal slaughter and exploitation cannot be justified under any circumstances:

“Although the differences between humans and animals may be important for some purposes, they are completely irrelevant to the morality of using and killing animals, even if we do so “humanely”.”⁶⁶

⁶¹ Francione (n 30) 34. Howse cites this key idea from Francione.

⁶² Nikolas Sellheim, ‘Policies and Influence: Tracing and Locating the EU Seal Products Trade Regulation’ (2015) 17 International Community Law Review 3, 30.

⁶³ Francione (n 30) 34.

⁶⁴ *ibid.*

⁶⁵ Tom Regan, *Animal Rights, Human Wrongs: An Introduction to Moral Philosophy* (Rowman & Littlefield 2003) 97.

⁶⁶ Francione (n 30) 34.

The EU and *amici* explicitly mention, and dismiss as irrelevant, rights-based philosophy. Animal rights does not play an important role in the dispute and I will not seek to analyse this philosophical tradition in detail. However, the existence of multiple schools of animal ethics raises an interesting question for WTO lawyers. What should be the standard of review for determining which approaches from moral philosophy can be validly invoked by regulators?

9.3.2 The Standard of Review

In *Seals*, WTO adjudicators accepted EU arguments based on welfarist philosophy. However, adjudicators refrained from establishing rules for determining which philosophical approaches are admissible in Article XX(a) disputes. In this section, I will analyse what the appropriate standard of review should be for evidence from moral philosophy.⁶⁷ First, I will address the question of when moral philosophy should be considered relevant under Article XX(a). Second, I will consider which philosophical approaches should be allowed.

In early disputes, moral philosophy did not arise as a central matter. In *Seals*, however, adjudicators accepted the EU's arguments from moral philosophy which suggests that a regulator can ground its defence in such arguments. If the EU had not explicitly introduced welfarist arguments into the dispute, what role would moral philosophy have played? Can complainants invoke moral philosophy to discredit a measure? Can adjudicators resort to philosophy on their own initiative.⁶⁸

The use of moral philosophy in Article XX(a) defences might be a positive development more broadly as it can shed light on whether different explanations for

⁶⁷ This matter was also briefly addressed in the previous chapter.

⁶⁸ This may be viewed by some Members as judicial overreach, though I believe it would be appropriate for adjudicators to do so where moral philosophy can help them understand and frame the key issues under dispute. In this sense, adjudicators would be using moral philosophy as a tool to assist their legal reasoning, rather than as a tool for making moral value judgments.

moral measures (including the protection of animals) are coherent or based on “standards”.⁶⁹ Adjudicators could draw on moral philosophy as an incubator where many of the key controversies about animal exploitation have been debated and analysed. Rather than starting from scratch in a WTO dispute, adjudicators could efficiently identify the most controversial questions, and the weakest arguments, by taking a quick glimpse at moral philosophy.

How should adjudicators determine which school of philosophy to use to resolve a given dispute? On this matter, Herwig argues that:

“the concept of public morals should be interpreted to refer to deontological moral standards of right and wrong – such as are generally considered to be embodied in human rights”.⁷⁰

Herwig’s approach would require WTO adjudicators to only consider deontological (rights-based) approaches under Article XX(a), such as Francione’s abolitionism. It would preclude them from relying on consequentialist approaches, such as Singer’s welfarism.⁷¹ In *Seals*, adjudicators rightly eschewed this overly-prescriptive approach by allowing the EU’s welfarist defence.

On what basis did the *Seals* adjudicators choose welfarism over animal rights? There are two possible explanations. First, the EU’s arguments show that it has a (subjective) moral preference for welfarism. Second, the EU implied that welfarism is (objectively) a superior moral approach. It described the rights approach as a “minority” view, while the *amici* dismissed it as “radical”. They imply that not only is

⁶⁹ This approach would be consistent with Posner’s view that legal reasoning has “distinct affinities with philosophical analysis”; Richard Posner, ‘Animal Rights: Legal, Philosophical, and Pragmatic Perspectives’ in Cass Sunstein and Martha Nussbaum (eds), *Animal Rights: Current Debates and New Directions* (OUP 2006) 51.

⁷⁰ Herwig (n 3) 137.

⁷¹ I disagree with Herwig’s view that the language of Article XX(a) requires a rights-based approach. After all, WTO jurisprudence on the necessity test, and especially the use of the W&B test, show a willingness by adjudicators to use consequentialist reasoning.

welfarism a legitimate approach, but that it is actually the best (or only legitimate) approach to animal ethics.

Adjudicators provide no guidance on their standard of review for moral philosophy, though it would be hard to believe they made a (normative) determination that welfarism is superior to animal rights. It would not be constructive for adjudicators to get drawn into the ideological battle between Francione and Singer.

As discussed in the previous chapter, it is more likely that the adjudicators' acceptance of the EU's welfarist claim was based on deference to the regulator. This would resemble the standard of review for science where regulators may choose their preferred approach, provided it comes from a "qualified and respected source".⁷² This standard of review seems appropriate. Herwig's strict interpretation fails to recognise that for many moral issues (including animal ethics), both consequentialist and deontological approaches might be credible.

The *amici's* attempts to discredit the animal rights movement go too far in the other direction. While welfarism may represent the leading view on animal ethics,⁷³ there can be little doubt that the rights-based approach and its leading thinkers (including Francione) are "qualified and respected". Their philosophy is peer-reviewed, widely-cited, intellectually-rigorous and morally-compelling.

If a government decided to change its policy settings to comply with the animal rights view of the world (highly unlikely in the near term), there is no reason to believe that they could not defend their minority view under Article XX(a). The EU may prefer

⁷² Appellate Body Report, *European Communities – Measures Affecting Asbestos and Asbestos-Containing Products*, WT/DS135/AB/R, adopted 5 April 2001 (*Asbestos*) para 178.

⁷³ Even Francione, a leading rights advocate, would acknowledge that welfarism is "the prevailing paradigm for thinking about our moral and legal obligations to nonhuman animals" (Francione, n 30, 24). Francione further recognises Singer as the "leading figure" in "new welfarism" (at 28).

welfarism, but surely the principle of “pluralism” would allow governments to follow Francione and Regan’s rights-based approach where that is their preference.⁷⁴ In *Seals*, adjudicators seem to adopt a deferential (and appropriate) standard for evidence from moral philosophy which replicates their approach for scientific evidence.

However, in other ways, adjudicators deviated from their approach to science. Even though welfarism plays such a central role in the dispute, the adjudicators and EU fail to cite studies or authors which support the EU’s moral credentials. The EU does not mention Singer, Bentham, utilitarianism or any philosophical literature, such as “Animal Liberation”.⁷⁵ In WTO disputes where science is a major issue, adjudicators will cite specific studies or even call up experts to provide evidence in person.

In *Seals*, this nonchalant approach to citing philosophy enabled the EU to adopt a vague and equivocal definition of what welfarism means. The EU cites the Suffering/Utility Standard as a generic concept which emanates from welfarism, but overlooks other key features such as the level of protection, consistency requirement and utility. Adjudicators fail to notice these significant gaps which undermine the coherence and consistency of the EU’s moral arguments.

This undermines the ability of adjudicators to coherently take advantage of welfarism’s moral reasoning. They accept the EU’s Suffering/Utility Standard but, if they had been more rigorous (by citing specific authors), they would have realised that the seal products ban cannot be justified by moral reasoning. It is an orphan which would be rejected by both welfarist and rights philosophers. This should have

⁷⁴ Indeed, under the Category/Content Distinction (discussed in the previous chapter), the *amici* would allow governments to adopt abolitionist approaches. If animal exploitation is a moral “category” (which is a non-controversial proposition), WTO law should not interfere with the right of governments to choose their preferred normative “content” (including abolitionist approaches).

⁷⁵ This can be contrasted with the *amici* who cite each of these authors, movements and philosophical works.

raised red flags that the EU's claim was based on logical inconsistencies worthy of deeper analysis.⁷⁶

9.3.3 Summary

The rights-based approach to animal ethics has several strengths. As a general defence under Article XX(a), it is clearly based on moral standards and finds support from a credible movement in moral philosophy (albeit a minority one). To explain the abolitionist elements of the seal products ban, the rights-based approach is arguably better placed than welfarism. However, the EU was (strategically) correct to avoid a rights-based defence as its animal exploitation practices are speciesist and inconsistent with Francione's categorical abolitionism.⁷⁷

In terms of the standard of review, adjudicators provide little guidance. They accept arguments from moral philosophy (in general) and welfarism (specifically), but the legal basis is unclear. In my view, they adopted a deferential standard whereby the regulator's preference is accepted, provided it is "qualified and respected". Strangely, they managed to accept evidence from moral philosophy without citing a single author or study.

9.4 Conclusion

In *Seals*, adjudicators recognised moral philosophy as a type of evidence which can be useful for defending a measure under Article XX(a), especially for the purpose of demonstrating "moral standards". While two different approaches to animal ethics were raised in the literature, adjudicators affirmed the right of the EU to rely on a welfarist approach.

⁷⁶ These inconsistencies will be further explored in chapter 11.

⁷⁷ In theory, the speciesist nature of the EU's ban should be equally problematic for its welfarist defence, but adjudicators, the EU and *amici* somehow manage to ignore welfarism's consistency requirement.

Even though the EU's welfarist argument is consistent with the Suffering/Utility Standard, it also suffers from some major weaknesses. Singer would denounce the inconsistent treatment of seals relative to factory farmed animals (as this is speciesist). He would further argue that the EU's level of animal protection is far too low.

Despite these major weaknesses, adjudicators accept the EU's right to selectively draw on the credibility of Singer and welfarism. They ignore those aspects of the EU's measure which are clearly anathema to welfarist philosophy. In order to prevent confusion, should the EU be required to renounce the welfarist branding of its measure and offer an explanation which actually explains the special protection it affords to seals? The next chapter will consider what that explanation would look like.

CHAPTER 10: INTUITIONISM

10.1 Introduction

The previous chapter described some of the major weaknesses in the EU's welfarist justification for the seal products ban. This chapter will address a second formulation of the EU's defence: that it is legitimate to offer special protection to certain species based on "intuition". This section will describe intuitionism and assess its legitimacy in moral philosophy and WTO law.

The first section will focus on intuitionist arguments as they appear in the *Seals* dispute and surrounding literature. The EU and *amici* make implicit arguments in favour of intuitionism, while Perišin and Sellheim explicitly oppose it.

I will also consider the broader animal rights debate where there has been fierce disagreement between Posner (a proponent)¹ and Singer (an opponent)² about whether "intuition" is a valid basis for moral decisions. I will summarise the debate between these two intellectual heavyweights which elaborates on many of the key issues which were addressed in *Seals*.

I consider whether Article XX(a) allows governments to introduce trade-restrictive measures on the basis of "intuition". I suggest that current jurisprudence does not allow such restrictions, primarily because of the "moral standards" requirement. The academic approach of "evidentiary unilateralism" would permit intuitionist measures, but this would require jurisprudential reform, as discussed in chapter 8.

¹ Richard Posner, 'Animal Rights: Legal, Philosophical, and Pragmatic Perspectives' in Cass Sunstein and Martha Nussbaum (eds), *Animal Rights: Current Debates and New Directions* (OUP 2006).

² Peter Singer, 'Ethics beyond Species and beyond Instincts: A Response to Richard Posner' in Cass Sunstein and Martha Nussbaum (eds), *Animal Rights: Current Debates and New Directions* (OUP 2006).

Finally, I use deconstruction to question whether Posner's position is truly intuitionist. In practice, he seems to rely (at least partly) on moral reasoning which resembles the Suffering/Utility Standard.

10.2 Intuitionism and the Seals Dispute

10.2.1 EU/Amici Arguments for Intuition

The EU subscribes to intuitionism when it argues that "humans do not regard all animals as equal from a moral point of view".³ These types of speciesist arguments would be directly contrary to welfarism.⁴

The EU cites a study which formalises the intuitionist approach into a "socio-zoological scale" where:

"people rate animals as morally more or less important, and therefore more or less worth protecting, according to a number of factors. These include how useful the animal is, how closely one collaborates with the individual animal, how cute and cuddly the animal is, how harmful the animal can be, and how 'demonic' it is perceived to be".⁵

This suggests there are intuitionist criteria (such as cuteness and cuddliness) which can establish the higher moral worth of certain species. These criteria appear highly subjective and non-scientific. They would be rejected by welfarists as morally-irrelevant as they are not based on sentience.

To illustrate the intuitionist approach, the EU provides an example:

"In some Asian countries dogs are killed for food and fur, a practice which is perceived as morally abhorrent by the EU public".⁶

³ EU, 'First Written Submission by the European Union' (*EC – Seals*, 21 December 2012) <https://trade.ec.europa.eu/doclib/docs/2012/december/tradoc_150190.pdf> (consulted 5 April 2020) ("EU First Panel Submission") para 73.

⁴ Especially the consistency requirement as described in chapter 9.

⁵ EU First Panel Submission (n 3) para 73 (citing Peter Sandøe and Stine Christiansen, *Ethics of Animal Use*, 2008).

⁶ ibid para 74.

The EU public's intuitionist conclusion that dog exploitation is "morally abhorrent" cannot be justified by welfarism's Suffering/Utility Standard. Such bans are not sentience-based as they deem exploitation to be "abhorrent" even where it is pain-free.

Unlike the EU's welfarist claims, its intuitionist arguments do not draw on moral philosophy. The EU briefly refers to the "socio-zoological scale" (which makes descriptive claims), but provides no theoretical support for the notion that intuitions are a valid reason for normative action.⁷

This is unfortunate as intuition has been extensively addressed in a number of fields, including philosophy, psychology and law.⁸ In fact, the EU's arguments based on "intuition" and "sentiment" have a significant academic pedigree. The EU's precise line of argument was eloquently articulated by Posner a decade prior to the *Seals* dispute.

There are some allusions to intuitionism in the *Seals* scholarship. Herwig describes it as an approach "in which feelings become the decisive element in grounding our moral judgments".⁹ Further, the *amici* make intuitionist arguments when they refer to

⁷ While they are not rationalist in nature, this does not mean that intuitionist arguments cannot influence judges. These intuitionist arguments have the potential to be highly influential if they manage to bypass System 2 and speak directly to the judge's System 1.

⁸ While there are schools of moral philosophy which endorse intuitionism as a guide to moral action, I will not address this philosophical approach as it does not appear in the *Seals* dispute or academic literature. This chapter will address the opposition to intuitionism from rationalist moral philosophy as this is an important issue in *Seals*.

⁹ Alexia Herwig, 'Regulation of Seal Animal Welfare Risk, Public Morals and Inuit Culture under WTO Law: Between Techne, Oikos and Praxis' (2015) 6 European Journal of Risk Regulation 382, 387.

“widely held moral intuitions among Europeans about what is cruel”¹⁰ and “moral aesthetics”.¹¹

10.2.2 Perišin/Sellheim Opposition to Intuition

The EU’s intuitionist claim has received far less academic attention than welfarism, possibly because the EU’s intuitionist claims are under the surface.¹² However, a handful of scholars, including Perišin and Sellheim, deny that the seal products ban is welfarist. They recognise intuitionism as the true claim and attack its weaknesses.

Perišin explicitly addresses intuitionism when she notes that many people are “intuitively biased” in favour of seals because of their “cuteness”.¹³ She wonders what the “real aims” of the measure are¹⁴ and suggests that it is based on “the public’s emotional attachment to seals” rather than “rational grounds”.¹⁵

She questions the welfarist defence by noting that the seal products ban “does not seem to be based on any rational differentiation between seals and other animals”.¹⁶ In this sense, she follows a true welfarist approach which would only allow discrimination based on morally-relevant criteria and rationalist reasoning.¹⁷

Perišin notes that the EU’s sentience-based arguments are only used selectively:

¹⁰ Robert Howse and Joanna Langille, ‘Permitting Pluralism: The Seal Products Dispute and Why the WTO Should Accept Trade Restrictions Justified by Noninstrumental Moral Values’ (2012) 37(2) *Yale Journal of International Law* 367, 388-9.

¹¹ *ibid* 372 (footnote 25).

¹² In order to critique the intuitionist claim, a scholar must undertake a deconstruction of the EU’s substantive claims to uncover the intuitionist nature of some aspects of its defence. While I am not the first author to recognise this fact, I will try to introduce a more detailed and explicit structure for understanding the intuitionist claim. I use Posner as a proxy for the EU’s claims as he explicitly articulates the key elements of an intuitionist defence.

¹³ Tamara Perišin, ‘Is the EU Seal Products Regulation a Sealed Deal? EU and WTO Challenges’ (2013) 62(2) *International & Comparative Law Quarterly* 373, 375.

¹⁴ *ibid* 400.

¹⁵ *ibid* 395.

¹⁶ *ibid* 395.

¹⁷ Perišin therefore adopts an approach to welfarism which is faithful to Singer and Bentham as she rejects speciesism and calls for consistency.

“The EU measures have a significant problem of coherence in that the basis for differentiating between permissible and impermissible products is not humane hunting”.¹⁸

After dismissing welfarism, Perišin identifies the EU’s true rationale for the ban:

“the public seems to be concerned for seals’ welfare because they are more attractive (‘cuter’) to the public than many other species”.¹⁹

She further explains why commentators are unwilling to criticise the seal products ban or expose its intuitionist underpinnings:

“Raising arguments against the Regulations is, however, not popular. Discussions reveal that many feel emotionally touched by the ‘cuteness’ of seals and that this makes them intuitively biased in supporting the EU measures. Setting aside these emotional attitudes towards seals, it is difficult to pin down what the rational basis behind special protection of seals (in contrast to that of other animals) is”.²⁰

Sellheim is equally dismissive of the EU’s welfarist claim which stands on “shaky grounds”.²¹ He dismisses the relevance of welfarism:

“Here lies the primary difference between the European moral standard pertaining to animal welfare and the EU Seal Regime... it decouples the seal from welfare frameworks by ascribing it a right to life”.²²

Sellheim argues that “the Seal Regime has no further implications for the wider moral standard of animal welfare” as it is based on the notion that “one species is singled out for total protection”.²³ Seals only receive special protection because they are “cute” with “big eyes” and a “friendly face”: they are “charismatic megafauna”.²⁴

¹⁸ Perišin (n 13) 404.

¹⁹ ibid 395.

²⁰ ibid 375.

²¹ Nikolas Sellheim, ‘The Legal Question of Morality: Seal Hunting and the European Moral Standard’ (2016) 25(2) Social and Legal Studies 141, 147-8.

²² ibid 156. Sellheim uses philosophical reasoning to recognize that the seal products ban resembles a rights-based rather than a welfarist measure.

²³ ibid 153.

²⁴ ibid 141-2.

He argues that the seal product ban was motivated by “evocative symbolism like red blood on white ice”,²⁵ rather than “objective, scientific knowledge”.²⁶

While Perišin and Sellheim both accept that animal welfare is a legitimate objective under WTO law, they reject the notion that the seal products ban can be characterised as a welfarist measure. After all, the ban is abolitionist, speciesist and based on subjective perceptions rather than sentience. Perišin and Sellheim apply Singer’s conception of welfarism and deem the seal products ban unjustifiable.

10.3 Philosophy: The Singer/Posner Debate

There is a dividing line in the *Seals* dispute between those who imply that intuitionism is legitimate (EU/*amicus*) and those who state explicitly that it is illegitimate (Perišin/Sellheim). This replicates a broader trend in the animal rights literature, including the debate between Singer and Posner.²⁷

The Singer/Posner debate does not directly address the seal products ban.²⁸ Rather, they engage in a wide-ranging debate on some of the major issues in animal ethics, including whether (and how) humans can justify full protection for some animals (such as dogs) and the painful exploitation of other animals on factory farms. The Singer/Posner debate provides some lessons which are highly relevant to *Seals*. I will summarise Posner’s argument followed by Singer’s retort.

²⁵ Nikolas Sellheim, ‘Policies and Influence: Tracing and Locating the EU Seal Products Trade Regulation’ (2015) 17 International Community Law Review 3, 28.

²⁶ Sellheim (n 21) 150.

²⁷ The Singer/Posner debate took place in a popular context (in Slate Magazine) and continued in an academic context where each addresses the other’s arguments in a chapter of Cass Sunstein and Martha Nussbaum (eds), *Animal Rights: Current Debates and New Directions* (OUP 2006). My analysis draws from their academic debate.

²⁸ The Singer/Posner debate pre-dates the seal products ban, though I have also cited some other literature from Singer where he explicitly questions the moral validity of offering special protection to *Seals*.

10.3.1 Posner's Defence of Intuitionism

Posner's analysis is a major contribution to the literature. While intuition plays an important role in reasoning about animal exploitation, it is rare to see it acknowledged openly. Most proponents, including the EU and *amici*, prefer to rely on intuitionism indirectly (by allusion only).

Posner, however, is not afraid to acknowledge intuitionism as the cornerstone of his thinking about animals. He openly states that his normative framework for animal exploitation is based on nothing more than a "purely sentimental attachment to animals".²⁹ He articulates a view which seems widespread, but rarely articulated. Thanks to Posner's openness about his true reasons, it is possible to study and critique the merits of intuitionism.

10.3.1.1 Intuition Trumps Logical Argument

While Posner claims to pursue a "legal, philosophical, and pragmatic" perspective on animal ethics,³⁰ his focus is ultimately pragmatic as he "eschew[s] on the one hand philosophical argument and a legal-formalist approach on the other".³¹ His pragmatic approach is based on "reasoning methods we use in dealing with practical problems outside of law",³² especially "intuition", "sentiment" and "instinct".³³

Posner prefers "intuition" over "logical argument":

"We realize that animals feel pain, and we think that to inflict pain without a reason is bad. Nothing of practical value is added by dressing up this intuition in the language of philosophy; much is lost when the intuition is made a stage in a logical argument".³⁴

²⁹ Posner (n 1) 62.

³⁰ ibid 50.

³¹ ibid 51.

³² ibid 58.

³³ Posner uses these terms interchangeably.

³⁴ Posner (n 1) 69.

Further, where there is conflict, Posner believes that intuition should prevail:

“ethical argument is and should be powerless against tenacious moral instincts”.³⁵

In addition to being descriptive, Posner’s claim is normative: instinct “should” trump ethics. In this sense, his framework implies that the special protection of seals for intuitionist reasons is normatively justifiable. What is Posner’s view when intuition leads to practices which are incoherent or contradictory?

10.3.1.2 Intuition Trumps Consistency/Cohherence

Posner’s intuition-based argument serves a critical role in his debate with Singer: it allows him to reject welfarism’s consistency requirement.³⁶ Posner explicitly makes the case that Singer’s consistency requirement is so contrary to human intuition that it cannot possibly be valid. Posner endorses the speciesist doctrine of “human supremacy” when he argues we must “maintain a bright line between animals and human beings”.³⁷ He further adds:

“the superior claim of the human infant than of the dog on our consideration is a moral intuition deeper than any reason that could be given for it and impervious to any reason that anyone could give against it”.³⁸

Posner implies that even welfarists would struggle to resist their natural intuition to prioritise humans over non-human animals:

“Philosophers who embrace weird ethical theories do not act on those theories even when they could do so without being punished”.³⁹

Posner appears to doubt that, in “his heart of hearts”, Singer would resolve a real world dilemma by prioritising a dog over a human baby.⁴⁰ He implies that Singer’s “human supremacy” instinct would kick in and he would prioritise the baby.

³⁵ ibid 66.

³⁶ This is the same role that intuitionism serves for the EU/*amici*.

³⁷ Posner (n 1) 61.

³⁸ ibid 65.

³⁹ ibid.

If you believe that intuitionist arguments are valid, Posner makes a powerful case. However, human supremacy is not the only type of speciesism which welfarists decry. Indeed, in *Seals*, the EU's intuitionist arguments relate to the special protection of seals as a charismatic species (a more complex moral issue). What are Posner's "intuitive" views on moral hierarchies between animals? Do humans possess an intuition that dogs (and seals) matter more than cows? If so, can this intuition justify special protection?

Posner addresses moral hierarchies between animals, but his analysis is incomplete. He makes the descriptive claim that humans seem to care more about certain animals. He notes the human "affection for certain 'cuter' animals"⁴¹ and observes that humans have "greater affection for our pets than for our simian first cousins".⁴²

However, he does not vigorously transition to the normative claim that we should care more about "cuter" animals. If he believes that dogs have a higher moral status than other animals, he does not provide a compelling or explicit normative justification for this view. This is in stark contrast to Posner's views on human supremacy where he makes both a descriptive claim (humans care more about other humans) and a normative claim (that humans matter more than other animals).

On moral hierarchies between animals, Posner avoids his earlier technique of appealing to his reader's "tenacious moral instinct". Posner may have an instinct that dogs should not suffer, but he does not rely on instinctive arguments to justify the suffering which occurs on factory farms. Does Posner possess a "tenacious

⁴⁰ ibid 64.

⁴¹ ibid 63. This resembles the EU's descriptive claim that its citizens care more about seals, including due to their cuteness.

⁴² ibid.

moral instinct” that animals who are not cute should suffer? If he does, I doubt it is widely-shared.

It is more likely that Posner lacks tenacious moral instincts in favour of commodity animals and that he is therefore relatively indifferent about changing the status quo (painful exploitation on factory farms).⁴³ However, Posner does not offer a justification for the status quo; he simply takes it for granted. This leaves a gap in his analysis: what is his justification for the suffering of commodity animals on factory farms?⁴⁴

The notion that “cuter” animals deserve special protection is philosophically weak and would be summarily dismissed by both welfarist and rights philosophers. When our intuitions and moral reasoning are in conflict, Posner suggests that “tenacious moral instincts” should prevail. What about all those moral matters where we have “weak” moral instincts? Should we follow our instinct or defer to our human capacity for reason? Posner offers no guidance on such problems even though they are the most complex and philosophically-important ones.

10.3.1.3 Pure Intuitionism - Dog Product Bans

When seeking to protect cute animals, such as dogs, Posner rejects rationalism and suggests our moral approach is based purely on intuition:

“It is possible... to have a purely sentimental attachment to animals—to like them, or some species at any rate... or if not to “like” them, to sympathize with them sufficiently to feel their pain and want to alleviate it”.⁴⁵

⁴³ Perhaps Posner is suffering from a “status quo” bias and believes that no justification is required for the maintenance of the status quo. Rather, the onus is on those seeking to change the status quo to produce compelling reasons. Moral philosophers would be appalled by such status quo thinking. As Hume famously argued: “you cannot deduce an ought from an is”.

⁴⁴ I will return to this question in section 10.4.

⁴⁵ Posner (n 1) 62.

Posner's approach is "humancentric". He suggests that the desire to protect dogs (or other favoured species) is not because they have moral standing,⁴⁶ but rather because we have a sentimental desire to protect them. Although Posner alludes to alleviating pain, the human sentiment he describes is not really sentience-based as it can be engaged (at least for "cute" animals) even if there is no pain.

For charismatic species, our desire to protect need not be linked to cruelty. This can be demonstrated by the EU ban on "morally abhorrent" dog products. Europeans reject these products even if the dogs are exploited in a pain-free manner. The moral norm in the EU that dogs should not be exploited or consumed is not based on "suffering". It is based on "love", "cuteness" and "affection".⁴⁷

This leads to a strange inversion between Posner and Singer: while Singer is the far more strident defender of animals in general, Posner is a more strident defender of dogs. Posner endorses a categorical ban on dog exploitation (motivated by intuition), even where it is cruelty-free. Singer would allow dog exploitation in situations where it is cruelty-free. As Posner states:

"Singer, a strong proponent of vegetarianism, acknowledges that 'as a matter of strict logic... One could consistently eat animals who had lived free of all suffering and been instantly, painlessly slaughtered'."⁴⁸

Singer refrains from promoting strict abolition and, according to his consistency requirement, this would mean that some dog exploitation could be morally-acceptable. By contrast, Posner's philosophy allows for total protection of charismatic species. His approach could justify the EU's total ban on dog or seal products, even where no suffering is involved.

⁴⁶ For an analysis of animal ethics and moral standing, see James Rachels, 'Drawing Lines' in Cass Sunstein and Martha Nussbaum (eds), *Animal Rights: Current Debates and New Directions* (OUP 2006).

⁴⁷ Where dog exploitation also involves cruelty, this would also be "morally abhorrent" to European sensibilities, but cruel treatment is not the ultimate basis of the norm.

⁴⁸ Posner (n 1) 60.

10.3.1.4 Summary

Posner relies on intuition for the purpose of countering welfarism's consistency requirement. He makes this argument forcefully and in an intuitively-appealing way when suggesting that humans matter more than other animals. Posner further applies this argument to the special protection of "cuter" animals (the relevant type of speciesism in *Seals*).⁴⁹ However, when justifying discrimination between non-human animals, Posner fails to offer a normative justification or demonstrate the existence of "tenacious moral instincts".

On the matter of intuition, Posner makes strikingly similar arguments to the EU. They both suggest that intuitions are a valid basis for offering special treatment to "cuter" animals. In my view, these intuitionist arguments serve the same analytical purpose for both Posner and the EU: to overcome Singer's consistency requirement. The only meaningful difference is that Posner confronted Singer's consistency requirement directly, while the EU simply ignored it.⁵⁰

10.3.2 Welfarism's Rejection of Intuitionism

The previous chapter explained why welfarists reject intuition as morally-irrelevant. In his response to Posner, Singer explains why intuitionism can lead us astray and is incapable of serving as a valid basis for moral decisions:

"Posner would still have to explain why he thinks ethical argument... not only is, but should be powerless to change any tenacious instinct, no matter how aggressive, murderous, or xenophobic that instinct may be".⁵¹

Singer also makes this point in the specific context of seal exploitation when he opposes the moral hypocrisy of those who consume factory farm products, but

⁴⁹ We need not focus on his stronger arguments in favour of "human supremacy" to understand the *Seals* dispute.

⁵⁰ The EU's approach to welfarism's consistency requirement was addressed in the previous chapter.

⁵¹ Singer (n 2) 84.

“protest about bullfighting in Spain, the eating of dogs in South Korea, or the slaughter of baby seals in Canada”.⁵²

It is not surprising that welfarists explicitly reject intuition-based arguments, as they pose a direct threat to rationalist philosophy in general and to certain core tenets of welfarism in particular. Intuition-based arguments seek to circumvent welfarism’s consistency requirement and preference for morally-relevant criteria. When Posner and the EU argue that it is legitimate to provide special protection to those animals we “love”, they assert a right to engage in speciesist practices. This is morally repugnant to welfarists.

10.4 Is Posner a “Pure Intuitionist”?

In this section, I will deconstruct Posner’s approach to show that he is not really a pure intuitionist. Like the EU, Posner combines his intuitionist arguments with consequentialist moral reasoning based on the Suffering/Utility Standard. This section will reveal Posner’s true position which is strikingly similar to the EU’s defence in *Seals*.⁵³

10.4.1 Suffering Utility Standard

Posner claims to rely exclusively on intuition and to reject “conventional philosophizing”.⁵⁴ He suggests that “moral reasoning” has “severe limitations”⁵⁵ and states:

⁵² Peter Singer, *Animal Liberation* (3rd edition, HarperCollins 2002) 162. It’s important to note that Singer is talking about the moral behaviour of individuals, but the principles are equally-applicable to the moral decisions of governments.

⁵³ I will demonstrate that the substantive position of Posner and the EU is functionally the same, even though they frame their arguments in a very different way.

⁵⁴ Posner (n 1) 59.

⁵⁵ ibid 58.

“We should be able to agree without help from philosophers and constitutional theorists that gratuitous cruelty is bad”.⁵⁶

Despite Posner’s insinuation to the contrary (and use of layperson’s language), the notion that “gratuitous cruelty is bad” is a philosophical principle (not a moral instinct). More precisely, it is a consequentialist moral principle which appears in the writings of Singer, the EU and the *amici*. Posner’s moral principle is a mere reformulation of the Suffering/Utility Standard.

Posner implicitly acknowledges his use of consequentialism when he claims his approach:

“focuses on the consequences for us of recognizing animal rights. Those consequences are both good (benefits) and bad (costs—a word I am using broadly without limitation to pecuniary costs)”.⁵⁷

Posner further argues that:

“few of us are either so sadistic, or so indifferent to animal suffering, that we are unwilling to incur at least modest costs to prevent gratuitous cruelty to animals”.⁵⁸

In my earlier analysis, I noted a gap in Posner’s intuitionist analysis regarding the exploitation of commodity animals. How does Posner justify the painful exploitation of factory farmed animals? If Posner was a pure intuitionist, he could argue that he has a “tenacious moral instinct” that cows and chickens should suffer or, at least, that their suffering is acceptable. However, on this matter, he eschews intuitionist arguments in favour of moral reasoning.

He makes rationalist (and consequentialist) arguments about utility and suffering. Animal suffering should be condemned when it is “gratuitous”, but it is acceptable where it produces utility. Posner becomes a utilitarian when he argues that “most

⁵⁶ ibid 59.

⁵⁷ ibid 70.

⁵⁸ ibid 59.

people obtain great utility from eating meat” and that animal suffering for this purpose is therefore not “gratuitous”.⁵⁹

Singer rightly rejects Posner’s claim that his approach is free from philosophy.⁶⁰

While Posner’s justification of human supremacy and dog protection may rely on pure intuitionism (and therefore be philosophy-free), his justification for the painful exploitation of commodity animals relies on moral reasoning. Posner draws on Singer’s Suffering/Utility Standard except he adopts a much lower level of protection and discards the consistency requirement. In other words, he adopts the same hybrid position as the EU.

10.4.2 Posner’s Unified Theory

If Posner’s sole aim was to justify the special protection of dogs, he could have relied on pure intuitionism. However, he sets himself the ambitious goal of developing a “unified theory” which can explain animal ethics in general, from the special protection of dogs to the immense suffering of factory farmed animals.

Despite centuries of effort, moral philosophers have failed to develop a unified theory. Since moral philosophers reason from first principles, they are able to develop logically-coherent moral positions, but this leads to other “pragmatic” problems. Their normative recommendations may require serious changes to the status quo, such as Singer’s call to end all factory farming.⁶¹

⁵⁹ ibid 60.

⁶⁰ Singer states: “It is no more possible to reject the value of philosophical argument without taking a philosophical position than it is to win public office without being involved in politics”; Singer (n 2) 80.

⁶¹ Francione’s view that we should abolish all animal exploitation is even more extreme.

Further, philosophers may reach conclusions which are highly counter-intuitive, such as Singer's view that pig suffering is as morally significant as human suffering.⁶²

Philosophical approaches, from both consequentialist and deontological traditions, fail to offer a coherent explanation for the special protection of dogs and the factory farming of other animals.

Posner's approach avoids the shortcomings of philosophy. His pragmatism promotes moral conclusions which closely resemble the status quo on animal exploitation. His intuitionist approach enables him to avoid counter-intuitive conclusions.

However, the disadvantage of Posner's approach is that it lacks the logical rigour of philosophy. Since Posner eschews "logical argument", he cannot achieve the level of coherence/consistency which is available to philosophers like Singer. This is the price that intuitionist pragmatists must pay: to avoid the counter-intuitive and undesirable results of pure philosophy, they must take shortcuts in their logical reasoning.

However, Posner refuses to concede this weakness. He insists that he has found a coherent approach which can explain both dog product bans and factory farming:

"The humancentric approach establishes continuity between the animal rights movement and the love of pets, both seen as founded on sentiment rather than on philosophical idealism".⁶³

If Posner's claim is true, it would be a major achievement. However, even though Posner refuses to admit it, his approach fails to develop a coherent unified theory of

⁶² Sunstein is aware that utilitarianism can produce strange results in marginal cases. He notes that "utilitarianism, taken seriously, produces serious mistakes in some cases. In this view, utilitarianism is itself a heuristic, one that usually works well but leads to systematic errors". For this reason, Sunstein identifies with "weak consequentialism"; Cass Sunstein, 'Moral heuristics' (2005) 28 Behavioural and Brain Sciences 531, 534.

⁶³ Posner (n 1) 72.

animal exploitation. Rather, he has a twin-pronged approach based on intuitionist protection for charismatic species and moral reasoning to justify the exploitation of commodity species. Posner's hybrid approach (which resembles the EU's hybrid approach) leads to some serious logical inconsistencies.⁶⁴

At the outset, Posner's hybrid approach divides animals into two categories. Charismatic species (such as dogs and seals) are assessed under the intuitionist claim where he finds that their special (or even total) protection is justifiable.

By contrast, commodity animals are assessed under the Suffering/Utility Standard where Posner engages in consequentialist moral reasoning to determine which practices are legitimate. Commodity animals are entitled to protection from "gratuitous cruelty" but not to the special protection available to charismatic species. This raises an important question: what is Posner's method for determining which animals are charismatic and which ones are mere commodities? In the West, it may be obvious (based on centuries of tradition) that dogs are charismatic and that cows are commodities,⁶⁵ but are there hard cases (such as seals)?

It seems odd that Posner would promote dog product bans when his philosophy is based on intuition. Why not trust each individual to follow their own "tenacious moral instinct" about whether or not humans should consume dog products? Why should governments regulation interfere with the moral instinct of individuals who wish to eat dogs? Perhaps the norm against dog products is so strong in Western societies that Posner failed to perceive legislative interventions as an interference with moral instincts.

⁶⁴ These logical inconsistencies will be explored in detail in the next chapter.

⁶⁵ Perhaps, these categories were so intuitively obvious to Posner that he was not even consciously aware that he was engaged in a categorisation exercise.

However, there may be harder cases, such as seals, where norms are less widespread and are not backed by centuries of tradition. Until a few decades ago, the EU had no mechanisms in place for the protection of seals and it was among the world's largest market for seal products.⁶⁶ Over a relatively-short time period, the EU introduced rules to ban seal hunting and the sale of seal products.

Seals (somehow) managed to transition from commodity species to charismatic species. However, the status of seals remains contested by stakeholders who question the special protection of seals (unlike dog protection norms which appear near-universal).⁶⁷

Under Posner's twin pronged approach, are seals "cute" and "loved" (and granted intuitionist protection) or are they a commodity species (subject to exploitation under the Suffering/Utility Standard)? If seals made the successful transition to charismatic species, when and how did this happen? Is it legitimate for governments to ban seal products contrary to the "tenacious moral instincts" of many citizens within their societies?

These are thorny and important questions which Posner overlooks. His initial categorisation, whether a species is charismatic or a commodity, is the key factor which determines the fate of that animal. If Posner decides that seals are charismatic, they are entitled to full protection. If he decides that they are commodities, they can be exploited, provided there is no "gratuitous cruelty".

⁶⁶ As the Appellate Body notes (citing an EU argument): "the EU seal product market has traditionally 'occupied a central position within the global market"'; Appellate Body Reports, *European Communities – Measures Prohibiting the Importation and Marketing of Seal Products*, WT/DS400/AB/R / WT/DS401/AB/R, adopted 18 June 2014 (*Seals*) 5.245.

⁶⁷ For a discussion of the European public opinion polls on seal exploitation which were used in the *Seals* dispute (including their findings and main methodological problems), see Sellheim (n 21). Sellheim rejects the notion that Europeans universally oppose seal exploitation, but he recognises that "it is reasonable to conclude that the European public's support of a ban on trade in seal products numerically outweighs its opposition" (at 152).

However, Posner offers no explanation for how he categorises different species. Perhaps he is arguing in bad faith. Perhaps he truly believes in his “unified theory” and is completely oblivious to the (invisible) categorisation he undertakes before he starts reasoning.

10.5 WTO Jurisprudence

10.5.1 Does Article XX(a) Allow Posner’s “Intuitionist” Measures?

If we ignore the fact that Posner uses moral reasoning, and focus exclusively on his self-proclaimed “pragmatic” and intuition-based approach, he seems to offer a theoretical justification for the EU’s special protection of seals. Posner is a giant in the world of legal philosophy whose arguments could have provided tremendous support for the EU’s defence.⁶⁸ Why did the EU fail to cite Posner in support of an explicit intuitionist defence?⁶⁹

Even though Posner’s argument is substantively the same as the EU’s, he frames it in a manner which sounds inconsistent with WTO law. Posner’s framing rejects Singer’s welfarism (unlike the EU which claims to be welfarist) and openly acknowledges that intuitionism is not based on moral standards.

Intuitionism is Inconsistent with Welfarism

In direct contrast to the EU, Posner claims that a major advantage of his intuitionist approach is that it is not based on welfarism or rationalist philosophy. He directly

⁶⁸ Indeed, Posner had such confidence in his argument about animal ethics that he was prepared to go head-to-head with Peter Singer.

⁶⁹ This would have been a more powerful theoretical argument than the EU’s use of the “socio-zoological scale” (which seems to have been accorded no weight by adjudicators). It is also curious why the *amici* refrained from citing Posner. One explanation is that the EU and *amici* might have been unaware of Posner’s work. An alternative explanation, which I will describe below, is that Posner’s explicitly-intuitionist framing may have been a poor strategic choice for the EU.

rejects the philosophy of Singer and engages in a heated dispute suggesting that his non-welfarist approach is superior.

Even though the EU's substantive intuitionist claim finds direct expression in the articulate writings of Posner, the EU had good strategic reasons to avoid citing Posner. If the EU aligned its arguments with Posner, its non-compliance with Singer's philosophy would be immediately visible. In practice, the EU managed a seemingly-impossible balancing act by combining Posner's substantive arguments with Singer's credibility on animal ethics.

Intuitionism is not based on Moral Standards

A further problem with intuitionism, at least under WTO law, is that it is not based on "moral standards". In *Seals*, Canada suggested that any social norm based on intuition "is open-ended, and therefore not a standard at all".⁷⁰ It described the EU's "so-called moral standard" as "inherently subjective" and "arbitrary".⁷¹

Iceland agreed that intuitions are not standards-based:

"the Panel failed to explain what 'standard of right and wrong conduct' exists in the European Union that makes selling seal products a public morals concern".⁷²

Since Posner is writing about animal ethics in general (rather than WTO law), he is not afraid to admit that he eschews moral standards. In fact, he seems proud of his decision to reject the use of "ethical" and "logical" reasoning and philosophical "first principles".⁷³ Unlike Posner, the EU was constrained by the Article XX(a) jurisprudence. If the EU had sought to align with Posner, this would have amounted

⁷⁰ Appellate Body Reports, *Seals* (n 65) para 2.195.

⁷¹ *ibid*.

⁷² *ibid* 2.260.

⁷³ Posner (n 1) 59.

to an admission that its measure was not based on “moral standards”. Once again, the EU performed a seemingly-impossible task by advocating Posner’s substantive argument, but claiming it was based on moral standards.

10.5.2 Jurisprudential Reform

In order for Posner’s intuitionism to ground a defence under WTO law, adjudicators would have to drop the “moral standards” requirement and amend the public morals test to introduce evidentiary unilateralism. In practice, this would reduce the legal test to a descriptive “public concerns” requirement which relies on unilateral evidence to determine the values in the regulator’s society.⁷⁴

An alternative (and more nuanced) approach would be for adjudicators to retain the “moral standards” requirement for instrumental claims, but adopt evidentiary unilateralism for non-instrumental claims (such as dog and seal product bans).

In order to create a typology of different moral norms, adjudicators would have had to reform the Article XX(a) jurisprudence, including by establishing a legal test for distinguishing between instrumental and non-instrumental moral norms. As discussed in chapter 8, the *Seals* adjudicators did not introduce any jurisprudential reforms along these lines.

The continued use of the “moral standards” requirement suggests that Members are precluded from adopting intuitionist measures. If this is true, it is doubtful that dog and seal product bans are legitimate under WTO law,⁷⁵ but this did not prevent the *Seals* adjudicators from finding that the seal products ban was based on moral

⁷⁴ This would render other types of external standard (such as transnationalism, internationalism and moral philosophy) outside the scope of analysis.

⁷⁵ In my view, the current public morals jurisprudence does not permit the defence of intuitionist measures such as dog product bans. This does not mean that Members must proactively remove their dog product bans as there may be compelling diplomatic reasons why such measures are unlikely to be challenged.

standards. How did they achieve this result without accepting evidentiary unilateralism or intuitionist defences? The next two chapters will analyse some of the fallacious reasoning which enabled this outcome.

10.6 Conclusion

Posner's "intuitionist" approach to animal ethics is an incomplete philosophy. Even if it is capable of justifying special protection for dogs and seals, it cannot justify the cruel exploitation of commodity species. When Posner seeks to justify animal suffering, he abandons his intuitionist approach in favour of moral reasoning under the Suffering/Utility Standard.

Posner's substantive argument is ultimately the same as the EU and *amici*, especially in the following key respects:⁷⁶

- They rely on intuitionism to justify special protection for charismatic species (such as dogs and seals).
- They rely on consequentialist moral reasoning (the Suffering/Utility Standard) to justify the painful exploitation of commodity animals.
- They use intuitionist arguments to bypass certain aspects of welfarism, especially the consistency requirement.

In this chapter, I have explained why the pure intuitionist argument cannot succeed under Article XX(a) unless the Appellate Body adopts evidentiary unilateralism.⁷⁷

However, the EU bypasses this problem by framing its explicit defence as welfarism (while merely making implicit arguments from intuitionism). This is the opposite framing to Posner who is explicitly intuitionist (but incorporates implicit arguments from welfarism). In both cases, their substantive position is based on the same combination of intuitionism and welfarist moral reasoning. The next chapter will explain why this "hybrid" claim is inherently flawed.

⁷⁶ There is a difference in form as the EU brands its argument as "welfarist" while Posner claims that his approach emanates from outside moral philosophy.

⁷⁷ Either for all public morals disputes or for a subset of "non-instrumental" claims.

CHAPTER 11 - THE HYBRID APPROACH: WELFARISM MIXED WITH INTUITIONISM

11.1 Introduction

In the previous chapter, I argued that the EU/*amici* and Posner rely on a hybrid claim which combines both welfarism and intuitionism. Their substantive reasoning is remarkably similar, but it is framed in different ways. The EU explicitly endorses welfarism (but implicitly incorporates intuitionism). Posner starts at the other end of the spectrum by explicitly endorsing intuitionism (and implicitly incorporating welfarism). In both cases, the end point is the same: a welfarist/intuitionist hybrid claim.

From the EU's perspective, the main benefit of the hybrid claim is that it has the credibility of Singer's welfarism (and therefore appears to be based on "moral standards"), but the intuitionist elements liberate the EU from the more demanding aspects of welfarism, especially the consistency requirement. If the hybrid claim is valid, the EU can claim to be welfarist, but blatantly engage in speciesist practices.

In my view, intuitionism and welfarism cannot be coherently combined into a hybrid claim because they are mutually exclusive. Unfortunately, adjudicators adopt an equivocal definition of the EU's objective (seal welfare) which fails to clarify whether it is welfarist or intuitionist. In this chapter, I will show that this equivocal definition enables adjudicators to oscillate between the welfarist and intuitionist objectives in ways which lead to flawed logical reasoning.

The main problem with the hybrid claim is that it is riddled with contradictions and logical inconsistencies. This chapter will demonstrate that welfarism and intuitionism

are mutually-exclusive and incapable of being combined into a joint approach. I will highlight the logical fallacies and questionable techniques which lie at the heart of the hybrid claim including: (i) the use of equivocal language to characterise the EU's objective; (ii) contradictory views on key issues such as risk, science, suffering and instrumental reasoning; and (iii) the application of double standards under the provisional justification.¹

11.2 Welfarism and Intuitionism are Mutually Exclusive

This section will demonstrate that welfarism and intuitionism are mutually-exclusive. It is instructive to highlight some fundamental differences between them:

- Whereas welfarism regulates commodity species, intuitionism only applies to certain cuter (or charismatic) species.
- Whereas welfarism relies on instrumental reasoning related to sentience and suffering, intuitionism relies on sentiment.
- Whereas welfarism focusses on how animals are exploited (regulation), intuitionism focusses on whether animals should be exploited (prohibition).
- Whereas welfarism would permit dog and seal exploitation (and even suffering), intuitionism offers "total protection" (even from pain-free exploitation) but only for certain species.
- Whereas welfarism relies on science (to measure animal suffering) and risk assessments (to mitigate suffering), intuitionism relies on non-instrumental factors such as sentiment and perception.

There is a clear tension between welfarism and intuitionism with respect to many key themes. The table below highlights some issues where welfarism and intuitionism are diametrically-opposed:

	Welfarism	Intuitionism
Status of animal	Commodity	Special (Cuter or Charismatic)
Based on science and risk assessments	Yes	No
Based on moral standards and	Yes	No

¹ The next chapter will focus on a further problem with the hybrid claim: the mixed messages it delivers on the matters of speciesism and regulatory consistency.

consequentialist/instrumental reasoning		
Tolerates some animal suffering	Yes	No
Measures animal suffering	Yes	No

Depending on which objective the EU is pursuing, there are flow-on effects for our conceptual understanding of the seal products ban. Here are some key conceptual issues where the two claims are fundamentally different.

	Welfarism	Intuitionism
Moral and policy reason for protecting seals	Suffering and sentience	Cuteness and sentiment
Regulatory approach	Regulation of exploitation practices to reduce (but not necessarily eliminate) pain and suffering	Abolition of exploitation (even pain-free exploitation) through species-wide bans
Key analytical question	Which seal exploitation practices fail the Suffering/Utility Standard?	Is it categorically wrong to exploit seals?
Which other animals can seals be compared to?	Other commodity animals (cows, chickens)	Other charismatic species (dogs, cats)

This suggests that it would be impossible for adjudicators to undertake a meaningful analysis of the hybrid claim. After all, if they seek to apply the hybrid claim, they will have to hold logically contradictory views on key issues such as the status of seals (commodity or charismatic species) and the relevance of moral standards, suffering, science and risk.² In practice, adjudicators followed the EU's hybrid approach leading to a series of logical fallacies which I will describe in this chapter.

² An alternative way of understanding these logically contradictory views is that adjudicators are oscillating between the two different claims. When adjudicators allow instrumental reasoning (science, risk and suffering), they are reasoning based on the welfarist objective. When they reject the use of instrumental reasoning, they are using intuitionist reasoning.

11.3 Logical Fallacy I: Equivocal Language

11.3.1 Adjudicators' Approach

If the hybrid claim is incoherent, adjudicators were free to reject it.³ It is the panel's responsibility to characterise the policy objective and it could have chosen to split the EU's defence into two separate, coherent claims (intuitionism and welfarism).⁴

Instead, the panel invented the equivocal term "seal welfare" to describe a single overarching policy objective:

"For ease of reference, we will use in these Reports the phrase "addressing public moral concerns on seal welfare" as shorthand for the specific objective".⁵

The ambiguous term "seal welfare" fails to clarify whether the EU's defence is welfarist or intuitionist. It contains the word "welfare" which strongly hints at a welfarist approach. However, the term "seal welfare" might imply that seals are special. After all, there is no such thing as "cow welfare" or "chicken welfare"; there are merely "animal welfare" standards which apply generally to all commodity animals.⁶

Adjudicators provided an expanded definition of seal welfare:

"The objective of the EU Seal Regime is "to address the moral concerns of the EU public with regard to the welfare of seals". The Panel elaborated that these concerns have two specific aspects: (i) "the incidence of inhumane killing of seals"; and (ii) "EU citizens' individual and collective participation as

³ The WTO's approach to characterising the objective was discussed in section 3.2.3 and the *Brazil Tyres* case study (chapter 7).

⁴ These claims were analysed extensively in chapters 9 and 10. The panel also had the option of assessing both the welfarist and intuitionist claims as separate (parallel) defences.

⁵ Panel Reports, *European Communities – Measures Prohibiting the Importation and Marketing of Seal Products*, WT/DS400/R and Add.1 / WT/DS401/R and Add.1, adopted 18 June 2014 (*Seals*) para 7.410 (footnote 675). The Appellate Body also relied on this same equivocal "seal welfare" characterisation of the objective.

⁶ If "seal welfare" represents a special welfare standard just for seals, this would deviate from the standard welfarist understanding of animal protection.

consumers in, and exposure to ("abetting"), the economic activity which sustains the market for seal products derived from inhumane hunts".⁷

This complex definition fails to clarify whether the EU's defence is welfarist or intuitionist.

This full definition uses terms and concepts which clearly evoke welfarism, such as "welfare", "inhumane killing" and "inhumane hunts". However, there is no mention of animals or "animal welfare". This is strange as adjudicators often rely on animal welfare examples to demonstrate the legitimacy of the EU's measure.⁸ In fact, the Appellate Body uses the term "animal welfare" 127 times throughout its legal reasoning.⁹ If the EU's defence is welfarist, why does the Appellate Body refer to seals but avoid any explicit mention of "animals"?

Further, the definition fails to mention the "utility" of seal or animal products. If the EU's defence is based on welfarism's Suffering/Utility Standard, how can we determine whether seal suffering is justified without an analysis of "utility"?

The failure to mention "animal welfare" or "utility" suggests that seal welfare is not based entirely on welfarism. It seems to imply that seals are special and that the EU's objective is intuitionist. However, the characterisation of the objective fails to explicitly mention any key elements of an intuitionist claim such as "moral intuitions", "cuteness" or "subjective feelings of disgust".

While the characterisation of the objective lacks an explicit reference to intuitionism, it is nonetheless too vague to preclude it either. Indeed, one of the hallmarks of intuitionist arguments is that they can influence our decisions even without being

⁷ Appellate Body Reports, *European Communities – Measures Prohibiting the Importation and Marketing of Seal Products*, WT/DS400/AB/R / WT/DS401/AB/R, adopted 18 June 2014 (*Seals*) para 5.139.

⁸ I will elaborate on this point in the discussion of double standards below.

⁹ This is only slightly less than the term "seal welfare" which is used 139 times. I obtained these figures by undertaking a word search of the report.

explicitly invoked. I will show that, in practice, the EU manages to benefit from intuitionist arguments, especially for the purpose of justifying the fact that seals receive special treatment.¹⁰

WTO adjudicators should have treated the EU's competing objectives as a threshold question. At the outset, they should have determined whether the EU considers seals to be a commodity species (regulated under a welfarist framework) or a charismatic species (regulated under an intuitionist framework). This would have guided adjudicators on which claim to address.

The adoption of an equivocal objective, based on the hybrid claim, creates space for logical contradictions and double standards. This ultimately contaminates the subsequent reasoning.

11.3.2 EU's Approach

This section provides an example (from the EU's transnationalist evidence) of how the hybrid claim can undermine logical reasoning.¹¹

Transnationalist evidence is certainly useful in Article XX disputes, but only where the examples are analogous to the contested measure. To support seal welfare, should the EU furnish examples of welfarist measures or intuitionist measures?

The EU provides the following hybrid example which seems to blur the line between welfarism and intuitionism:

¹⁰ As discussed in chapter 10, the main role which intuitionist arguments play in the moral reasoning of the EU (and likeminded scholars) is to sidestep the consistency requirement and permit special treatment for charismatic species.

¹¹ The use of transnationalist evidence was discussed in chapter 8. Ultimately, it permits the EU to argue that welfarism/intuitionism is legitimate by providing examples of welfarist/intuitionist measures from other countries.

“the state of California bans the sale of foie-gras on animal welfare grounds, whereas the states of Illinois, New Jersey, Oklahoma and Texas prohibit the marketing of horse meat for human consumption”.¹²

This hybrid example contains one welfarist and one intuitionist objective. The EU explicitly states that the foie-gras ban is an “animal welfare” measure (a convincing claim), but is silent on the rationale for the horse meat ban. I will show below that such bans are intuitionist in nature.

The EU describes the foie-gras measure as a “ban”, so it sounds like the seal products ban, but they are quite different substantively. California bans a particular product (foie-gras), based on a practice (force-feeding) which it deems excessively inhumane.

However, the foie-gras measure is not a species-wide ban granting geese a “right to life”. Geese remain a commodity species which can be (painfully) exploited. They can be de-beaked, held in cages, separated from their offspring and slaughtered for commercial purposes. The foie-gras ban is therefore not analogous to the EU’s dog or seal product bans.

The horse meat ban is grouped with the foie-gras ban as a single example, so inattentive observers may think it must be an “animal welfare” measure. However, this view is implausible as these horses are not governed by welfarism’s Suffering/Utility Standard for commodity animals. Rather, the horse has obtained a special status in certain US States where it has been granted a “right to life”. These horses are protected based on intuition and sentiment (not sentience).

¹² EU, ‘First Written Submission by the European Union’ (*EC – Seals*, 21 December 2012) <https://trade.ec.europa.eu/doclib/docs/2012/december/tradoc_150190.pdf> (consulted 5 April 2020) (“EU First Panel Submission”) para 75.

The horse meat ban also represents different regulatory treatment compared to the foie-gras ban. The measure protects horses through a species-wide ban on slaughter and consumption. The horse meat ban does not merely ban a particular practice or product, but provides total protection to a species.¹³ The horse meat ban is an intuitionist (not a welfarist) measure which is analogous to the seal products ban.

By combining the foie-gras and horse meat measures into a single example, the EU muddies the waters and perpetuates the hybrid claim. This is based on incoherent reasoning which adjudicators should have rejected by establishing a clear position on what the EU's policy objective was. The next section will show how the hybrid claim leads to contradictory reasoning in practice.

11.4 Logical Fallacy II: Contradictory Reasoning

Welfarism and intuitionism have diametrically-opposed views on the relevance of suffering, science, risk and consequentialist reasoning? Under the hybrid claim, how do scholars and adjudicators deal with these contradictory views? I will briefly describe the *amici*'s approach as a prelude to the Appellate Body's approach.

11.4.1 The Amici and Instrumental Reasoning

In my view, the *amici* deal with the hybrid claim by holding contradictory views about key aspects of the EU's defence. In some contexts, they argue that the EU's defence is based on instrumental/consequentialist reasoning, science and risk. In other contexts, they argue the exact opposite. I will provide examples below.

The *amici* suggest that animal welfare is an instrumental claim:

¹³ It is worth noting that horses, dogs and cats are arguably still exploited by humans (for example as pets), though this raises different public policy (and moral) issues compared to the slaughter of commodity animals for consumer products.

“The implementation of instrumental moral legislation with respect to animal welfare, in which the goal is to reduce pain and suffering under a consequentialist moral framework, is relatively straightforward for a judge to understand and assess”.¹⁴

They further reinforce this instrumental view of animal welfare by stating that:

“pain and suffering can be substantiated by scientific and other types of empirical evidence that facilitates judicial analysis of the efficacy and thus the legality of legislation”.¹⁵

These arguments appear consistent with a welfarist defence (and the Suffering/Utility Standard) as they invoke science, suffering and instrumental/consequentialist reasoning. However, the *amici* subsequently make contradictory arguments suggesting that animal welfare is non-instrumental in nature.

They assert that “the instrumental rationality typically deployed in WTO dispute settlement” will be “incapable of doing full justice to animal welfare”.¹⁶ They further state that the Appellate Body report:

“recognizes that all legislation is not amenable to cost-benefit analysis or an assessment of its effectiveness in achieving material goals. And this is particularly so in the context of non-instrumental moral legislation such as that at stake in the Seal Products dispute”.¹⁷

They suggest that:

“appreciating the moral grounds for a regulation of this kind confirms the limits of science and instrumental policy reasoning in deciding a dispute of this nature”.¹⁸

¹⁴ Robert Howse, Joanna Langille and Katie Sykes, ‘Pluralism in Practice: Moral Legislation and the Law of the WTO after Seal Products’ (2015) 48 George Washington International Law Review 81, 84.

¹⁵ *ibid*.

¹⁶ Robert Howse and Joanna Langille, ‘Permitting Pluralism: The Seal Products Dispute and Why the WTO Should Accept Trade Restrictions Justified by Noninstrumental Moral Values’ (2012) 37(2) Yale Journal of International Law 367, 372.

¹⁷ Howse et al (n 14) 144.

¹⁸ Howse et al (n 16) 384.

These arguments suggest that animal welfare is a non-instrumental claim which is incapable of being resolved by reference to science, suffering and instrumental/consequentialist reasoning.

Surely, the EU's defence cannot be both instrumental and non-instrumental at the same time.¹⁹ The *amici*'s analysis appears riddled with confusion and inconsistencies. They appear to hold contradictory views on whether suffering, science and risk are relevant to the dispute.²⁰ The next section will show that they are not alone.

11.4.2 Adjudicators: Science and Risk

This section will show that the Appellate Body perpetuates the same contradictions as the *amici*. Rather than having a fixed view on the role of science and risk (either they are relevant or not), it seems to change its view from one legal element to the next. When Canada makes risk and science-based arguments, adjudicators reject them as irrelevant. However, when the EU makes such arguments, adjudicators accept them. I will illustrate this tendency with some examples below.

Appellate Body rejects Canada's arguments

Canada relies on science and risk to attack the EU's welfarist defence:

"Canada argues that the Panel failed to consider whether the risks associated with commercial seal hunts 'exceeded the accepted level of risk of compromised animal welfare, as reflected in the [European Union's] policies and practices in this field'."²¹

¹⁹ Although, as mentioned above, this would not prevent the EU from pursuing its intuitionist and welfarist claims as parallel defences.

²⁰ Perhaps they have an unstated system, which makes intuitive sense to them (and many readers), about when to treat the EU's defence as instrumental and when to treat it as non-instrumental.

²¹ Appellate Body Reports, *Seals* (n 7) para 5.194.

If the EU's defence is welfarist, Canada's argument seems reasonable. Since animal welfare standards rely on risk assessments related to animal suffering, adjudicators should assess the risk of seals being treated inhumanely relative to the EU's overarching risk tolerance for animal welfare.

However, the Appellate Body rejects outright the relevance of risk:

"risk-assessment methods do not appear to be of much assistance or relevance in identifying and assessing public morals".²²

The Appellate Body concludes that:

"We reject Canada's argument that the Panel was required to assess whether the seal welfare risks associated with seal hunts exceed the level of animal welfare risks accepted by the European Union in other situations such as terrestrial wildlife hunts".²³

The Appellate Body's view that risk is irrelevant to the Seals dispute (and public morals more broadly) seems categorical.

Appellate Body accepts EU's arguments

However, when raised by the EU, the Appellate Body treats science and risk-based claims as admissible and compelling. The EU's welfarist defence is clearly based on science.²⁴ Its arguments appear in a 44-page section titled "Scientific Grounds for the Public Moral Concerns".²⁵ The EU states that "it is appropriate to take into

²² ibid para 5.198.

²³ ibid para 5.201.

²⁴ Sykes argues that the instrumental and science-based credentials of the EU's claim are sufficiently strong for it to be defended under Article XX(b). They argue there is a "very strong doctrinal basis" for dealing with "the protection of individual animals from adverse welfare effects" under Article XX(b); Katie Sykes, 'Sealing Animal Welfare into the GATT Exceptions: the International Dimension of Animal Welfare in WTO disputes' (2014) 13(3) World Trade Review 471, 492.

²⁵ EU First Panel Submission (n 12) pages 28-72.

account available relevant scientific evidence, as the EU legislators did in this case".²⁶

The scientific nature of animal welfare measures can also be found in EU legislation, such as its animal slaughter regulation which states:

"Measuring the lack of consciousness and sensibility of an animal is complex and needs to be performed under [a] scientifically approved methodology".²⁷

The EFSA Report, which the EU cites extensively for demonstrating that seal slaughter is "inhumane", is described in its title as a "Scientific Opinion" produced by a "Scientific Panel".²⁸

The EFSA Report is also based on risk assessments. It states:

"In line with the terms of reference for EFSA, the working group carried out a Risk Assessment, but, due to the limited amount of data, only a qualitative Risk Assessment could be done".²⁹

The OIE guidelines (cited by the EU as internationalist evidence) also considers animal welfare to be a science-based matter when it states:

"The scientific assessment of animal welfare has progressed rapidly in recent years and forms the basis of these recommendations".³⁰

The OIE Guidelines also endorse risk assessments. For example, in the context of live transport, they state:

"Animals confined in groups, especially in pens, should be stocked at a high enough density to prevent injuries at take-off, during turbulence and at

²⁶ ibid para 82.

²⁷ Council Regulation (EC) No 1099/2009 of 24 September 2009 on the protection of animals at the time of killing (2009) OJ L 303/1 preamble (20th recital).

²⁸ EFSA, 'Scientific Opinion of the Panel on Animal Health and Welfare on a Request from the Commission on the Animal Welfare Aspects of the Killing and Skinning of Seals' (2007) 610 The EFSA Journal 1.

²⁹ Ibid 12.

³⁰ World Organisation for Animal Health, *Terrestrial Animal Health Code* (20th edition, 2011) (OIE Guidelines) Art 7.1.3.

landing, but not to the extent that individual animals cannot lie down and rise without *risk of injury or crushing*".³¹

"Risks during transport can be reduced by selecting animals best suited to the conditions of travel and those that are acclimatised to expected weather conditions".³²

When argued by the EU, adjudicators accepted the notion that animal welfare in general, and the seal products ban in particular, is risk-based.³³ For example, the Appellate Body noted "extensive information on the Panel record regarding the 'welfare *risks* of seal hunting':"³⁴

The Appellate Body also rejected a proposed alternative measure on the basis that:

"it was not reasonably available given *inter alia* the inherent animal welfare risks and challenges found to exist in seal hunting".³⁵

The Appellate Body also relies extensively on science, including the aforementioned EFSA Report and OIE Guidelines. It validates the scientific evidence submitted by the EU when it states that:

"there are inherent obstacles that make it impossible to kill seals humanely on a consistent basis, and that the European Union 'submitted extensive scientific evidence and argument to the Panel in order to substantiate that position'."³⁶

When raised by the EU, the Appellate Body appears to strongly agree that science and risk are highly relevant to animal welfare and the *Seals* dispute. As discussed in chapter 9, this seems appropriate for a welfarist claim.

³¹ ibid Art 7.4.3. These are just a few select examples of the relevance of risk to animal welfare (the Guidelines also refer to other risks such as those to human health).

³² ibid Art 7.2.7.

³³ The EU also made risk-based arguments when seeking to defend its exception: "it may be justified, or even required, from a moral point of view to tolerate a higher level of risk to the welfare of seals" for certain types of seal exploitation such as Inuit seal products. EU First Panel Submission (n 12) para 39. It is outside the scope of my analysis to assess the exceptions in detail as they were addressed under the chapeau.

³⁴ Appellate Body Reports, *Seals* (n 7) para 5.284.

³⁵ ibid para 5.289.

³⁶ ibid para 5.284.

But why did the Appellate Body reject Canada's science and risk-based arguments?

It offered the following (unconvincing) explanation:

"While the focus on the dangers or risks to human, animal, or plant life or health in the context of Article XX(b) may lend itself to scientific or other methods of inquiry, such risk-assessment methods do not appear to be of much assistance or relevance in identifying and assessing public morals. We therefore do not consider that the term "to protect", when used in relation to "public morals" under Article XX(a), required the Panel, as Canada contends, to identify the existence of a risk to EU public moral concerns regarding seal welfare".³⁷

With this statement, the Appellate Body rejects the relevance of science and risk not just for the *Seals* dispute, but for all public morals disputes. This view contradicts earlier public morals jurisprudence where adjudicators clearly alluded to instrumental reasoning, risk and science.

In *Gambling*, the Appellate Body referred to "risks to youth, including underage gambling"³⁸ and "health risks associated with addiction to gambling".³⁹ The panel considered "scientific evidence" on whether "remote access gambling is any more problematic than other forms of gambling when it comes to compulsive, problem or other forms of pathological gambling".⁴⁰

Instrumental reasoning, science and risk can (and sometimes do) play an important role in public morals disputes. The *Seals* adjudicators could have made the less extreme claim that some public morals objectives can be non-instrumental in nature, such as the EU intuitionist claim. However, it should have simultaneously recognised that many moral claims remain instrumental in nature, such as gambling and animal welfare.

³⁷ ibid para 5.198.

³⁸ Appellate Body Report, *United States – Measures Affecting the Cross-Border Supply of Gambling and Betting Services*, WT/DS285/AB/R, adopted 20 April 2005 (*US – Gambling*) para 283.

³⁹ ibid 94.

⁴⁰ Panel Report, *United States – Measures Affecting the Cross-Border Supply of Gambling and Betting Services*, WT/DS285/R, adopted 20 April 2005 (*US – Gambling*) 3.177.

Adjudicators could reasonably hold the view that science and risk are relevant for some claims (such as welfarism) but not for others (such as intuitionism). However, the *Seals* adjudicators seem to determine the relevance of risk and science based on which party is making the claim (the EU or Canada) rather than which claim is being discussed. This type of bias (based on which party makes the claim) is wholly inappropriate.

The *Seals* adjudicators held contradictory views on the relevance of instrumental reasoning, science and risk. In my view, this is because the hybrid claim permitted them to (unconsciously) switch between two different conceptions of “seal welfare”. When they accepted the EU’s instrumental arguments, they were thinking like welfarists. When they rejected Canada’s instrumental arguments, they were thinking like intuitionists. This is how the hybrid claim leads to confused and contradictory reasoning. The unanswered question is: how did adjudicators know when to switch from a welfarist to an intuitionist mode of thought?

Non-instrumental claims and judicial reform

The EU’s intuitionist defence appears to be the first non-instrumental claim in WTO case law. It therefore raises some new and important questions: are non-instrumental claims valid? If so, do they require a different legal test? If instrumental reasoning is not available, how can the legitimacy of non-instrumental norms be tested? What is the rule for determining which norms are non-instrumental (and therefore subject to a different legal test)?

The Appellate Body failed to address these questions or create a framework for distinguishing between instrumental and non-instrumental moral objectives. This is because the policy objective they assessed (seal welfare) was equivocal.

Adjudicators made the contradictory claim that all public moral matters are non-instrumental (and that science and risk are therefore irrelevant), while simultaneously accepting the EU's instrumental arguments.

If adjudicators decide to accept non-instrumental norms as a separate category under Article XX(a), it would have to reform the current jurisprudence. First, it would require a rule for determining which objectives are non-instrumental in nature. In the case of animal ethics, this would enable adjudicators to distinguish intuitionist claims (which are non-instrumental) from welfarist claims (which are instrumental).

Adjudicators should have separated the EU's claims into alternative (parallel) defences governed by different jurisprudential approaches.

Second, adjudicators would have to clarify the legal test they apply to non-instrumental norms. Since instrumental reasoning is not available (such as rationalist philosophy or external standards), how would adjudicators determine legitimacy? One option would be to adopt evidentiary unilateralism and to develop principles for how regulators can demonstrate that a contested social norm truly is widely-supported within their society.

Summary

The Appellate Body's reliance on the hybrid claim leads to incoherent reasoning. After all, welfarism and intuitionism are mutually-exclusive approaches based on completely different views about the relevance of science, risk and instrumental reasoning. Under the hybrid claim, the Appellate Body did not hold a fixed view about the EU's objective, but rather oscillated between welfarism and intuitionism from one issue to the next.

How did adjudicators know when it was appropriate to switch between the welfarist and intuitionist modes? On the matters of science and risk, they followed a simple heuristic: they accepted science and risk-based arguments from the EU, but rejected them from Canada.

11.5 Logical Fallacy III: Double Standards

This section will provide two examples where the equivocal meaning of “seal welfare” leads to double standards: the “design step” and W&B test. Within each test, the meaning of “seal welfare” changes significantly from one legal element to another.⁴¹

11.5.1 Design Step

The design step contains two elements: public concerns and moral standards.⁴² In *Seals*, adjudicators used a different meaning of “seal welfare” for each of these elements.⁴³ In fact, the double standard is signposted by a shift in language away from “seal welfare”:

“Given its finding that the concerns of the EU public on *animal welfare* involved standards of right and wrong within the European Union as a community, the Panel found that addressing public moral concerns on seal welfare is a ‘legitimate’ objective”.⁴⁴

The Appellate Body finds that “animal welfare” is based on moral standards, but that the public moral concerns related to “seal welfare”. Is this just sloppy drafting or

⁴¹ This analysis builds on the method I used to demonstrate systematic bias in the *Brazil Tyres* case study.

⁴² While the “moral standards” requirement has been part of public morals jurisprudence since the first dispute (*Gambling*), the public concerns element was introduced in *Seals*. This was analysed in chapter 8.

⁴³ In chapter 8, I already showed that adjudicators use different types of evidence for each legal element of the public morals test. Public concerns relied on evidentiary unilateralism, while moral standards relied on external standards. In this section, I will show that the meaning of the objective shifted from intuitionism (for public concerns) to welfarism (for moral standards).

⁴⁴ Appellate Body Reports, *Seals* (n 7) para 5.140. The panel initially undertook this as a “legitimate objective” analysis under the TBT Agreement. Both the AB and panel ultimately relied on this same analysis for their assessment of “standards of right and wrong” under Article XX(a).

does it signal that the Appellate Body is oscillating between welfarism and intuitionism?⁴⁵

Under the hybrid claim, adjudicators do not explicitly state when they are shifting from welfarism to intuitionism, so I will use the following method. Where adjudicators focus on subjective perceptions or speak exclusively about seals, I will treat this as the intuitionist claim. Alternatively, where adjudicators focus on sentience and incorporate broader evidence of animals (especially commodity animals) or “animal welfare”, I will treat this as the welfarist claim.

11.5.1.1 Public Concerns (Seal Welfare)

On public concerns, the evidence is seal-specific. There are no references to animal welfare, the Suffering/Utility Standard or the treatment of commodity animals.

The panel noted:

“During the last decade, the cruelty of seal hunting has been documented by videos from several authoritative television channels as well as by the personal observations of many members of national and European parliaments, scientists, celebrities and representatives of non-governmental organizations (NGOs). Such cruelty has generated a public morality debate in Europe”.⁴⁶

The Appellate Body further noted:

“The Commission received during the last years a massive number of letters and petitions on the issue expressing citizens' deep indignation and repulsion regarding the trade in seal products in such conditions”.⁴⁷

“Public survey results submitted by the European Union were also informative, although "to a limited extent", in demonstrating the EU public's concerns”.⁴⁸

⁴⁵ Further, it seems strange that adjudicators would even refer to “animal welfare” when the characterisation of the objective expressly avoided this term in favour of “seal welfare”.

⁴⁶ Panel Report, *Seals* (n 5) para 7.392.

⁴⁷ Appellate Body Reports, *Seals* (n 7) para 5.153.

⁴⁸ *ibid* para 5.135.

This evidence suggests the public concerns are based on strong subjective feelings (indignation, repulsion) by EU stakeholders such as Parliamentarians, NGOs, celebrities, the media and the public. This sounds like the intuitionist claim.

The evidence contains a passing reference to “cruelty” (which hints indirectly at welfarism), but there is no attempt to use science to quantify this cruelty or instrumental reasoning and risk analysis to compare it to a broader animal welfare standard. The concerns are therefore based on subjective perceptions, rather than an instrumental assessment, of cruelty. This is an intuitionist, not a welfarist, approach.

11.5.1.2 Moral Standards (Animal Welfare)

For moral standards, adjudicators rely on welfarist reasoning. This is hinted at by the use of the term “animal welfare”, but it is also inherent in the external evidence cited by adjudicators (internationalism, moral philosophy and transnationalism). Unlike the public concerns requirement, adjudicators rely on objective evidence, rather than subjective perceptions.

Internationalism

Adjudicators cited internationalist evidence, including the “recommendations of the Office International des Epizooties (OIE) (Guiding Principles for Animal Welfare)⁴⁹. These Guiding Principles focus on the treatment of commodity animals under the Suffering/Utility Standard, a welfarist approach. Even though the EU concedes that the Guiding Principles do not address seal exploitation (perhaps because they are not universally viewed as a commodity species), adjudicators conclude nonetheless

⁴⁹ Panel Report, *Seals* (n 5) para 7.408.

that “the existence of those international instruments confirms that preoccupation with the welfare of animals is a widely shared moral concern.”⁵⁰

The EU fails to produce any internationalist evidence supporting its intuitionist claim. This fact should undermine the EU’s defence. As Lester notes, “in the absence of such [international] agreements on seal hunting, the claimants may argue that unilateral trade action is not permitted.”⁵¹ However, the EU and adjudicators circumvent this weakness in the intuitionist defence by finding internationalist evidence which supports the welfarist defence.

Expert Evidence (Moral Philosophy)

As discussed in Chapter 9, the welfarist claim has strong credentials in moral philosophy. However, the EU produces no arguments from moral philosophy justifying special protection for certain animals based on intuitionist considerations (such cuteness or intuitive reactions to suffering).

The EU finds support for intuitionist judgments when it invokes the so-called “socio-zoological scale”. However, it is important to note that the “socio-zoological scale” does not make normative claims about how people should treat animals, but merely seeks to describe how people actually behave.⁵²

⁵⁰ EU First Panel Submission (n 12) para 71.

⁵¹ Simon Lester, ‘The WTO Seal Products Dispute: A Preview of the Key Legal Issues’ *ASIL Insights* 14(2) (13 January 2010).

⁵² In this sense, the evidence would be more useful under the descriptive “public concerns” discussion than the normative “moral standards” discussion.

In fact, its proponents explicitly recognise that the scale has weaknesses “on both scientific and ethical grounds”.⁵³ They note that their analysis is “based on traditions and prejudices” and seeks to describe the “social reality”.⁵⁴

The EU makes arguments from moral philosophy which strongly support a welfarist defence, however it provides no philosophical evidence supporting its intuitionist claim. It provides non-normative arguments suggesting that people care more about “cuter” animals but, like Posner, it fails to explain why this should support a normative conclusion that it is morally-right to offer them special protection.

Transnationalism

From internationalism and philosophy, the EU furnished exclusively welfarist evidence. Under transnationalism, it provides both welfarist and intuitionist evidence.⁵⁵ The panel explicitly cites measures in other countries, such as bans on dog products, which are clearly intuitionist rather than welfarist.⁵⁶ The panel also cites:

“certain other WTO Members’ measures on seal products based on moral grounds (e.g. Chinese Taipei; Russia; and Switzerland)”.⁵⁷

Even though these measures are seal-specific, the panel is unclear about whether these measures are hunting bans or product bans.⁵⁸

The EU’s intuitionist evidence is mixed up with welfarist examples in a way which is likely to lead to confusion.⁵⁹ The panel appears to follow the EU’s practice of mixing

⁵³ EU First Panel Submission (n 12) para 73 (citing Peter Sandøe and Stine Christiansen).

⁵⁴ ibid.

⁵⁵ I analysed this above in the discussion of the fois-gras and horse meat bans.

⁵⁶ See eg Panel Reports, *Seals* (n 5) para 7.406 (and accompanying footnotes).

⁵⁷ ibid para 7.408.

⁵⁸ See eg ibid (and footnote 670) describing the Russian measure. This distinction is important as hunting bans are far less controversial as they do not prevent imports. Indeed, when the EU banned seal hunting (a few decades before the seal products ban), this did not lead to trade frictions with other countries.

welfarist and intuitionist examples when it cites as relevant transnationalist evidence “various measures on animal welfare and seal products adopted by other WTO Members”.⁶⁰

Some of the evidence seems completely irrelevant to the seal products ban. For example, the panel refers to a “comprehensive body of legislation on the welfare of farm animals within the framework of its Common Agricultural Policy”.⁶¹ While this evidence could help support a welfarist defence, it does not seem at all relevant to the intuitionist defence.

Summary of Evidence

Adjudicators find (convincingly) that the EU’s welfarist defence is based on moral standards. To reach this finding, they rely extensively on evidence related to the treatment of commodity species under the Suffering/Utility Standard, especially internationalist evidence (OIE Guiding Principles) and moral philosophy (welfarism).

The only evidence which can support the EU’s intuitionist claim is a handful of transnationalist examples related to seals or other charismatic species. These are interspersed with transnationalist examples supporting the welfarist defence so it is hard to extract the evidence supporting intuitionism or gauge how compelling it is.

The panel concludes that “the evidence as a whole sufficiently demonstrates that animal welfare is an issue of ethical or moral nature in the European Union”.⁶² When analysing “moral standards”, these references to “animal welfare” correspond with a substantive analysis of the welfarist claim.

⁵⁹ See eg the earlier analysis of US State bans on foie gras and horse products where the EU was vague about whether this evidence should support the welfarist or intuitionist defence.

⁶⁰ Panel Reports, *Seals* (n 5) para 7.420.

⁶¹ ibid para 7.406.

⁶² ibid para 7.409.

While, it is clear that welfarism is based on moral standards, adjudicators failed to assess the EU's intuitionist defence under this element. In my view, the intuitionist defence would fail the moral standards requirement.⁶³

11.5.1.3 Summary

There is little doubt that welfarism is a legitimate objective under WTO law. It can be justified using evidence from transnationalism, internationalism, science and moral philosophy. There is no credible scholarship arguing that WTO Members gave up their right to introduce trade-restrictive animal welfare measures.⁶⁴

However, it is highly questionable if intuitionist measures are legitimate in WTO law. Under the moral standards requirement, the EU avoided this problem by framing its defence in welfarist terms (though it did allude to a handful of intuitionist measures among its transnationalist evidence).

The use of double standards enabled adjudicators to avoid this serious weakness in the EU's intuitionist claim. Rather than assessing intuitionism, it cited a wide range of external evidence showing that the EU's welfarist objective was based on moral standards. This had the dual advantage of allowing the EU to produce a greater weight of evidence (internationalist, transnationalist and expert) and to furnish objective types of evidence based on science and risk, such as the OIE Guiding Principles.

However, the adjudicators' use of double standards leaves some significant questions unanswered. Most significantly, it failed to assess whether intuitionist

⁶³ As discussed in the previous chapter, Posner willingly acknowledges that his intuitionist approach is not based on moral standards. I believe that intuitionist defences could only succeed under Article XX(a) if adjudicators adopted an "evidentiary unilateralism" standard.

⁶⁴ For example, Perišin and Sellheim do not deny the legitimacy of welfarist measures but insist, rather, that the seal products ban was not a welfarist measure.

bans are based on moral standards.⁶⁵ Adjudicators avoided the difficult jurisprudential question which lies at the heart of *Seals*: are intuitionist claims valid under Article XX(a), even where there is no external evidence demonstrating they are based on “moral standards”?

11.5.2 Weighing and Balancing Test

In *Seals*, the Appellate Body described the W&B test⁶⁶ as:

“a process of “weighing and balancing” a series of factors, including the importance of the objective, the contribution of the measure to that objective [and] the trade-restrictiveness of the measure”.⁶⁷

This section will show how adjudicators used double standards for the “importance” and “contribution” elements.⁶⁸

11.5.2.1 Importance (Animal Welfare)

When discussing importance, adjudicators refer to animals, not seals. The panel’s analysis is two sentences long, so I cite it in full:

“The European Union submits that the “moral concern with regard to the protection of animals” is regarded as a value of high importance in the European Union. We consider, and the parties do not dispute, that the protection of such public moral concerns is indeed an important value or interest”.⁶⁹

For this element, the panel analysed the “protection of animals”, rather than “seal welfare”. This is an extremely broad objective and the panel was able to conclude, seemingly on the basis of intuition or common sense, that “protection of animals” is important. The analysis of importance contains no evidence or reasoning so it is

⁶⁵ It also failed to consider whether the EU’s public concerns were welfarist in nature, though this issue would not have been controversial if it had been addressed.

⁶⁶ The broader jurisprudence on the W&B test was analysed extensively in chapters 3-7, including its jurisprudential evolution and academic commentary from leading scholars.

⁶⁷ Appellate Body Report, *Seals* (n 5) para 5.169.

⁶⁸ The Appellate Body was also required to analyse the “trade-restrictiveness of the measure”, but since this element does not require any analysis of the objective (“seal welfare”), it is not relevant to my “double standards” analysis.

⁶⁹ Panel Report, *Seals* (n 5) para 7.632.

difficult to confirm whether “protection of animals” represents the welfarist or intuitionist claim.⁷⁰ It is so broad that it could be an umbrella term which incorporates any policy reason for protecting animals.⁷¹

The panel justified its lack of substantive reasoning by claiming that the complainants had conceded “importance”. Canada, however, insists that the panel “misinterpreted” its argument and that it “did not agree that the specific public moral concern in issue was considered important”.⁷² Canada had agreed that public morals were important, but it disputed that “seal welfare” was important. It seems likely that Canada was contesting the importance of the EU’s intuitionist objective.⁷³

The EU could likely demonstrate that welfarism is important, but the intuitionist claim would be harder to defend. What types of non-instrumental reasoning could the EU use to show that intuitionist objectives are important? How can a non-instrumental objective (protecting “cute” animals) be weighed against an instrumental matter (trade-restrictiveness)?

The Appellate Body criticised the panel’s analysis, noting that it “did not provide much elaboration of its assessment of the importance of the objective”.⁷⁴ This is an understatement as the panel provided no analysis whatsoever. However, the Appellate Body appears unperturbed by the use of the wrong objective (“protection of animals”) and refrains from overturning the panel’s findings.

On the contrary, the Appellate Body doubles down on the panel’s erroneous finding by referring to “the importance of public moral concerns regarding *animal welfare*”

⁷⁰ In my view, the reference to “animals” suggests that adjudicators are assessing the welfarist claim, possibly alongside other policy reasons for protecting animals.

⁷¹ Such as welfarism, intuitionism, endangered species and religious reasons.

⁷² Appellate Body Reports, *Seals* (n 7) para 5.202.

⁷³ In my view, it is unlikely that Canada was contesting the importance of the welfarist defence.

⁷⁴ Appellate Body Reports, *Seals* (n 7) para 5.203.

and concluding that “the *protection of animals* is a value of high importance”.⁷⁵ The Appellate Body replicates the panel’s error of talking about animals (not seals) and substituting an incorrect standard (protection of animals) for the correct one (seal welfare).

In my view, adjudicators found that the welfarist claim (or something similar) was important. However, they failed to assess whether the EU’s intuitionist objective was important. This was a serious oversight.

11.5.2.2 Contribution (Seal Welfare)

The contribution analysis was highly complex at both the panel and appellate stage.⁷⁶ With respect to terminology, both adjudicating bodies switched their description of the objective from “protection of animals” (for importance) to “seal welfare” (for contribution). The panel concluded, and the Appellate Body agreed, that the “EU Seal Regime... contributed to a certain extent to its objective of addressing the EU public moral concerns on seal welfare.”⁷⁷

How did adjudicators determine the measure’s contribution when it is unclear what “seal welfare” actually means? This lack of clarity might explain why the panel’s analysis was, according to the Appellate Body, “not very detailed”.⁷⁸

Norway is less diplomatic when it describes the panel’s findings as:

“so beset with equivocation, vagueness, and imprecision that it is impossible to form an understanding as to how the findings justify the overall conclusion that there is even ‘some’ contribution”.⁷⁹

⁷⁵ ibid.

⁷⁶ This analysis will only consider a selection of the issues analysed under contribution with a focus on how the meaning of “seal welfare” evolves.

⁷⁷ Panel Report, *Seals* (n 5) para 7.638.

⁷⁸ Appellate Body Reports, *Seals* (n 7) para 5.228.

⁷⁹ ibid para 2.65.

For the contribution analysis, adjudicators defined seal welfare as “reducing the number of inhumanely killed seals”.⁸⁰ The Appellate states that:

“it was not unreasonable for the Panel to assume that a decrease in demand, and hence a contraction of the seal product market, would have the effect of reducing the number of seals killed, and thus the number of inhumanely killed seals”.⁸¹

While this methodology for “contribution” might seem superficially logical, it fails to clarify whether adjudicators were assessing the welfarist or intuitionist claim. This methodology has certain aspects which are consistent with both approaches, but it is not fully coherent with either.

Under a welfarist approach, the EU could seek to reduce cruelty, for example by regulating the slaughter process to ensure minimum standards of humane slaughter.⁸² As long as those standards were met, it would not seek to reduce total seal deaths. The desire to reduce the number of seals killed does not make sense under a welfarist framework.⁸³

Further, the EU’s measure is not truly welfarist as it does not seek to regulate slaughter techniques to make them more humane. Marceau notes that the measure is not really an “anti-cruelty ban” as it “does not restrict access to the market for seal products based on how they are killed (humanely rather than inhumanely)”.⁸⁴

⁸⁰ Ibid, para 5.244.

⁸¹ Ibid, para 5.247.

⁸² In the next chapter, I will elaborate on the arguments which the EU uses to link the number of seals killed to the concept of animal welfare. For example, it argues that seal slaughter is “inherently inhumane”. I will highlight the problems with this claim, including why it cannot be understood as a welfarist claim.

⁸³ Animal welfare laws (for cows, chickens, etc) do not seek to regulate or reduce the number of animals killed. Indeed, the number of animals slaughtered has increased drastically since governments first started introducing animal welfare laws.

⁸⁴ Gabrielle Marceau, ‘A Comment on the Appellate Body Report in EC-Seal Products in the Context of the Trade and Environment Debate’ (2014) 23(3) Review of European Community & International Environmental Law 319, 319.

Norway explicitly makes the point that the panel “never considered to what extent these hunts actually involved inhumane killing”.⁸⁵

However, the contribution analysis is not really consistent with intuitionism either. An intuitionist approach would pursue the elimination of seal slaughter (even where it is pain-free) rather than a mere reduction in the number of seals killed.

The adjudicators focus on “reducing the number of inhumanely killed seals” is unsatisfactory under both the intuitionist and welfarist models. It would fail the intuitionist test because it does not eliminate seal consumption in the EU. It would fail the welfarist test because it is based on the number of seals killed rather than how they are killed.

The contribution analysis is highly-confused. The key problem (in terms of my analytical approach) is that adjudicators change their characterisation of the objective for the purpose of the contribution analysis. For this legal element, they refer to “seal welfare” as “reducing the number of inhumanely killed seals”. This is certainly not the same objective which was considered for importance (or moral standards).

11.5.2.3 The Balancing Step

Adjudicators erred by applying a “protection of animals” standard for importance, followed by “seal welfare” for contribution. They had an opportunity to notice and rectify this error under the balancing step. Instead, the Appellate Body appeared to cover up the double standard by claiming the panel:

“evaluated the *importance* of the objective of addressing EU public moral concerns regarding *seal welfare*, the trade-restrictiveness of the EU Seal Regime [and] the contribution of the EU Seal Regime to the objective”.⁸⁶

⁸⁵ Appellate Body Reports, *Seals* (n 7) para 2.62.

This summary gives the impression that adjudicators analysed “the importance of the objective of... seal welfare”, but this is false. Did adjudicators not realise they had actually assessed the importance of “protection of animals”?

The adjudicating bodies failed to apply a coherent standard throughout the W&B test. Both bodies applied a “protection of animals” standard for importance and a dubious “seal welfare” standard (based on reducing the number of seal deaths) for contribution.

The adjudicators’ use of double standards leaves some significant questions unanswered. While it is likely they considered the importance of the welfarist claim, they failed to assess the importance of intuitionist product bans. Regarding contribution, adjudicators considered whether the measure reduced seal deaths, but it is unclear how this makes a meaningful contribution to animal welfare.

11.5.3 Summary

Between the design step and W&B test, adjudicators applied four different interpretations of seal welfare as illustrated in the table below.⁸⁷ Some interpretations appear welfarist, others intuitionist, while some approaches are simply unclear.

Legal Element	Definition of EU Objective	EU Claim
Public Concerns	“Seal welfare” (subjective perceptions about the cruelty of seal slaughter)	Intuitionist
Moral Standards	“Animal welfare” (the Suffering/Utility Standard)	Welfarist
Importance	“Protection of animals”	Probably welfarist
Contribution	“Reducing the number of inhumanely killed seals”	Unclear

⁸⁶ ibid para 5.289.

⁸⁷ Due to limitations of space, I have refrained from analysing further interpretations used by adjudicators under the AM test.

The *Seals* adjudicators relied on a very similar technique to *Brazil Tyres*. They found in favour of the regulator through a seemingly value-neutral analysis, but they surreptitiously made value judgments by changing the definition of the objective. In *Brazil Tyres*, the objectives fluctuated along a scale from broad to narrow. In *Seals*, adjudicators oscillated between two substantively different objectives.

If adjudicators had assessed the EU's intuitionist claim, it would have raised the types of problems foreseen by scholars regarding culturally-specific non-instrumental norms. However, adjudicators managed to resolve the *Seals* dispute with little jurisprudential reform. How did adjudicators resolve the non-instrumental issues in *Seals* using their standard instrumental tools?

Adjudicators avoided the problem created by the intuitionist claim by using double standards. For those elements which were likely to expose the weakness of the intuitionist claim (moral standards and importance), adjudicators defined "seal welfare" as the welfarist claim. For importance and moral standards, they were able to assess welfarism using their standard instrumental tools.

In reality, the *Seals* dispute should have raised some unique and difficult questions about how Article XX(a) operates for non-instrumental claims. Through the use of logical fallacies, adjudicators managed to endorse the legitimacy of the seal products ban, but they somehow managed to avoid providing clear answers to the novel questions raised by the dispute.

11.6 Conclusion

In *Seals*, adjudicators should have addressed welfarism and intuitionism as alternative (parallel) defences. Instead, they followed the EU's lead of combining

both defences into a hybrid claim based on a foundation of logical fallacies, such as equivocal language and double standards. These fallacies have no place in legal reasoning. Of more concern, they were applied in a way which seemed to systematically favour the EU.

The *Seals* adjudicators are not alone. A number of legal scholars, including Posner and the *amici*, share the view that there is a hybrid claim which can simultaneously justify the special protection of charismatic species and the suffering which occurs on factory farms. Moral philosophers would summarily dismiss the validity of this hybrid claim as it is based on blatant logical fallacies and contradictions. Adjudicators and legal scholars should have done likewise.

CHAPTER 12: WELFARISM AND REGULATORY CONSISTENCY

The philosophical literature on animal welfare directly addresses the question of consistency. Singer insists that “all animals are equal” and that welfarism requires them to be treated consistently. He compares the inconsistent treatment of different species (which he calls “speciesism”) to racism and sexism.¹ Singer would find support from HLA Hart who argued that the “leading precept” of justice is “treat like cases alike and different cases differently”.²

Posner, on the other hand, argues that it is legitimate to treat different species inconsistently if we have an “intuition” that a certain charismatic species deserves special protection. According to his approach, animals can cease to be “like” based on human intuitions about how cute they are or how much we believe they deserve protection.

This chapter will consider whether the EU was subject to a consistency requirement under WTO law. It will analyse three possible entry points for consistency analysis under a welfarist public morals defence and three different approaches to consistency used by the *Seals* adjudicators.

12.1 Regulatory Consistency

12.1.1 Consistency, Article XX(a) and Animal Welfare

In principle, the question of regulatory consistency could arise in *Seals* in three ways:

- (i) as a general requirement in Article XX; (ii) under the “moral standards”

¹ Rights philosophers, such as Francione, also oppose speciesist practices which result in the inconsistent treatment of animals.

² HLA Hart, *The Concept of Law* (2nd edition, OUP 1994) 159. Hart illustrates this precept with an example based on racism: “when, in the name of justice, we protest against a law forbidding coloured people the use of the public parks, the point of such criticism is that such a law is bad, because in distributing the benefits of public amenities among the population it discriminates between persons who are, in all relevant respects, alike.”

requirement of Article XX(a); or (iii) under the EU's welfarist defence. I will briefly describe each of these scenarios.

Article XX

While the SPS Agreement explicitly contains a consistency requirement,³ there is a live debate about whether Article XX implicitly contains a similar requirement (or whether it should).⁴ Lydgate suggests that WTO disputes "reveal an emerging emphasis on regulatory consistency".⁵ She "celebrates" the use of consistency requirements by adjudicators, but also cautions that "there is a danger of imposing too much of a consistency requirement".⁶ Her concern relates to how the consistency requirement is applied, not whether it should exist.⁷

Sykes also recognises the growing emphasis on consistency. He refers to "newer consistency requirements" which "allow challenges to domestic regulation based on disparate policies toward different products and industries (such as beef and pork, or salmon and baitfish)." In his view, "consistency analysis based on inter-industry comparators is exceedingly unlikely to afford a persuasive basis for a challenge".⁹ Sykes appears dismissive of consistency arguments, though he does not categorically rule them out. It therefore remains an ongoing debate whether Article

³ The SPS Agreement states: "With the objective of achieving consistency in the application of the concept of appropriate level of sanitary or phytosanitary protection... each Member shall avoid arbitrary or unjustifiable distinctions in the levels it considers to be appropriate in different situations, if such distinctions result in discrimination or a disguised restriction on international trade"; Agreement on the Application of Sanitary and Phytosanitary Measures, Marrakesh Agreement Establishing the World Trade Organization, Annex 1A (adopted 15 April 1994, entered into force 1 January 1995) 1867 UNTS 493 ("SPS Agreement") Art 5.5.

⁴ It is also disputed whether such a requirement does (or should) exist under other Agreements such as GATS and the TBT Agreement.

⁵ Emily Lydgate, 'Is it Rational and Consistent? The Wto's Surprising Role in Shaping Domestic Public Policy' (2017) 20 Journal of International Economic Law 561, 561.

⁶ *ibid* 580.

⁷ This suggests that adjudicators must establish clear principles around how they apply consistency requirements. I will consider the standard of review for consistency below.

⁸ Alan Sykes, 'Regulatory Consistency Requirements in International Trade' (2017) 49 Arizona State Law Journal 821, 821.

⁹ *ibid* 863.

XX contains an overarching requirement for governments to pursue “regulatory consistency”.¹⁰

Moral Standards

Even if there is no regulatory consistency requirement under Article XX in general, it could be argued that the “moral standards” requirement under Article XX(a) contains an implicit consistency requirement. This requirement seems to suggest that only rationalist moral approaches (based on “first principles” and “moral reasoning”) are permitted, such as welfarism and animal rights.¹¹

Singer subscribes to the view that consistency is a core component of moral reasoning:

“We can distinguish the moral from the nonmoral by appeal to the idea that when we think, judge, or act within the realm of the moral, we do so in a manner that we are prepared to apply to all others who are similarly placed”.¹²

As long as the “moral standards” requirement is maintained under Article XX(a), it could be interpreted to require rationalist approaches based on moral reasoning and the use of objective, morally-relevant criteria. In this sense, the Article XX(a) legal test would contain a consistency requirement which does not necessarily exist for other general exceptions.¹³

Welfarist Defences

¹⁰ Unfortunately, I do not have space in this dissertation to enter into a detailed discussion of whether the Appellate Body does (or should) require “regulatory consistency” for all Article XX disputes. I will therefore treat the consistency requirement under Article XX as an open question.

¹¹ This approach would preclude intuitionist defences as they are neither rationalist nor standards-based. In intuitionism, the moral conclusion derives from a descriptive social fact (that people care more about dogs and seals) rather than normative moral reasoning.

¹² Singer P, ‘Ethics beyond Species and beyond Instincts: A Response to Richard Posner’ in Cass Sunstein and Martha Nussbaum (eds), *Animal Rights: Current Debates and New Directions* (OUP 2006) 81.

¹³ Alternatively, if adjudicators dropped the “moral standards” requirement, and adopted evidentiary unilateralism, intuitionist defences would be acceptable and regulators would be exonerated from a consistency requirement.

I have already provided a detailed analysis of the fact that welfarist philosophy contains a consistency requirement.¹⁴ The EU and *amici* invoke welfarism's Suffering/Utility Standard, but fail to acknowledge its consistency requirement. This is a serious gap, as it is unclear how a good faith reading of Bentham, Singer or welfarism could fail to acknowledge the importance of consistency.

It seems disingenuous to make speciesist arguments (in favour of seals) while invoking the welfarist movement. If the EU wished to argue that it is permitted to regulate inconsistently, it should have renounced welfarism and relied on a "pure intuitionism" defence. In my view, even if there is no consistency requirement under Article XX, or the "moral standards" element, any defence which invokes the moral credibility of welfarism is required to respect Singer's consistency requirement.

12.1.2 Consistency Requirements in Seals

12.1.2.1 Academic Arguments

Seals commentators, such as Perišin and Sellheim treat regulatory consistency as a key issue and suggest that the EU's defence was fundamentally undermined by a lack of consistency. Sellheim accuses the EU of lacking "consistency in argumentation" and suggests that its welfarist claims are "opportunistic".¹⁵

Perišin argues that the EU's measure lacks "coherence or consistency".¹⁶ She appears to endorse Singer's consistency requirement when she questions the EU's special treatment of seals for not being "based on any rational differentiation

¹⁴ Bentham and Singer explicitly endorse consistency and oppose speciesism. I have also described how animal rights proponents share this consistency requirement.

¹⁵ Nikolas Sellheim, 'The Legal Question of Morality: Seal Hunting and the European Moral Standard' (2016) 25(2) Social and Legal Studies 141, 151.

¹⁶ Perišin T, 'Is the EU Seal Products Regulation a Sealed Deal? EU and WTO Challenges' (2013) 62(2) International & Comparative Law Quarterly 373, 400.

between seals and other animals".¹⁷ She promotes the consistent application of objective and morally-relevant criteria, rather than subjective intuitions. Sellheim and Perišin clearly believe the EU was under a consistency obligation.¹⁸

Lester notes that the "question might be raised why the EU did not pass a broad animal welfare law that sets out rules for both foreign and domestic products, instead of focusing only on the narrow sub-category of seal products".¹⁹ His comment implies that regulators should consistently apply the same standards to all "policy-like" animals.²⁰ He insinuates that WTO Members are under a consistency obligation for their animal welfare measures and questions the EU's favouritism of seals.

The *amici*, on the other hand, defend the EU by suggesting that it was not subject to a consistency requirement. Their arguments will be assessed further below.

12.1.2.2 Canada's Consistency Argument

Canada explicitly questions the EU's failure to apply its welfarist standard consistently to seals. Here are some examples as summarised by the Appellate Body:

"Canada argues that the Panel failed to consider whether the risks associated with commercial seal hunts 'exceeded the accepted level of risk of compromised animal welfare, as reflected in the [European Union's] policies and practices in this field'."²¹

"According to Canada, identifying such a risk requires the identification of a precise standard of animal welfare in the European Union, and an

¹⁷ *ibid* 395.

¹⁸ Though they do not specify where this obligation arises in WTO law.

¹⁹ Simon Lester, 'The WTO Seal Products Dispute: A Preview of the Key Legal Issues' *ASIL Insights* 14(2) (13 January 2010).

²⁰ This is similar to a philosopher who uses "first principles" as a starting point.

²¹ Appellate Body Reports, *European Communities – Measures Prohibiting the Importation and Marketing of Seal Products*, WT/DS400/AB/R / WT/DS401/AB/R, adopted 18 June 2014 (*Seals*) para 5.194.

assessment of the incidence of suffering in commercial seal hunts against that standard".²²

"Canada recalls that it presented evidence to show that "EU policies and practices with respect to animal welfare included a tolerance for a certain degree of animal suffering, both for slaughterhouses and wildlife hunts".²³

Canada's consistency claim uses a number of different formulations. It argues that the EU must regulate seals consistently with other commodity animals on matters such as suffering, risk, moral standards and the level of protection.²⁴ In addition to being a standalone matter, the question of consistency therefore infuses many of the overarching themes in *Seals*.

Adjudicators acknowledge Canada's consistency argument and describe it as follows:

"By suggesting that the European Union must recognize the same level of animal welfare risk in seal hunts as it does in its slaughterhouses and terrestrial wildlife hunts, Canada appears to argue that a responding Member must regulate similar public moral concerns in similar ways".²⁵

The academic debate, from both law and philosophy, suggests that consistency was a central issue in *Seals*. Importantly, Canada explicitly raised consistency in the dispute²⁶ and adjudicators acknowledged that it was a matter which required their assessment.

The discussion of regulatory consistency by adjudicators should have clarified (i) whether the EU was subject to a consistency requirement²⁷; (ii) where that

²² ibid 5.196

²³ ibid 5.194.

²⁴ The Appellate Body's equivocal understanding of these themes was addressed in Chapter 11.

²⁵ Appellate Body Reports, *Seals* (n 21) para 5.200.

²⁶ As with Perišin and Sellheim, Canada is vague about precisely where the consistency obligation arises in WTO law.

²⁷ If adjudicators expressly determined that there was no consistency requirement, they would not be required to address the subsequent questions.

requirement is located in WTO law (Article XX, “moral standards” or “welfarism”)²⁸; and (iii) whether the EU violated the consistency requirement. I will show in the next three sections that adjudicators failed to provide clear answers to any of these questions.

12.2 Avoiding the Consistency Requirement

This section will consider several techniques which adjudicators (and scholars) used to avoid the consistency question. It will consider (i) the Appellate Body’s failure to clarify whether the EU’s policy objective is welfarist or intuitionist,²⁹ (ii) the Appellate Body’s use of a “same treatment” rather than “regulatory consistency” requirement; and (iii) the *amici*’s claim that Canada’s argument is hypocritical.

12.2.1 Silence on Consistency and Policy likeness

The equivocal characterisation of the policy objective as “seal welfare” makes it unclear whether the EU’s defence is welfarist or intuitionist.³⁰ This ambiguity significantly affects the consistency analysis as each defence has a different view on whether seals and commodity animals are policy-like. Welfarism considers all animals policy-like and demands consistent treatment. The intuitionist claim, on the other hand, allows special treatment for charismatic species, such as seals, on the grounds they are not policy-like commodity animals.

Since the consistency analysis will vary for each claim, adjudicators must adopt a clear view on whether seals are policy-like commodity animals. In other words, the question of “policy likeness” serves as a proxy for whether the EU’s defence is

²⁸ If adjudicators determined that welfarism contains a consistency requirement, they would have been within their rights to exercise judicial economy on the broader questions related to consistency under Article XX or the public morals exception.

²⁹ This can also be understood as a question about whether seals are “policy-like” commodity animals or not.

³⁰ This matter was analysed in chapter 11.

welfarist or intuitionist. Unfortunately, the Appellate Body seeks to remain silent on this question. It states:

“Even if Canada were correct that the European Union has the same moral concerns regarding seal welfare and the welfare of other animals...”³¹

The Appellate Body refuses to take a clear position on whether seals raise “the same moral concerns” as commodity animals. This ultimately means it lacks a fixed view on whether the EU’s defence is welfarist or intuitionist. The Appellate Body suggests that Canada is the proponent of the view that seals and other animals are “policy-like”, but this misrepresents the situation. The EU was the proponent of “policy-likeness” each time it made arguments based on welfarism and the Suffering/Utility Standard (moral philosophy), the OIE Guidelines (internationalist evidence) or examples of animal welfare measures around the world (transnationalist evidence).

Adjudicators accepted these welfarist arguments to determine that animal welfare was “important” and based on “moral standards”.³² If seals are not policy-like commodity animals, why did the Appellate Body regularly refer to “animal welfare” (instead of “seal welfare”) and accept evidence which relates to commodity animals but not seals (such as the OIE Guidelines)?

By contrast, when Canada attacked the EU’s inconsistent application of the Suffering/Utility Standard, adjudicators questioned whether it was useful to compare seals with commodity animals. It claimed to be silent on whether seals and commodity animals are policy-like.

The Appellate Body had no right to be silent on whether seals are policy-like commodity animals as this was a key matter under dispute. It was under an

³¹ Appellate Body Reports, *Seals* (n 21) para 5.200.

³² This matter was analysed extensively in Chapter 11.

obligation to clearly characterise the policy objective which meant that it had to have a clear view about policy-likeness. The Appellate Body's vagueness about policy-likeness enables it to oscillate between a welfarist and intuitionist interpretation of seal welfare.

Adjudicators allowed the EU to rely on welfarist evidence related to commodity animals to prove its claims. However, when Canada sought to expose the weakness of the EU's welfarist defence, the Appellate Body seemed to suggest that it was not useful to consider evidence about how commodity animals are treated. How can animal welfare laws be admissible as evidence to support the EU, but inadmissible for the purpose of challenging its defence?

Due to the Appellate Body's equivocation on the policy objective, its findings on consistency are unclear. One interpretation is that regulators are not required to consistently apply their welfarist measures to all commodity animals. However, an equally plausible interpretation is that adjudicators rejected Canada's calls for consistency on the basis that seals do not raise the "same moral concerns" as commodity animals. This second interpretation would suggest that there may be an obligation to apply animal welfare laws consistently, but that such laws do not apply to seals as they are not a commodity species.

12.2.2 The "Same Treatment" Argument

The Appellate Body's avoidance of the consistency question is aided by a further technique where it misrepresents the consistency requirement as a "same treatment" requirement. It states:

"Even if Canada were correct that the European Union has the same moral concerns regarding seal welfare and the welfare of other animals... we do not

consider that the European Union was required by Article XX(a), as Canada suggests, to address such public moral concerns in the same way".³³

The *amici* replicate the Appellate Body's "same treatment" approach:

"Canada's argument implies that the European Union does not have the right to treat certain animals differently from others if it wishes to do so".³⁴

Adjudicators and the *amici* misrepresent Canada's welfarist argument (that all commodity animals should be regulated consistently) as a "same treatment" argument.³⁵ In reality, consistency and "same treatment" are different concepts. To illustrate this point, I will highlight two scenarios (an intuitionist and a welfarist one) where different treatment could be legitimate, even where a consistency requirement is present.

Intuitionist example

A consistency requirement would not prevent the EU from introducing species-wide bans for non-welfarist policy reasons related to endangered species,³⁶ charismatic species³⁷ or religious bans.³⁸ It would merely require consistent treatment within the family of animals that the EU treats as commodities and therefore policy-like.

If the EU was prepared to explicitly state that its defence was intuitionist, it would be liberated from Canada's consistency requirement. However, it would also have to renounce the use of welfarist arguments and fully defend its measure based on

³³ Appellate Body Reports, *Seals* (n 21) para 5.200.

³⁴ Robert Howse, Joanna Langille and Katie Sykes, 'Pluralism in Practice: Moral Legislation and the Law of the WTO after Seal Products' (2015) 48 George Washington International Law Review 81, 115.

³⁵ In critical reasoning, this is the "straw man" fallacy.

³⁶ For example, the EU argues that "the United States bans imports of fins of sharks" in order to "improve shark conservation" under the Shark Conservation Act; EU, 'First Written Submission by the European Union' (EC – Seals, 21 December 2012)

<https://trade.ec.europa.eu/doclib/docs/2012/december/tradoc_150190.pdf> (consulted 5 April 2020) ("EU First Panel Submission") para 75.

³⁷ Such as the dog and cat product bans referred to by the EU; *ibid*.

³⁸ Such as the EU's claims that "many Islamic countries ban trade in meat products from animals which have not been killed according to the Halal rules" and that "Israel bans the importation of non-kosher products"; *ibid*.

intuitionism.³⁹ It could not invoke welfarism, Singer, the Suffering/Utility Standard, the OIE Guidelines or *foie-gras* bans.

Welfarist example

Alternatively, if adjudicators accepted that seals are a commodity species, Canada is correct that a consistency requirement would require them to establish “a precise standard of animal welfare in the European Union” and undertake “an assessment of the incidence of suffering in commercial seal hunts against that standard”.⁴⁰

This would require the EU to seek the same level of suffering for seal slaughter as it does for commodity animals, but it would not require the “same treatment”. The regulatory treatment for seals could vary due to contextual differences such as slaughter methods, the operating conditions or the physical features of seals.⁴¹

This is how animal welfare laws tend to operate in practice. For example, regulators might pursue the same goal for a variety of different commodity animals (such as pain-free slaughter by rendering the animal unconscious), but permit different treatment for different species in order to achieve that goal. For example, the techniques and modalities for stunning may vary across species (depending on their size and physiology), but this different “treatment” does not necessarily amount to regulatory inconsistency.

When adjudicators affirm that the EU is not obliged to treat all animals “the same way”, this is not necessarily a rejection of consistency requirements. It could mean

³⁹ This defence was analysed in detail in chapter 10 where I concluded that for an intuitionist defence to succeed, the EU would have to convince adjudicators to adopt “evidentiary unilateralism” for Article XX(a) or at least for a subset of non-instrumental claims.

⁴⁰ Appellate Body Reports, *Seals* (n 21) para 5.196.

⁴¹ For example, the EU argues that “seals have unique anatomic and physiological features that enable them to stay under water for very long periods. As a result, seals may experience suffering in ways which are peculiar to that species”; EU First Panel Submission (n 36) para 407.

(i) that they are treating “seal welfare” as the intuitionist claim or (ii) that the EU is required to apply its moral standards consistently, but that this requires different treatment for seals due to important contextual difference. In Section 12.4, I will explain why this justification for the seal products ban is not plausible.

12.2.3 The Hypocrisy Argument

The *amici* produce a further argument against consistency testing. They accuse Canada of hypocrisy because its proposed consistency requirement:

“flies in the face of Canada’s own legislation on animal welfare, which in practice provides protection to cats and dogs from cruelty but not other animals with a similar level of sentience”.⁴²

This argument is superficially-appealing: if banning dog products is morally-desirable, but consistency requirements would somehow render dog product bans WTO-inconsistent, the *amici* conclude that consistency requirements have no place in WTO law.

This mode of thinking is the antithesis of rationalism as it starts with a moral conclusion (dog product bans must be legitimate) and reasons backwards to determine that consistency requirements could not possibly exist in WTO law.

Singer would abhor such thinking. He would start with moral principles (including consistency) and reason forwards to see what conclusions they produce, even if those conclusions are contrary to common intuitions (dog product bans are not morally-justifiable).

Singer’s approach seems more akin to the rationalist view of legal decision-making where judges start with principles and reason forwards, rather than seeking to justify a conclusion that was reached intuitively. Either consistency requirements exist or

⁴² Howse et al (n 34) 115.

they do not. If they render dog product bans WTO-inconsistent, then so be it. It may seem intuitively appealing to many observers that dog product bans should be allowed, but this does not mean that they are legitimate under WTO law.

Perhaps the non-challenge of dog product bans at the WTO is a matter of tacit acceptance and diplomatic agreement, rather than a confirmation that they are legally justified under Article XX(a).⁴³ In fact, in the absence of evidentiary unilateralism, I doubt there is a legal basis for intuitionist measures like dog product bans.⁴⁴

In any case, it is far from certain that Canada's argument is hypocritical⁴⁵ or that a consistency requirement would lead to a flood of challenges to dog product bans. If Canada is prepared to acknowledge that its dog product ban is intuitionist (and refrains from misrepresenting such bans as welfarist), the amici's hypocrisy argument ceases to apply. Species-wide bans for charismatic species are only hypocritical if you pretend that they were adopted for welfarist reasons, as the EU does when it hides its (intuitionist) seal products ban behind a veneer of welfarism.

12.3 – Rejecting Consistency Requirements

12.3.1 Adjudicators Arguably Reject Consistency

In addition to the “avoidance” strategy described above, the Appellate Body also adopted a second approach to consistency when it made statements which have

⁴³ It's worth noting that measures which are inconsistent with WTO rules can survive provided that no Member seeks to challenge them. This does not necessarily mean that they are WTO-consistent.

⁴⁴ The Suffering/Utility Standard cannot justify a species-wide ban of dog products. Such bans could be justified by evidentiary unilateralism (for those Western societies where this social norm exists), but the Appellate Body has yet to recognise unilateralist approaches under Article XX(a). That said, if a dog product ban was challenged at the WTO, I strongly suspect that adjudicators would find a legal basis for its legitimacy, even though this would require a significant change to the current jurisprudence.

⁴⁵ In critical reasoning, it is well recognised that accusations of hypocrisy (*the tu quoque fallacy*) rarely amount to a compelling argument.

been interpreted by some scholars as a rejection of consistency requirements. The Appellate Body stated:

“Members have the right to determine the level of protection that they consider appropriate, which suggests that Members may set different levels of protection even when responding to similar interests of moral concern”.⁴⁶

Based on this statement, Lester posted a blog entitled “There’s No ‘Consistency’ Requirement for Animal Welfare and Public Morals”.⁴⁷ Lester’s language might be excessively categorical, but he is correct that, if Members may vary their level of protection in policy-like situations, this is tantamount to rejecting consistency requirements.

However, there are reasons to be cautious about Lester’s conclusion. Apart from this one statement, everything else about the Appellate Body’s analysis is highly equivocal on the question of consistency.⁴⁸

The Appellate Body’s supposed declaration that there is no consistency requirement is made vaguely and softly. It appears in a single sentence. The Appellate Body does not mention the term “consistency”. It formulates its view as a “suggestion” rather than as a declaration or clear conclusion.

This supposed rejection of consistency requirements is so indirect that many observers seem to have overlooked it (though Lester astutely noticed it). The statement does not seem sufficiently powerful that it would tie adjudicators’ hands in future cases. While adjudicators seemed to avoid consistency analysis in *Seals* (at

⁴⁶ Appellate Body Reports, *Seals* (n 21) para 5.200.

⁴⁷ Simon Lester, ‘There’s No “Consistency” Requirement for Animal Welfare and Public Morals’ (International Economic Law and Policy Blog, 22 May 2014) <<https://ielp.worldtradelaw.net/2014/05/theres-no-consistency-requirement-for-animal-welfare.html>> accessed 15 January 2020.

⁴⁸ For example, section 12.2 described a variety of poor reasoning techniques which sought to avoid the consistency question altogether. Section 12.4 will show that the Appellate Body engaged in some analysis which strongly resembles consistency testing.

least for some matters), the reasoning does not suggest that adjudicators sought to categorically reject consistency testing in WTO law.

A further problem with the Appellate Body's consistency statement is that the reasoning is poor. It confirmed that Members have the right to set their own level of protection (a non-controversial claim). It then jumped to a further conclusion that Members "may set different levels of protection" even when addressing "similar" (policy like) situations.

It is unclear why the Appellate Body treats the "level of protection" and "consistency" as linked concepts? At face value, they appear to address two different matters. This is why there is a lively scholarly debate about consistency, even though the right of Members to set their level of protection is well-settled in WTO law and the academic literature.

The *amici* endorse the Appellate Body's approach:

"the AB rejected Canada's [consistency] argument, stating that Member states have a right to set the level of protection that they desire".⁴⁹

The *amici* share the Appellate Body's view that consistency requirements are linked to the level of protection and therefore have the potential to impinge on the right to regulate. If this is true, consistency requirements would pose a serious threat to the WTO's model based on "negative integration" and "pluralism".

However, it is not convincing that consistency requirements pose a threat to regulatory autonomy. The "pluralism" principle allows each Member to choose its

⁴⁹ Howse et al (n 34) 114.

own preferred regulatory approach, including the EU's choice to supposedly pursue high animal welfare standards.⁵⁰

When Canada (and Perišin) invoke consistency requirements, they do not suggest that the EU must adopt a lower level of animal welfare protection to be compliant with WTO law. They accept that each Member has a sovereign right to set its own level of protection.

However, they call on the EU to apply its chosen level of protection consistently to all commodity animals. They deny the EU's right to have a different Suffering/Utility Standard for cows than they do for pigs. If the EU considers seals a commodity species, they would also have to apply the same animal welfare standard to seals.

In this sense, Canada's consistency requirement calls on the EU to respect its own EU-endorsed level of protection. It does not seek to impose on the EU an external norm about how much animal welfare protection is legitimate.

The *amici* further clarify their position when they state that:

"it is not the WTO's place to second guess the appropriate level of protection for each species".⁵¹

This argument betrays the *amici*'s position that consistency requirements have no place in WTO law. While the *amici* strongly oppose a single level of protection for all animals, they appear to advocate a consistency requirement at the species level.

This argument (that regulators should pursue consistency at the species level) contradicts the *amici*'s own views. In the moral standards discussion, the *amici* suggest that animal welfare laws are consistent (and therefore standards-based)

⁵⁰ Though, as discussed in Chapter 9, welfarists would consider the EU's standards quite low.

⁵¹ Howse et al (n 34) 115.

when they allude to a single “level of protection against animal suffering”. When citing Singer and Bentham on this matter, why did the *amici* fail to clarify their view that morality requires a different level of protection for each species?⁵²

If the reader’s reference point is a seal or dog, it may seem intuitively obvious that there should be a different level of protection for each species. However if we exclude charismatic species from the analysis, and focus solely on commodity animals, does it make sense to have a different standard for each species?

Should we offer cows, sheep and pigs different levels of protection based on our intuitive feelings about their suffering? When the matter is reframed in these terms, it seems less intuitively obvious that there should be different levels of protection.

In practice, the EU and other governments tend to take a comprehensive and consistent approach to the regulation of commodity animals. For example, the EU Regulation on animal slaughter seeks to establish “minimum rules for the protection of animals at the time of slaughter”.⁵³ It does not establish separate levels of protection for each species. The Regulation states:

“Business operators or any person involved in the killing of animals should take the necessary measures to avoid pain and minimise the distress and suffering of animals during the slaughtering or killing process”.⁵⁴

The EU seeks to regulate all “vertebrate animals” as part of a single package and does not establish separate principles or benchmarks for each species.⁵⁵ Where it

⁵² Under a welfarist framework, this argument lacks credibility, though it might be reasonable under an intuitionist framework.

⁵³ Council Regulation (EC) No 1099/2009 of 24 September 2009 on the protection of animals at the time of killing (2009) OJ L 303/1 preamble (first recital). I use this Regulation as an example as the Appellate Body chooses “effective stunning rates” at the time of slaughter as a benchmark for humane treatment of animals (a matter I will discuss in section 12.4).

⁵⁴ *ibid* preamble (second recital).

⁵⁵ *ibid* preamble (nineteenth recital).

chooses to exclude a species from the Regulation, it states this explicitly and offers a policy justification for the exclusion. For example:

“Recommendations on farm fish are not included in this Regulation because there is a need for further scientific opinion and economic evaluation in this field”.⁵⁶

If the EU and *amici* wish to offer a higher level of protection to seals, it is disingenuous to pretend that it is normal to have different moral or regulatory standards for each species. Welfarist approaches, in moral philosophy and in practice, tend to take broadly consistent approaches to commodity animals. The special protection of seals cannot be justified using welfarist standards. It can only be explained under an intuitionist view that seals are special.

The reasons put forward for rejecting Canada’s consistency arguments seem unconvincing. While it is well-established that Article XX allows Members to choose their preferred level of protection, there is no well-established principle under Article XX that Members have a right to vary the level of protection among “policy like” situations. If the Appellate Body wished to establish this new principle (and therefore reject consistency requirements), it should have done so clearly and explicitly.

In my view, Lester overstates the *Seals* finding on consistency. The Appellate Body’s supposed finding against consistency is based on poor reasoning and is expressed in a highly-indirect manner. The *Seals* dispute has not dealt a death blow to consistency testing.

12.3.2 Academic Arguments for Rejecting Consistency

The *amici* provide further reasons for rejecting consistency. They argue WTO law should not require Members to be “fanatical about achieving [their] moral purposes”⁵⁷

⁵⁶ *ibid* preamble (sixth recital). The Regulation further suggests that research on the stunning of fish is far less developed than for other farmed species (eleventh recital).

or “perfectly philosophically consistent” regarding their morally-motivated legislation”.⁵⁸

Building on their “fanatical consistency” argument, the amici suggest that consistency requirements would hamper “any incremental change of moral positions through legislation”.⁵⁹ In their view, consistency requirements would freeze piecemeal animal welfare reform.

It is important to note that the *amici*’s arguments appear to be theoretical in nature, rather than addressing the actual scenario in *Seals*. After all, the seal products ban was not an “incremental” reform being stifled by “fanatical” consistency requirements. Quite the contrary, it represented a drastic reform which deviated markedly from the EU’s approach to commodity animals. An observer need not be “fanatical” to notice the inconsistency between the EU’s treatment of seals and cows.

Nevertheless, if we consider that the *amici* are making a systemic claim (unrelated to *Seals*), their argument merits attention. It is instructive to consider a real world scenario to test this claim. Let’s imagine that, since animals are sentient beings, a government determines that they should not be subjected to the painful removal of body parts without anaesthetic.

This principle would not apply to painless practices (akin to a human cutting their hair or clipping their nails), but would apply to painful practices such as tail docking of sheep, dehorning of cattle, debeaking of chickens or castration of pigs. Under the Suffering/Utility Standard, welfarists would insist that these practices can only be moral if they are truly painless (with the aid of anaesthetic, for example).

⁵⁷ Howse et al (n 34) 99.

⁵⁸ ibid 115.

⁵⁹ ibid 145.

Welfarists would also require that the principle be applied consistently to each of the affected species. They would denounce a government that protected cows from painful dehorning but ignored the suffering of pigs being castrated. This welfarist view is probably intuitively-appealing to many people. What possible reason could justify the inconsistency in these scenarios?

The *amici* seem to imply that, in the real world, animal protection is implemented through piecemeal reforms (one-species-at-a-time). They suggest that a consistency requirement would require governments to take an all-or-nothing approach and that, in practice, governments would choose nothing. Reform would be stifled and animals would ultimately be the victims of a consistency requirement. Is their assessment plausible?

This is ultimately an empirical matter, however the EU's approach to animal welfare, including its slaughter rules, seems to indicate that comprehensive packages are realistic. Indeed, consistency requirements could raise welfare standards by requiring governments to be more comprehensive in their reforms. It might push them up to "all" rather than down to "nothing".

Let's suppose that California's recent measure to phase out "extreme methods of farm animal confinement" was subject to a legal challenge, would California fail a consistency test and would this lead to a lowering of standards?⁶⁰ If the regulation was egregiously inconsistent – such as banning cage eggs but allowing products

⁶⁰ California Proposition 12, 'Farm Animal Confinement Initiative' (approved on 6 November 2018) <[https://ballotpedia.org/California_Proposition_12,_Farm_Animal_Confinement_Initiative_\(2018\)](https://ballotpedia.org/California_Proposition_12,_Farm_Animal_Confinement_Initiative_(2018))> accessed 5 April 2020 ("California Proposition 12"). This law is being challenged by a number of other US States, including for claims that it breaches the Constitution's Dormant Commerce Clause; David Lieb, '13 States Launch New Legal Challenge to California Egg Law' (US News, 4 December 2017) <www.usnews.com/news/best-states/california/articles/2017-12-04/12-states-launch-new-legal-challenge-to-california-egg-law> accessed 3 February 2020.

from pig gestation crates or veal cages – it is conceivable that California could lose a consistency challenge.

From a moral point of view, it makes sense that a regulation aimed at preventing cruel confinement should protect pigs and calves, not just chickens. Under a legal consistency requirement, such egregiously-inconsistent measures might fail a WTO challenge.

Let's imagine that California lost a consistency challenge on this basis, would this lead to lower animal welfare standards? It would depend on how California pursued compliance. It could achieve compliance by allowing cage eggs again (thus decreasing welfare) or by banning gestation crates and veal cages (thus increasing welfare). In this sense, consistency requirements are value-neutral. Whether they drive standards up or down is ultimately a policy decision for the regulator.

In practice, California's rules are comprehensive (and consistent) as they seek to create minimum space requirements for each species. California has created a consistent principle by banning any confinement practice which "prevents the animal from lying down, standing up, fully extending the animal's limbs, or turning around freely".⁶¹ This does not mean that the requirements are identical ("same treatment") for each animal. Under California's rules, larger species are (logically) entitled to more space.

Where individual species have specific regulations, how do you determine if the regulations are consistent? It would seem reasonable that cows should be entitled to more space (or more powerful stun guns) than smaller animals. But at what point

⁶¹ California Proposition 12 (n 60) Section 4(e).

does different treatment amount to regulatory inconsistency? Where should adjudicators draw the line in practice?

This is where adjudicators must establish a standard of review. I agree with the *amici* that “fanatical consistency” would be excessive, but why not a standard of “reasonable consistency” which provides some deference to regulators? This is how adjudicators approach other matters such as necessity (where they do not require a measure to be “fanatically” necessary). The *amici*’s opposition to “fanatical consistency” is legitimate, but they fail to make a compelling case against a reasonable or moderate consistency requirement. Moral philosophers and regulators seem to concur that consistency is a key component of animal welfare.

12.3.3 Summary

The “fanatical” consistency argument is not relevant to *Seals* as the EU’s ban was a drastic (not “incremental”) reform. In any case, consistency requirements are unlikely to chill animal welfare reform in practice. Much welfare reform is pursued comprehensively (and consistently) rather than in a piecemeal fashion.

There is no reason why measures would fail a consistency challenge provided they are “reasonably consistent”. Even if a regulator lost a case, it could implement by improving welfare for those animals which did not benefit from the first round of reforms (such as adding a ban on gestation crates and veal cages).

12.4 Applying the Consistency Requirement

Adjudicators adopted a third approach to consistency when they assessed whether there were any contextual factors which justified the EU’s application of a higher level of protection to seals. This analytical approach resembled Singer’s conception

of welfarism as it included (i) the Suffering/Utility Standard and (ii) a consistency requirement.⁶² Is it plausible for the EU to adopt a coherent and consistent moral standard for all animals which allows the suffering that occurs on factory farms but requires a species-wide ban for seals?

The EU pursues this argument using two formulations: (i) that seals suffer more than commodity animals and (ii) that seal hunting is “inherently inhumane”. I’ll analyse both formulations, including the adjudicators’ response.

12.4.1 Seals Suffer more than Commodity Animals

The EU seeks to demonstrate that seal exploitation is less humane than the exploitation of commodity animals:

“Unlike the animals stunned in a slaughterhouse, which are restrained and immobile, seals are freely moving targets and can react in unpredictable ways when alarmed by an approaching seal hunter”.⁶³

It cites scientific evidence that seal hunting is riskier and less humane:

“Veterinary advice and the regulations that result from it have thus not solely focused on how to ensure that the killing is humane (as required in established commercial slaughter), but rather how to make it less inhumane by adopting methods that are practical on the ice (but which would be considered primitive in a slaughterhouse on land)”.⁶⁴

For these arguments, the EU engages in direct comparisons between seal hunting and commodity animal slaughter.⁶⁵ It refers to “a slaughterhouse on land”, and “animals stunned in a slaughterhouse”.⁶⁶

⁶² Though Singer would still question whether the EU’s level of protection for factory farmed animals is moral.

⁶³ EU First Panel Submission (n 36) para 127.

⁶⁴ *ibid* para 108.

⁶⁵ It should be recalled from section 12.3 that Canada was denied the right to make arguments comparing seals with commodity animals.

⁶⁶ In the interest of completeness, it is worth mentioning that the EU makes arguments that seals are different because of their physical features, such as the following: “seals have a number of physiological and anatomical adaptations that bring into question whether ‘conventional’ thinking on slaughter can be applied to these

Under a welfarist claim, it is reasonable to seek to justify higher protection by showing that a particular practice leads to a higher degree of suffering. For example, this is the logic which underpinned California's welfarist ban on *foie-gras* even though it allowed other types of goose exploitation.

However, despite its theoretical logic, this argument poses a different problem for adjudicators. It requires them to accept that the EU has the same moral concerns for seals and commodity animals (they are "policy-like") and that it is necessary to compare their treatment. Section 12.2 already showed that adjudicators rejected Canadian attempts to compare seals with commodity animals.⁶⁷

Canada argued that the EU has "a tolerance for a certain degree of animal suffering, both for slaughterhouses and wildlife hunts".⁶⁸ It called on adjudicators to "identify a relevant standard or benchmark of animal welfare" and determine if the risks associated with seal slaughter "exceeded the accepted level of risk of compromised animal welfare".⁶⁹

Canada asked adjudicators to acknowledge that factory farmed animals suffer, to assess how much suffering occurs on factory farms and to determine if seal slaughter is worse than factory farming. Canada's request seemed like a logical and rigorous way to test for consistency, however adjudicators rejected it. They made their famous statement (highlighted by Lester) denying the utility of consistency testing:

"Even if Canada were correct that the European Union has the same moral concerns regarding seal welfare and the welfare of other animals... we do not

⁶⁷ "unusual animals"; EU First Panel Submission (n 36) para 96. This argument was not pursued in any meaningful way and adjudicators ignored it.

⁶⁸ This matter was also addressed in Chapter 11.

⁶⁹ Appellate Body Reports, *Seals* (n 21) para 5.194.

⁶⁹ *ibid.*

consider that the European Union was required by Article XX(a), as Canada suggests, to address such public moral concerns in the same way".⁷⁰

However, when the EU argues that seal slaughter is less humane than commodity animal slaughter (a consistency argument in the EU's favour), adjudicators appear willing to consider it. Adjudicators compare seals to commodity animals and determine that seal slaughter produces more suffering:

"The evidence did not establish comparable effective stunning rates in seal hunts and commercial abattoirs".⁷¹

This finding seems to agree with Canada's earlier arguments that the EU tolerates the suffering of commodity animals, that adjudicators should identify a "relevant standard or benchmark" (effective stunning rates) and that they should determine if seal slaughter is truly less humane. While adjudicators were not prepared to undertake such analysis at Canada's request, they were prepared to do it in favour of the EU.

However, there is a serious gap in the adjudicators' approach to consistency testing. They seem to have reached a conclusion that seal slaughter is less humane, but they have not really made the effort to understand how much suffering occurs on factory farms. How can you determine that seal slaughter is less humane when you have no facts on the record which identify how inhumane factory farming is?

In this sense, the Appellate Body's methodology is deeply-flawed. It is implicit in the Appellate Body's statement that some factory farmed animals are slaughtered without being effectively stunned, but it does not state what the percentage is. It has no data or sources to support its claim that seal slaughter has lower stunning rates.

⁷⁰ ibid para 5.200.

⁷¹ ibid para 5.278. It further adds that "the two situations differ significantly in areas of great relevance to the application of humane killing methods".

Perhaps adjudicators simply assume that stunning in slaughterhouses must be highly effective and that it entails negligible suffering. However, studies suggest that slaughterhouses struggle to comply with rules on humane slaughter. Workers have testified that 30% of cows in slaughterhouses are not properly stunned which leads to them being “routinely scalded, bled, skinned, dismembered and/or eviscerated while awake and fully conscious”.⁷²

A recent academic study from Sweden suggested that 12.5% of commercially-slaughtered cattle were not stunned effectively.⁷³ Adjudicators also failed to make a factual finding on the “effective stunning rates” for seal slaughter, though Sellheim cites a study claiming that 42% of seals are skinned while conscious.⁷⁴

It is doubtful that “effective stunning rates” truly drive EU animal welfare policy. The EU imposes no regulations which mandate minimum effective stunning rates. Even though “it should be possible to ensure adequate stunning in almost 100% of animals”, studies suggest ineffective stunning ranges from 9-35% in commercial slaughter.⁷⁵ If the EU is concerned by animals being slaughtered while conscious, why does it fail to effectively implement effective stunning rates in slaughterhouses?

Further, EU rules on livestock and meat statistics do not mandate any data or statistics on effective stunning rates.⁷⁶ If the EU’s public concerns are based on “effective stunning rates”, why are there no minimum benchmarks or data collection on this matter?

⁷² David Simon, *Meatonomics* (Conari Press 2013) 48.

⁷³ Sophie Atkinson, Antonio Velarde and Bo Algers, ‘Assessment of Stun Quality at Commercial Slaughter in Cattle Shot with Captive Bolt’ (2013) 22 Animal Welfare 473, 473.

⁷⁴ Sellheim (2016), p144. However, it is important to note that Sellheim questions the impartiality of the science on seal hunting.

⁷⁵ Atkinson et al (n 73) 473.

⁷⁶ See eg Regulation (EC) No 1165/2008 of the European Parliament and of the Council of 19 November 2008 concerning livestock and meat statistics (2008) OJ L 321/1.

For farmed animals, the EU appears to be far more nonchalant about “effective stunning rates” than for seals. In the US, chickens are excluded from the legislation requiring stunning before slaughter.⁷⁷ US chicken products entering the EU have a far worse stunning record than Canadian seals. Why doesn’t the EU ban US chicken alongside Canadian seal products?⁷⁸

The Appellate Body appears (intuitively) partial to the EU’s “consistency” claim that seal slaughter is less humane. However, it appears unwilling to obtain any empirical data to support this claim, especially any fact-finding about what actually occurs in factory farms and commercial abattoirs. What would an empirical investigation reveal?

This is a poor methodology for comparison and consistency testing. Canada’s proposed method seemed far more rigorous and appropriate. However, the psychology literature suggests that the Appellate Body’s approach reflects how most humans “reason” about animal exploitation.

Psychological Barriers to Logical Reasoning

Comparing seal suffering to factory farm suffering should not be difficult as a factual matter. After all, animal suffering is an instrumental matter which can be assessed using science and risk assessment techniques. However, for consumers of animal products (and perhaps WTO adjudicators), it can be psychologically difficult to impartially consider the suffering which occurs on factory farms. Many humans

⁷⁷ Simon (n 72) 47 where he discusses the Humane Methods of Slaughter Act (which also excludes other poultry and fish).

⁷⁸ Further, why does the Appellate Body group all commodity animals together for the purpose of identifying a single “effective stunning rate”? If the Appellate Body agrees with the *amicus*’s view that animal welfare is species-specific, why does it fail to identify the effective stunning rates for a range of different species (rather than grouping all commodity animals under a single benchmark for stunning)? Perhaps seals have a higher effective stunning rate than some species which are slaughtered in abattoirs.

adopt non-rational “reasoning” techniques which prevent them from finding out how much suffering occurs in slaughterhouses and factory farms. This section will consider “avoidance” and “denial of animal pain”.

Avoidance

Many people simply avoid information about the suffering of commodity animals. In a survey on meat-eating, “67% of respondents indicated that they do not think about animal suffering in factory farms when they purchase meat”.⁷⁹ Meat-eaters deliberately avoid visiting abattoirs, watching documentaries about meat production or otherwise obtaining information about where their meat comes from.

The avoidance strategy makes the suffering of animals (intentionally) invisible. As Stewart and Cole highlight, “farming has led to a progressive removal of animals from public view, through relocating farms, increasing security measures and the use of the laws of trespass to inhibit the exposure of cruel practice”.⁸⁰ Many governments, including European Member States, have passed laws aimed at making it impossible for people to know what happens in slaughterhouses.⁸¹

Denial of animal pain

Rothgerber has demonstrated that there is a correlation between humans who consume more meat and the perception that “animals don’t really suffer when being

⁷⁹ Hank Rothgerber, ‘Efforts to Overcome Vegetarian-Induced Dissonance among Meat Eaters’ (2014) 79 *Appetite* 32, 33.

⁸⁰ Kate Stewart and Matthew Cole, ‘The Conceptual Separation of Food and Animals in Childhood’ (2009) 12(4) *Food, Culture & Society* 457, 461.

⁸¹ Simon (n 72) 31 refers to these types of “ag-gag” laws when he notes that the animal agriculture industry has been “remarkably successful at passing laws that prevent [critics of factory farms] from investigating, criticising or suing them”.

raised or killed for meat".⁸² A study by Piazza et al further established that meat-eaters tend to deny animals' "capacity to suffer or experience pleasure".⁸³

In other words, people first decide whether they consider a species a commodity and then reason backwards to decide whether they believe that animal suffers.⁸⁴ You would think that whether and how much animals suffer is a factual question which a court, policymaker or consumer could determine empirically.⁸⁵ The psychology literature suggests that for most people animal suffering is an emotional question rather than an empirical one.

Jacques Cousteau summarised this phenomenon when he stated:

"The harp seal question is entirely emotional. We have to be logical... Those who are moved by the plight of the harp seal could also be moved by the plight of the pig - the way they are slaughtered is horrible".⁸⁶

If the Seals adjudicators (or academic commentators) structure their reasoning to render the suffering of farmed animals irrelevant or invisible, they may be engaging in the judicial/intellectual equivalent of "avoidance". Should judges rely on common (but misguided) assumptions about what happens in slaughterhouses or should they undertake a fact-based interrogation of the reality? Should they use System 2 thinking or give in to their powerful System 1 intuitions?

In my view, the adjudicators' finding that seal slaughter is less humane than factory farming has two major flaws. First, adjudicators applied a double standard by

⁸² Rothgerber (n 79) 33.

⁸³ Jared Piazza, Matthew Ruby, Steve Loughnan, Mischel Luong, Juliana Kulik, Hanne Watkins and Mirra Seigerman, 'Rationalizing Meat Consumption: the 4Ns' (2015) 91 *Appetite* 114, 117.

⁸⁴ This is similar to Posner's technique of separating out the commodity animals who are then subject to a sentience (rather than sentiment) based normative framework.

⁸⁵ In Chapter 11, I showed that the Appellate Body, EU and *amicus* all argued that animal suffering is an instrumental matter which can be determined by reference to science and risk.

⁸⁶ Jacques Cousteau cited in Nikolas Sellheim, 'Policies and Influence: Tracing and Locating the EU Seal Products Trade Regulation' (2015) 17 *International Community Law Review* 3, 24.

rejecting Canada's arguments based on comparison, while accepting the EU's right to make such arguments.

Second, it undertook no analysis of what happens on factory farms and provided no data to support the claim that seals suffer more.⁸⁷ Empirically, it is impossible to tell whether this conclusion is true or false, although many observers of the dispute might have a strong intuition that seal slaughter must entail a higher degree of suffering..

Even if seal slaughter is less humane, this finding could only justify higher standards of protection for seals, but it could not justify a species-wide ban. The next section will consider a further claim by the EU which seeks to justify the seal products ban by arguing that seal slaughter is not only less humane, but that it is “inherently inhumane”.

12.4.2 Seal Slaughter is “Inherently Inhumane”

The EU makes a further, highly-dubious claim that seal hunting is “inherently inhumane”. It states:

“Commercial seal hunts are inherently inhumane because humane killing methods cannot be effectively and consistently applied in the field environments in which they operate”.⁸⁸

“Canada’s commercial seal hunt can never be made acceptably humane because of the conditions in which the hunt takes place”.⁸⁹

“Deteriorating ice conditions, extreme and unpredictable weather, high winds and ocean swells are all deterrents to humane killing and accuracy in clubbing and shooting, and in timely retrieval of animals in the case of shooting”.⁹⁰

⁸⁷ It also fails to explain how it chose “effective stunning rates” as the appropriate benchmark. This indicator had not been raised previously (for example in the “moral standards” discussion) and there is no justification for the choice. I should note that this choice of standard appears designed to discount the suffering of factory farmed animals as it focusses only on suffering at the time of slaughter, but deems irrelevant all of the other suffering throughout the life cycle of the animal. In the case of seals, the human-caused suffering only occurs at the moment of slaughter.

⁸⁸ EU First Panel Submission (n 36) para 54.

⁸⁹ ibid para 94.

“Even in the case of clubbing it is unlikely the four step killing process can consistently be effectively implemented. Veterinary studies conducted over the past half century confirm this is the case”.⁹¹

The “inherently inhumane” argument plays an important role in the EU’s defence. If true, this claim would liberate adjudicators from having to analyse what occurs in slaughterhouses.⁹² Adjudicators could seemingly determine that seal hunting is “inhumane” in an objective sense, rather than by comparing it to commercial slaughter. This is facilitated by the use of absolute (rather than comparative) language, such as the suggestion that seal slaughter is “inherently inhumane” and “can never be made acceptably humane”. More importantly, if seal hunting is “inherently inhumane”, it means the EU was justified in banning seal products instead of merely regulating them through stringent welfare requirements.

The Appellate Body ultimately accepts the EU’s “inherently inhumane” argument. It upholds the legitimacy of the seal products ban by noting “extensive scientific evidence” which showed “inherent obstacles that make it impossible to kill seals humanely on a consistent basis”.⁹³

There are some serious problems with the “inherently inhumane” argument. First, even though it sounds non-comparative (especially when expressed in absolute terms), all of the EU’s justifications are clearly comparative in nature. It relies on the exact same evidence which it used to argue that seal slaughter was less humane than other types of animal slaughter (“field environments” for seal hunting, “deteriorating ice conditions”, “extreme and unpredictable weather”).

⁹⁰ ibid para 94.

⁹¹ ibid.

⁹² This was one of the major weaknesses considered in the previous section.

⁹³ Appellate Body Reports, *Seals* (n 21) 5.284.

The “inherently inhumane” argument appears to be a mere repackaging and reframing of the “less humane” argument. The EU shifts language so that the comparative nature of the argument is implied rather than explicit. It is unclear how this evidence proves that seal slaughter is “inherently inhumane”. It seems to merely demonstrate that seal slaughter occurs in a less controlled environment than an abattoir.

Second, it is difficult to understand how a practice can be “inherently inhumane” under the EU’s chosen moral standard. The Suffering/Utility Standard is, by its very nature, a relative standard.⁹⁴ Singer, who generally seeks a much higher level of animal protection than the EU, would not consider seal exploitation “inherently inhumane”. He would allow pain-free exploitation and may even allow painful exploitation (where the utility is sufficiently high).⁹⁵

In the alternative measures discussion, the complainants attacked the EU’s explicitly welfarist defence and proposed an alternative based on “strict animal welfare standards”⁹⁶ including “a certification system that would operate to exclude all inhumanely killed seals”.⁹⁷ Canada argued that a species-wide ban was excessive:

“the types of measures applied with respect to the welfare of other animals – including setting animal welfare requirements, certification, labelling, monitoring, and enforcement – raise doubts with respect to the necessity of the more restrictive EU Seal Regime”.⁹⁸

⁹⁴ This is the defining feature of a utilitarian approach. Unlike rights-based approaches which can deem a practice categorically immoral, a utilitarian must weigh and balance all of the positive and negative consequences of an act. This is why Singer concludes that dog exploitation could be justifiable in some circumstances (a view which Posner opposes).

⁹⁵ Interestingly, Singer would allow the painful exploitation of seals by indigenous communities. He states that “Inuit living a traditional lifestyle in the far north where they must eat animals or starve can reasonably claim that their interest in surviving overrides that of the animals they kill. Most of us cannot defend our diet in this way”; Peter Singer, *Practical Ethics* (3rd edition, CUP 2011) 54.

⁹⁶ Appellate Body Reports, *Seals* (n 21) para 2.79.

⁹⁷ *ibid* para 5.267.

⁹⁸ *ibid* para 2.25.

The Appellate Body's "inherently inhumane" conclusion allowed it to dismiss "strict animal welfare standards" as an alternative measure. However, its factual finding that seal slaughter is "inherently inhumane" is highly dubious. It seems plausible that seals can be slaughtered in ways which are no less humane than the standards the EU tolerates for commodity animals. Indeed, the Preamble to the Seal Regulation concedes that "it might be possible to kill and skin seals in such a way as to avoid unnecessary pain, distress, fear or other forms of suffering".⁹⁹

In fact, within the EU, the measure was originally designed as a welfare standard aimed at preventing "avoidable pain" before EU Parliamentary Committees converted it into a "total ban".¹⁰⁰ Perišin suggests that "it is not really clear from all the data taken together why a ban was chosen as the most suitable type of measure".¹⁰¹ She rejects the "inherently inhumane" view and criticises the fact that "the Regulation does not allow countries or individual traders to prove that their seal products are derived from humane hunts".¹⁰²

Sellheim questions the science supporting the seal products ban, suggesting that there had been "politicization" of the process in favour of "a pre-determined stance".¹⁰³ He notes:

"there was a clear tilt towards a seemingly pre-determined opposition towards commercial sealing which favoured a ban before an impact assessment or an assessment of the hunt itself was conducted".¹⁰⁴

⁹⁹ Regulation (EC) No 1007/2009 of the European Parliament and of the Council of 16 September 2009 on trade in seal products (2009) OJ L286/36 preamble (eleventh recital). Although, the Regulation further states that this is "not feasible in practice or, at least... very difficult to achieve in an effective way"

¹⁰⁰ Perišin (n 16) 385. For a detailed discussion of the EU policy and Parliamentary process, see Sellheim (n 86).

¹⁰¹ Perišin (n 16) 387.

¹⁰² *ibid* 398.

¹⁰³ Sellheim (n 86) 35.

¹⁰⁴ *ibid* 19

Sellheim appears highly-sceptical of how the EU defines what is humane or inhumane. He notes that the EU “legalizes killing methods for animals other than seals that are considered cruel for the killing of seals, but not for other species”.¹⁰⁵ While adjudicators conclude that the clubbing of seals is “inherently inhumane”, they seem unperturbed by the fact that clubbing is a legal slaughter method in the EU for piglets and lambs.¹⁰⁶

The “inherently inhumane” argument is rhetorically powerful. It allows adjudicators to ignore what happens on factory farms and to conclude that the species-wide ban was not excessive (as welfare standards are insufficient to address “inherently inhumane” practices). However, the reasoning and fact-finding which underpin the “inherently inhumane” argument suffer from major weaknesses.

Despite the flaws in its reasoning and fact-finding, the Appellate Body’s amenability to the “more inhumane” and “inherently inhumane” arguments is highly-significant. By relying on these arguments, it engages in consistency testing. In order to justify greater (or even total) protection for seals, the EU was required to demonstrate important contextual differences relative to commodity animals. Such arguments would have been redundant if consistency testing was not relevant to the dispute.

12.5 Conclusion

It is difficult to extract a clear message about consistency from the *Seals* dispute. In my view, there are three possible conclusions.

First, adjudicators sought to avoid addressing the consistency question. While they seem to hold a clear view that there is no obligation to treat seals the same as

¹⁰⁵ ibid 26-27

¹⁰⁶ ibid 26-27 (including footnote 98). Sellheim describes this as a “double standard”.

commodity animals, their rationale is not clearly stated. Is it because there is no consistency requirement or because seals and commodity animals are not “policy-like”? If there is no consistency requirement under Article XX(a), adjudicators appear extremely reluctant to state this explicitly.

Second, to the extent that adjudicators suggest that there is no consistency requirement, they do so indirectly and on the basis of weak reasoning. They suggest that the non-existence of a consistency requirement flows logically from the right of regulators to choose their preferred level of protection. This rationale for rejecting consistency would be logically flawed.

Third, despite insinuating that there is no consistency requirement, adjudicators considered (and endorsed) the EU’s claim that its regulatory approach passes the consistency test. Adjudicators found that a consistent application of the Suffering/Utility Standard can lead to a conclusion that seal slaughter is “inherently inhumane” and that a species-wide ban is justifiable. The reasoning and facts to support this conclusion are highly-questionable.

Lester argues that *Seals* confirms there is no consistency requirement under Article XX(a), but I believe the report is far more equivocal. The Appellate Body appears to reject Canada’s consistency arguments, but simultaneously allows the EU’s consistency arguments. Adjudicators rely on “effective stunning rates” as a benchmark in favour of the EU, but provide no data or sources to support this factual finding. While the outcome of this analysis might be intuitively-appealing to many people, it is not based on the type of coherent reasoning which is expected of adjudicators.

CHAPTER 13 - CONCLUSION AND FURTHER REFLECTIONS

13.1 Conclusion

This dissertation argued that there is an unresolved tension in WTO law about where to draw the line between retained and ceded regulatory autonomy. Since GATT drafters failed to clearly define which measures are necessary, adjudicators must fill this gap despite the fact it is considered taboo for them to make normative value judgments which second-guess the regulatory choices of Members. How can adjudicators draw the line between legitimate and WTO-inconsistent measures without making such value judgments?

In hard cases, adjudicators face a seemingly impossible task to determine whether moral, health and environmental objectives can justify trade restrictive measures. Since the Appellate Body has not been equipped with the tools of constitutional adjudication, it has turned to reasoning fallacies to hide the normative judgments which underpin its decisions. The real reasons go unstated.

In certain sensitive disputes (such as *Seals*), the real reasons might even be based on unconscious (intuitive) judgments rather than instrumental reasoning. Indeed, this phenomenon likely exists in judicial decision-making more broadly, not just WTO adjudication. On matters which inspire powerful intuitions, I make the socio-legal claim that judges (like other humans) can potentially have their capacity for System 2 normative reasoning overwhelmed by System 1.

Korea Beef

Early GATT panels handled Article XX disputes simply by testing if Pareto improvements could be found under the AM test, however *Korea Beef* revealed that

a Pareto approach would not be effective in all cases. Even though Korea's dual retail system was widely perceived as illegitimate under WTO law, a formalist application of the AM test likely would have found it WTO-consistent.

Adjudicators ultimately determined that the “desirable result” was more important than legal formalism which led to two highly-significant developments in WTO law. First, adjudicators introduced a stricter level of necessity testing by adopting the W&B test (at least in theory). This was highly-controversial as it represented a philosophical shift towards more intrusive review by WTO adjudicators (in the form of CBA).¹

Second, the *Korea Beef* adjudicators surreptitiously applied the AM test as a “crude CBA” by accepting alternative measures which lowered the level of protection and increased costs for the regulator. While, these methods are highly-intrusive, they spark far less controversy, possibly because the intrusiveness is less visible.

The Appellate Body seems to have learned an enduring lesson from *Korea Beef*. what you say matters more than what you do.² The seemingly-intrusive W&B test attracts strident criticism, while the highly-intrusive application of the AM test raises far less controversy.

Since *Korea Beef*, adjudicators have discovered other techniques for making disguised value judgments. When finding in favour of the regulator, a useful technique is to adopt an equivocal definition of the policy objective. My analysis considered this technique in both *Brazil Tyres* and *Seals*.

¹ In practice, the W&B test has not been operationalised in a meaningful way and has only been decisive under Article XX for “easy” cases, for example where the measure makes no contribution to the objective.

² This is one of the reasons I relied on deconstruction to make visible what the Appellate Body actually does, rather than what it claims to do.

Brazil Tyres

In *Brazil Tyres*, adjudicators used a spectrum of different policy objectives from extremely broad (public health) to extremely narrow (non-generation of waste tyres). This created a variable benchmark for different steps of the Article XX analysis. I compared this situation to behavioural economics where a large gap between “willingness to pay” and “willingness to accept” is inconsistent with the assumption of rationality. If adjudicators lack a fixed definition of the policy objective, does this suggest they are not behaving “rationally”?

Further drawing from behavioural economics, I suggested that this non-rational behaviour might also reveal “systematic bias” where a floating objective (for different steps of the analysis) consistently favours one party. I argued that my method can be used to test for systematic bias in any “necessity” dispute where adjudicators adopt a floating definition of the policy objective. Finally, as a general observation about WTO law, I suggested that providing some authority to adjudicators to make value judgments might be superior to a system of disguised value judgments through flawed reasoning.

Seals

Whereas the *Brazil Tyres* policy objective floated along a spectrum from broad to narrow, the *Seals* adjudicators adopted an equivocal objective based on two fundamentally-different moral approaches to animal protection. It contained a welfarist objective, based on the moral reasoning of Singer and Bentham, and an intuitionist objective.

I drew from moral psychology (and dual-process theory) to highlight the significant differences between these approaches to normative decision-making. Welfarism is a rationalist approach based on System 2 and philosophical reasoning, while intuitionism is a non-rationalist approach based on automatic and unconscious System 1 judgments.

Traditional models of judicial decision-making suggest that judges are able to see past their intuitions to ensure the “rationalist” integrity of their decisions. I argued that there are classes of social norm where the assumption of judicial rationality is questionable.³

In *Seals*, I sought to investigate whether powerful System 1 intuitions are capable of overpowering an adjudicator’s professional duty and training. In simple terms, are adjudicators humans or judges? Or, to ask a more nuanced version of this question, can psychology help us map those situations and issues where intuition is likely to trump reason, even for adjudicators?⁴

Framing Effects

A key reason why System 1 intuitions fail the rationality test is because the same fact pattern can lead to different conclusions depending on how it is framed. How is it possible, the pure rationalist might ask, that the same factual scenario can lead to two different intuitions (and therefore different moral conclusions) purely on the basis of how the information is framed? A judge who is influenced by framing effects would be no better than an economic agent who has a negative intuitive reaction to credit card surcharges, but not to equivalent discounts for cash purchases.

³ My challenge to the rationalism of judicial decision-making was inspired by the challenge of behavioural economics to the rationalism of standard economic theory.

⁴ Such scholarship would have a practical application in making judges aware of their “biases and prejudices”, as Frank sought to do in the 1930s (referred to in chapter 1).

Psychologists have identified this framing effect in moral reasoning. Haidt notes that it is possible to influence moral judgments by “reframing a problem to trigger new intuitions”.⁵ In a legal context, Sunstein notes that a “purely semantic shift in framing is sufficient to produce different outcomes” which should “raise obvious questions about the rationality of moral intuitions”.⁶

Under a rationalist model, judges are capable of ignoring the parties’ framing and ensuring that it does not influence outcomes. Judges use the parties’ substantive arguments as raw data to be reframed by the judge (if necessary) and assessed on its merits. However, the rationalist model does not seem to accurately describe the real world.

Indeed, the influence of framing effects, and the role of advocacy in developing them, is evident in the *Seals* dispute.⁷ When the EU (and *amici*) frame the debate in terms of tremendous seal suffering (under the “inherently inhumane” argument), this framing makes the seal products ban seem intuitively reasonable.

However, when Canada (and Perišin/Sellheim) frame the debate in terms of “animal suffering” (which would seem appropriate under a welfarist defence), the ban becomes harder to justify. Most animal exploitation leads to tremendous suffering: why does the EU seek to end the clubbing of Canadian seals but not the clubbing of European piglets and lambs? This may well be due to moral intuitions (about the cuteness of seals), but this explanation creates a dissonance between our intuitionist and rationalist reactions. Our moral reasoning is acutely aware that “cuteness” is not a valid reason for special protection.

⁵ Jonathan Haidt, ‘The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment’ (2001) 108(4) *Psychological Review* 814, 823.

⁶ Cass Sunstein, ‘Moral heuristics’ (2005) 28 *Behavioural and Brain Sciences* 531, 535.

⁷ The *Seals* dispute is an ideal site for this analysis as psychologists such as Herzog have shown that animal ethics is a topic where powerful intuitive reactions can overwhelm moral reasoning.

What is the real reason for the ban: suffering or cuteness? Even though it has significant legal implications, the legitimacy of the ban depends on how adjudicators choose to frame the matter and whether they allow seals to be compared with commodity animals for the purpose of consistency testing. If judges make decisions based on intuition and framing, what does this suggest about the certainty and predictability of judicial decisions? Indeed, what does it suggest more broadly about the determinacy of law?

It should be clear that intuitionist and rationalist approaches cannot unite coherently as part of a “hybrid claim”. Where decisions are made intuitively by System 1, the reasons produced by System 2 are mere post hoc rationalisations. Indeed, those reasons are likely to be riddled with logical fallacies and contradictions.

My analysis used deconstruction to reframe the reasoning provided by adjudicators (and the EU, *amici* and Posner) in the hope of producing different intuitions in the reader. This reframing revealed significant contradictions and inconsistencies which, I argue, provides compelling evidence that adjudicators rationalised an intuitive judgment rather than using coherent legal reasoning.

Despite its logical flaws, the Appellate Body’s reasoning may appear superficially convincing to those who share the underlying intuition that seals deserve special protection.⁸ My analysis sought to reframe the Appellate Body’s reasoning to make the logical fallacies and contradictions visible. This technique was intended to produce a cognitive dissonance in readers who would be torn between their intuition

⁸ Though the reasoning would seem intuitively suspect to observers who do not share this underlying intuition, including diverse groups from proponents of rationalist animal ethics (such as welfare and rights advocates) to proponents of seal hunting.

in favour of special protection for seals and their rationalist understanding of how WTO law and legal reasoning should operate.

If seals deserve special protection, I hope the reader will ask, why was the Appellate Body unable to reach this conclusion by applying coherent legal reasoning to current WTO jurisprudence? If my technique works, certain arguments in *Seals* which previously seemed (intuitively) reasonable may start to seem (logically) absurd. How do we resolve this tension between System 1 and System 2, especially when they are both equally adamant about their moral conclusions?

Main Findings

My analysis combined socio-legal and doctrinal techniques to bring a fresh perspective to WTO law and judicial decision-making more broadly. I developed a behavioural technique to determine whether the use of double standards under Article XX's necessity test is random or reflects the systematic bias of adjudicators.

I presented my findings in a way which can be validated or contested by other legal scholars. This behavioural technique, which draws from Thaler's use of "anomalies" in behavioural economics, can be applied to other "necessity" disputes at the WTO or could even be adapted to other bodies of law (by scholars who possess the relevant subject matter expertise).

My socio-legal analysis sought to expose an unresolved tension in WTO law which denies adjudicators the mandate they need to develop a coherent body of law. This was particularly evident in *Brazil Tyres* where adjudicators were denied the tools of constitutional adjudication (such as balancing) and therefore resorted to flawed reasoning as a way to disguise their underlying value judgment.

My analysis suggests that WTO law suffers from arrested development as adjudicators have consistently failed to make decisive statements to clarify the law over time. The web of vagueness in doctrinal law allows them to produce the ‘desirable result’ even in the absence of clear legal tests or explicit value judgments. While such approaches are antithetical to legal reasoning, this problem might well be inherent to systems based on negative integration.

There are some key examples where WTO law is stunted. The W&B test remains on the books, but is not used to undertake any meaningful (cost-benefit) analysis. The public morals jurisprudence has failed to clarify if evidentiary unilateralism can singlehandedly ground a claim, even in culturally-specific and non-instrumental disputes. We do not know if consistency testing exists under Article XX or any of its sub-paragraphs. WTO jurisprudence should have provided clear answers to these questions by now. My doctrinal analysis sought to highlight these gaps in WTO jurisprudence as a result of the Appellate Body’s “equivocal” approach to developing case law.

The *Seals* decision relied on the same technique as *Brazil Tyres* (non-rational double standards), but I believe there were different motivating forces at play. Since *Brazil Tyres* was an instrumental dispute, adjudicators were likely aware of the value judgments which motivated their decision (even if they were denied the mandate to explicitly state or rely on them).

In *Seals*, by contrast, it is unclear if adjudicators knew the real reasons for their decision. If the outcome of *Seals* was truly made by an instantaneous System 1 judgment, it is possible that the flawed reasoning by adjudicators was intended to fool not just an external audience, but also themselves.

In chapter 9, I asked what makes public morals different from the other GATT general exceptions. While it may be impolitic to say so, perhaps the fundamental difference is that there is a class of moral matters (including animal exploitation) where an adjudicator's System 2 reasoning is capable of being overwhelmed by powerful System 1 intuitions. On such matters, perhaps it can be concluded that even judges are capable of "moral dumbfounding".

If this is true, it has implications for judicial decision-making more broadly. Even where judges are provided with a mandate to make explicit value judgments, they might struggle to do so in cases where their decisions are driven by intuitions they are not consciously aware of.

I would like to finish this dissertation by focussing on the relationship between judicial decision-making and moral progress, especially the contrast between moral reasoning and rationalisation. I will consider this matter from the angle of moral philosophy, psychology, WTO law and constitutional adjudication.

13.2 Further Reflections on Law, Morality and Social Norms

In chapter 1, I introduced Haidt's three-part description of normative decision-making:

- Rationalism based on normative "reasoning" (System 2);
- Intuitionism based on automatic reactions (System 1); and
- Post hoc "rationalisation" (where System 2 reasoning is used to justify a System 1 decision).

The rest of this chapter will focus on the implications of these methods for judicial decision-making and moral progress. First, I will consider the "rationalist" and

“intuitionist” approaches under the banner of “coherent” reasoning.⁹ Second, I will consider “rationalisation” under the banner of “incoherent” reasoning.¹⁰ I will consider the appropriateness of these methods under both negative integration system (such as the WTO) and constitutional adjudication.

Haidt’s framework can be applied to normative decision-making at both an individual and a societal level. In the case of intuitionism, individual and collective decision-making operate in slightly different ways. When an individual makes an intuitionist decision, this means that they made an instantaneous judgment based on their System 1. For example, a citizen (or judge or parliamentarian) might have an intuitive reaction that seal exploitation is morally wrong.

However, it is also possible for public decisions to be made on the basis of “intuitionism”. For example, a regulator may choose to ban seal products (or a WTO adjudicator may accept the legitimacy of that ban) on the basis that it reflects a widespread social intuition. In this scenario, the regulator/adjudicator is not making a decision based on personal intuition. Rather, they are making a public decision based on an empirical investigation of whether a certain intuition is widespread within a given society.¹¹

13.2.1 “Coherent” Approaches: Reasoning and Intuition

Philosophers and moral psychologists would endorse rationalist approaches. On animal exploitation, the leading schools of moral philosophy (welfarist and rights

⁹ I describe these approaches as “coherent” in the sense that the real reasons and the stated reasons are the same. My use of the term “coherent” does not represent a normative claim about the desirability of each approach.

¹⁰ I describe rationalisation as “incoherent” because the real reasons for the decision are different from the stated reasons.

¹¹ For example, an EU regulator or WTO adjudicator might not share the intuition that seals deserve special protection (in their personal capacity), but they might declare the legitimacy of the seal products ban (in their official capacity) as they recognise that it truly does reflect a widespread social intuition within the EU.

approaches) are based on rationalism. Haidt endorses rationalist decision-making when he highlights that intuitionism in public decisions can lead to “non-optimal or even disastrous consequences”.¹² Shermer makes the broader claim that the “moral progress” of humanity depends on our use of “reason”. While these social claims may be compelling, what is their significance for judges?

Judges are not free to make their decisions based on first principles. Unlike moral philosophers, judges are constrained by external factors such as their mandate (under a treaty or constitution) or the substantive legal rules they are bound to apply. The scope of a judge to engage in “moral rationalism” will vary based on the nature of the system they operate in and the legal document they are interpreting. The judge’s first duty is to the law. If the law tolerates injustices (even those which cannot be justified by moral reasoning), the fault ultimately lies with the legislator and broader society. It is not the judge’s role to impose moral progress on society.

Constitutional Adjudication

However, there may be certain contexts, such as the interpretation of a Bill of Rights, where it would be appropriate for judges to use moral reasoning. When constitutional courts rely on an overarching “right” or “principle” to recognise gender equality, civil rights or native title for indigenous communities, they take a “first principle” derived from law (such as the constitution) but use moral reasoning to expand its interpretation to protect oppressed communities.

Such reasoning can be used to end oppressive practices which are morally-unjustifiable, even if those practices remain widely perceived as legitimate. In such

¹² Haidt (n 5) 815. This echoes Frank’s call to overcome “biases and prejudices”. While Haidt argues that there are serious impediments to the use of moral rationalism by individuals, he believes that public decisions (including law) can be based on systems and structures which promote rationalism.

cases, judges use moral reasoning to poke their head out in front of society, rather than waiting for social or political dynamics to change. While remaining within their mandate, judges can perform a valuable social role as they are among the few stakeholders capable of using rationalism to see past social intuitions. These types of ground-breaking judicial interventions can promote moral progress and change societal perceptions of what is appropriate and just. However, such interventions are used sparingly and only by the highest national courts.

WTO Law (and Negative Integration)

Unlike a constitutional court, WTO adjudicators do not have a mandate to make “rationalist” decisions contrary to the social intuitions of Member societies. Under the WTO’s negative integration system, the key question in *Seals* was whether the measure reflected a widespread social intuition, rather than whether it could be justified by rationalist moral philosophy. It would have been highly-intrusive and controversial for WTO adjudicators to reject the seal products ban on the basis that Peter Singer rejects it.¹³

If WTO adjudicators ruled that seal (or dog) product bans were immoral and therefore WTO-inconsistent, they might be applauded at moral philosophy conferences, but they would be shunned by their own Members. For the WTO and Appellate Body, the constant questioning of their institutional legitimacy means the scope for rationalist reasoning is constrained. Their constituency is the membership, not moral philosophers.

¹³ While adjudicators cannot impose Singer’s philosophy against a regulator’s will, the issue in *Seals* was more complex as the EU claimed to base its measure on Singer’s philosophy. In this sense, the EU created a welfarist benchmark for itself (rather than having the benchmark imposed by adjudicators).

13.2.2 The “Incoherent” Approach: Rationalisation

The third approach in Haidt’s framework, rationalisation, is based on a dissonance between the real (intuitive) reasons for a decision and the (seemingly-rationalist) public justification. This problem was foreshadowed by the Realists, a century ago, when they argued that law and adjudication should seek to reconcile the real reasons with the stated ones. Weiler observed this problem in a WTO context when he observed that “in law, a right result might be a necessary but not a sufficient condition”.¹⁴ Results matter, but so do reasons.

Rationalism (rightly) has its supporters as the gold standard for reasoning. At the other end of the spectrum rationalisation has its opponents as the antithesis of reasoning. In an animal ethics context, Joy argues that people often bypass coherent moral reasoning when they rely on the “Three Ns” of rationalisation – that animal exploitation is necessary, natural and normal. Joy notes that the “Three Ns” are a common theme in many systems of rationalised oppression and exploitation, including African slavery and denying women the right to vote.¹⁵

One of the most troubling aspects of rationalisation is that it discourages evidence and fact-based debate. As Joy points out, “violent ideologies rely on promoting fiction as fact and discouraging any critical thinking that threatens to expose this truth”.¹⁶ She further notes:

¹⁴ JHH Weiler, ‘Law, Culture, and Values in the WTO – Gazing into the Crystal Ball’ in Daniel Bethlehem, Isabelle Van Damme, Donald McRae, and Rodney Neufeld (eds), *The Oxford Handbook of International Trade Law* (OUP 2009) 139.

¹⁵ Melanie Joy, *Why We Love Dogs, Eat Pigs, and Wear Cows: An Introduction to Carnism* (Conari Press 2009) 97. In this respect, she cites Voltaire: “if we believe absurdities, we shall commit atrocities”.

¹⁶ ibid 96. Piazza et al endorse Joy’s view: “when an ideology is widely endorsed, as meat-eating is in most parts of the world today, the justifications supporting the ideology generally go unchallenged”; Jared Piazza, Matthew Ruby, Steve Loughnan, Mischel Luong, Juliana Kulik, Hanne Watkins and Mirra Seigerman, ‘Rationalizing Meat Consumption: the 4Ns’ (2015) 91 Appetite 114, 115. Piazza et al add “nice” as a fourth “N” of rationalisation.

“When an ideology is in its prime, these myths rarely come under scrutiny. However, when the system finally collapses, the Three Ns are recognised as ludicrous”.¹⁷

How can we, as a society, get better at identifying these “myths” and realising that they are based on “ludicrous” rationalisations rather than moral reasoning? If questions about the morality of slavery, gender inequality, civil rights or animal exploitation arise in legal disputes, which moral approach should judges adopt?

If the use of rationalisation is inherently incoherent and has no place in judicial decision-making, this leaves judges with two “coherent” options. In certain rare cases of constitutional adjudication, judges can rely on rationalism to expose unjustifiable myths. They can drive moral progress by ruling in favour of gender equality or civil rights.

In negative integration systems, this approach is not available. Since WTO adjudicators are precluded from using moral rationalism, they might have no option but to accept the legal validity of social intuitions (such as seal product bans) even where such norms cannot be justified by moral reasoning.

However, these adjudicators still face an important moral choice in terms of how they justify the legal validity of such norms. Should they (falsely) declare that such norms are morally-justifiable or should they (honestly) declare that they merely reflect a widespread social intuition? This choice – between rationalisation and intuitionism – is a critical one.

13.2.3 Rationalisation, Public Morals and the *Seals* Dispute

While moral rationalism was not available to the *Seals* adjudicators, they had to make an important choice between rationalisation and intuitionism in two places.

¹⁷ Joy (n 15) 97.

First, they had to consider it as a systemic matter under the public morals legal test. Second, they had to consider it with respect to the seal products ban. I will address each of these issues.

Public Morals Jurisprudence

To the extent that the WTO is a negative integration system, it should permit regulators to impose measures on the basis of widespread social intuitions (such as seal product bans).¹⁸ Indeed, many scholars have promoted “unilateralism” as a way to achieve this outcome¹⁹ and WTO adjudicators have made statements endorsing such “unilateral” approaches.

In *Gambling*, the panel stated that public morals “can vary in time and space” depending on the “prevailing social, cultural, ethical and religious values” of different societies.²⁰ It further noted that “Members should be given some scope to define and apply for themselves the concepts of ‘public morals’… according to their own systems and scales of values”.²¹

However, adjudicators have also stated that moral measures must be based on “standards of right and wrong”. This “moral standards” requirement suggests that regulators must justify their measures as more than mere social intuitions. In early disputes, adjudicators relied on transnationalist and internationalist evidence as proof of moral standards.

¹⁸ In my view, the description of the WTO as a negative integration system oversimplifies the reality. There are situations, like *Korea Beef*, where outcomes which rely on “balancing” have been widely-accepted as legitimate (provided there is no explicit acknowledgment that balancing actually occurred).

¹⁹ I referred to Charnovitz and Marwell as leading proponents of this view.

²⁰ Panel Report, *United States – Measures Affecting the Cross-Border Supply of Gambling and Betting Services*, WT/DS285/R, adopted 20 April 2005 (*US – Gambling*) para 6.461.

²¹ *ibid*. The GATS public morals exception also includes a reference to “public order” which does not appear in Article XX(a).

In *Seals*, for the first time, adjudicators turned to rationalist moral philosophy (welfarism) for the purpose of demonstrating “moral standards”. If the public morals test contains a moral rationalism requirement, this implies that adjudicators do not merely declare a contested measure WTO-consistent (a legal finding), but that they also declare it morally-justifiable. Are they mandated to make such moral findings? What purpose does it serve?

In reality, the jurisprudence remains equivocal. While adjudicators relied on moral rationalism in *Seals*, the overarching legal test remains ill-defined, especially the relationship between “moral standards” and “unilateralism”.

The Seal Products Ban

The *Seals* adjudicators bypassed coherent approaches (rationalism and intuitionism) and ultimately relied on rationalisation to justify the seal products ban.²²

Rationalism plays virtually no role in the *Seals* dispute. While WTO adjudicators, the EU and *amici* invoke Singer’s welfarism, they do not seek to apply it in any meaningful way. In chapter 9, I explained why moral rationalism cannot justify the seal products ban, as both the welfarist and rights schools of animal ethics would reject its credentials. In a WTO law context, both Perišin and Sellheim explained clearly why the ban could not be justified as a welfarist measure, primarily because of the failure to consistently apply a welfarist standard to all animals.²³

²² The same claim can also be made about the EU and *amici* whose reasoning follows the same approach as the Appellate Body.

²³ Based on the fact that the seal products ban cannot be justified by welfarism, both Perišin and Sellheim conclude that it therefore cannot be justified by WTO law. In this sense, they arguably suggest that Article XX(a) requires rationalism and that intuitionism cannot serve as the basis for a moral measure. Perišin and Sellheim make a very important contribution to the literature, but their failure to acknowledge intuitionism as a potential defence represents a gap in their analysis. In their defence, the EU’s claim is explicitly welfarist in nature. Perišin and Sellheim were entitled to assess the EU’s explicit claim rather than deconstructing it for the purpose of identifying (and assessing) any hidden claims.

Intuitionism also fails to play a direct role in *Seals* as the EU refrained from pursuing an explicitly intuitionist defence. It made some indirect arguments which were intuitionist in nature, but these were merely alluded to in certain contexts where the EU's explicitly welfarist defence was weak (such as the matter of regulatory consistency).²⁴

From a legal perspective, the intuitionist defence seems highly compelling and it is hard to understand why adjudicators ignored it. The EU produced compelling evidence that Europeans really are appalled by the exploitation of seals for commercial purposes.²⁵ Adjudicators could have determined that, under the WTO's negative integration system, all that matters is whether there is an EU norm demanding higher protection for seals, not whether that norm is morally-justifiable.

However, rather than following the unilateralist approach, adjudicators made the dubious finding that the EU's measure was based on moral standards and that even Peter Singer would endorse it. They were not satisfied with declaring the legal validity of the ban, but also sought to declare its moral validity. Why did they seek to rationalise the seal products ban in this way?

The *Seals* dispute and surrounding literature primarily take place under the third limb of Haidt's framework: post hoc rationalisation. This is where the real reasons for a social norm are intuitionist, but judges and scholars hide these real reasons behind seemingly-rationalist moral reasoning. In *Seals*, adjudicators, the EU and *amici* falsely claimed the seal products ban was based on "moral standards" and could be justified by Singer's rationalist moral approach.

²⁴ WTO adjudicators and the *amici* adopt the same strategy of indirectly inserting intuitionist arguments into a defence which is explicitly welfarist.

²⁵ Therefore, as an evidentiary matter, the EU was well-placed to prove that its social intuition did exist. Indeed, the EU managed to successfully demonstrate this fact under the "public concerns" element.

Why did adjudicators rely on rationalisation when they had a legitimate pathway to the “desirable result” simply by accepting “intuitionism” and “evidentiary unilateralism”? The sole benefit of rationalisation was that it allowed adjudicators and other stakeholders to (falsely) believe that the seal products ban was morally-justifiable (in addition to being legally-justifiable). It provided a psychological benefit, not a legal one.

13.2.4 Law, Moral Progress and Behavioural Adjudication

The use of rationalisation is deeply pernicious and has no place in judicial or moral decision-making. From a legal perspective, it detracts from the judicial project of establishing coherent reasoning and doctrine.

In *Seals*, the use of rationalisation by adjudicators undermined WTO law by using flawed reasoning (such as equivocal language and double standards) to produce flawed doctrine. Due to this method, WTO jurisprudence continues to lack clarity on the status of consistency testing, evidentiary unilateralism or what the “moral standards” requirement means in practice.

From a social perspective, rationalisation offers a moral seal of approval to widespread social practices (such as the seal products ban) which lack moral credentials. The *Seals* adjudicators (and scholars) should have easily uncovered the fact that Singer denounces special protection for seals.

The *Seals* adjudicators lacked a mandate to find the seal products ban WTO-inconsistent, but they also lacked a mandate to declare it morally-justifiable. They may have produced the correct result, but their choice of reasoning still matters. When judges engage in rationalisation, they stifle moral progress by offering a stamp of both legal and moral validity to social practices which cannot be morally justified.

Philosophers and moral psychologists focus on reason as the key to achieving moral progress. While moral rationalism is a noble ideal, it is nonetheless limited in real-world contexts, including judicial decision-making. Haidt notes that individuals rarely make decisions according to the rationalist model. Most judges, including WTO adjudicators, are precluded from engaging in moral rationalism. Even those judges who have a mandate to use moral rationalism may struggle to do so for sensitive social issues where their System 1 intuitions overwhelm their System 2 reasoning.

In those contexts where judges are not mandated to use moral rationalism, I argue that moral progress can be advanced by eliminating rationalisation and requiring intuitive decisions to be justified according to the real reasons. There is no legal barrier to WTO adjudicators relying on the real (intuitionist) reasons to justify the legitimacy of a seal or dog product ban. The only barrier is psychological. However, this psychological barrier is significant because it hampers critical scrutiny of social norms which may well be exposed as “ludicrous”. The elimination of rationalisation is a major step on the path to moral progress.

The tools of behavioural adjudication have a key role to play in exposing (and preventing) rationalisation by judges. Whereas Haidt uses the tools of psychology to study how people make decisions, my method is designed to reverse engineer the decisions of judges to see what the real reasons were. Using a judicial decision as the starting point of analysis, I have developed tools to test whether the “reasons” given by judges actually motivated a rationalist decision or whether they merely sought to rationalise an intuitionist one.

I argue that the egregious use of logical fallacies and contradictions in judicial decisions may be a strong indicator of rationalisation. According to my method, the

first step is to identify “anomalies” and other types of logical fallacies. Behavioural lawyers should become increasingly systematic at documenting these “anomalies” in ways which can be contested or validated.

Rather than treating “anomalies” as an end point (“the judge got it wrong” or “the judge’s reasoning was poor”), they should serve as the raw data for deeper analysis. In addition to showing that a judicial decision was not rational, can legal scholars go one step further to determine if it was based on systematic bias?²⁶

According to Haidt, judges represent one of the few safeguards in our society who can use moral reasoning to rationally discuss social norms rather than being blinded by current social prejudices. My analysis seeks to demonstrate the circumstances under which this safeguard systematically fails.

Behavioural economists have successfully demonstrated how human judgments can systematically misfire through the intervention of System 1. In the realm of normative reasoning, I believe that behavioural lawyers can start to map those situations where human moral reasoning systematically fails.

The fact that WTO adjudicators have no mandate for moral rationalism does not excuse their decision in *Seals*. Even under such constraints, their mandate permitted them to articulate the real reasons for their decision, rather than engaging in rationalisation. Judges (including WTO adjudicators) do much damage to moral progress when they give a “moral” stamp of approval to (and stifle critical scrutiny of) social practices which have no rationalist justification.

²⁶ This has been a major accomplishment in behavioural economics which behavioural lawyers could seek to emulate.

In behavioural law, we are starting to develop the tools to call judges out when they rely on rationalisation rather than coherent and honest reasoning.²⁷ Even where moral rationalism remains out of reach, much can be accomplished simply by stamping out rationalisation.

The EU's seal products ban cannot be justified by Singer's welfarism or by "moral standards". It may not seem like much, but it would have been a major step towards moral progress if WTO adjudicators (and the EU and *amicus*) had been able to admit that their real reason for opposing seal exploitation was: "I don't know, I can't explain it, I just know it's wrong".

²⁷ Such coherent reasoning can be rationalist (under constitutional adjudication) or intuitionist (under WTO adjudication).

TABLE OF GATT PANELS AND WTO DISPUTES

GATT Panel Report, *Japan – Customs Duties, Taxes and Labelling Practices on Imported Wines and Alcoholic Beverages*, L/6216, adopted 10 November 1987 (*Japan Alcohol I*).

GATT Panel Report, *United States Section 337 of the Tariff Act of 1930*, L/6439, adopted 7 November 1989 (*Section 337*).

GATT Panel Report, *Thailand – Restrictions on Importation of and Internal Taxes on Cigarettes*, DS10/R, adopted 7 November 1990 (*Thai Cigarettes*).

Appellate Body Report, *United States – Standards for Reformulated and Conventional Gasoline*, WT/DS2/AB/R, adopted 20 May 1996 (*Gasoline*).

Panel Report, *United States – Standards for Reformulated and Conventional Gasoline*, WT/DS2/R, adopted 20 May 1996, as modified by Appellate Body Report WT/DS2/AB/R (*Gasoline*).

Appellate Body Report, *Japan – Taxes on Alcoholic Beverages*, WT/DS8/AB/R, WT/DS10/AB/R, WT/DS11/AB/R, adopted 1 November 1996 (*Japan Alcohol II*).

Appellate Body Report, *European Communities – Regime for the Importation, Sale and Distribution of Bananas*, WT/DS27/AB/R, adopted 25 September 1997 (*Bananas III*).

Appellate Body Report, *Canada – Certain Measures Concerning Periodicals*, WT/DS31/AB/R, adopted 30 July 1997 (*Canada Periodicals*).

Panel Report, *Japan – Measures Affecting Consumer Photographic Film and Paper*, WT/DS44/R, adopted 22 April 1998 (*Japan Film*).

Appellate Body Report, *United States – Import Prohibition of Certain Shrimp and Shrimp Products*, WT/DS58/AB/R, adopted 6 November 1998 (*Shrimp*).

Panel Report, *United States – Import Prohibition of Certain Shrimp and Shrimp Products*, WT/DS58/R and Corr.1, adopted 6 November 1998, as modified by Appellate Body (*Shrimp*).

Appellate Body Report, *Chile – Taxes on Alcoholic Beverages*, WT/DS87/AB/R, WT/DS110/AB/R, adopted 12 January 2000 (*Chile – Alcohol*).

Appellate Body Report, *European Communities – Measures Affecting Asbestos and Asbestos-Containing Products*, WT/DS135/AB/R, adopted 5 April 2001 (*Asbestos*).

Appellate Body Report, *Korea – Measures Affecting Imports of Fresh, Chilled and Frozen Beef*, WT/DS161/AB/R, WT/DS169/AB/R, adopted 10 January 2001 (*Korea Beef*).

Panel Report, *Korea – Measures Affecting Imports of Fresh, Chilled and Frozen Beef*, WT/DS161/R, WT/DS169/R, adopted 10 January 2001, as modified by Appellate Body Report WT/DS161/AB/R, WT/DS169/AB/R (*Korea Beef*).

Panel Report, *European Communities – Protection of Trademarks and Geographical Indications for Agricultural Products and Foodstuffs, Complaint by the United States*, WT/DS174/R, adopted 20 April 2005 (EC – *Trademarks and Geographical Indications (US)*).

Panel Report, *European Communities – Trade Description of Sardines*, WT/DS231/R and Corr.1, adopted 23 October 2002, as modified by Appellate Body Report WT/DS231/AB/R (*Sardines*).

Appellate Body Report, *United States – Continued Dumping and Subsidy Offset Act of 2000*, WT/DS217/AB/R, WT/DS234/AB/R, adopted 27 January 2003 (*Byrd Amendment*).

Panel Report, *European Communities – Conditions for the Granting of Tariff Preferences to Developing Countries*, WT/DS246/R, adopted 20 April 2004, as modified by Appellate Body Report WT/DS246/AB/R (EC – *Tariff Preferences*).

Panel Report, *Canada – Measures Relating to Exports of Wheat and Treatment of Imported Grain*, WT/DS276/R, adopted 27 September 2004, upheld by Appellate Body Report WT/DS276/AB/R (*Canada Wheat*).

Appellate Body Report, *United States – Measures Affecting the Cross-Border Supply of Gambling and Betting Services*, WT/DS285/AB/R, adopted 20 April 2005 (US – *Gambling*).

Panel Report, *United States – Measures Affecting the Cross-Border Supply of Gambling and Betting Services*, WT/DS285/R, adopted 20 April 2005 (US – *Gambling*).

Appellate Body Report, *Dominican Republic – Measures Affecting the Importation and Internal Sale of Cigarettes*, WT/DS302/AB/R, adopted 19 May 2005 (DR – *Cigarettes*).

Panel Report, *Dominican Republic – Measures Affecting the Importation and Internal Sale of Cigarettes*, WT/DS302/R, adopted 19 May 2005, as modified by Appellate Body Report WT/DS302/AB/R (DR – *Cigarettes*).

Panel Report, *Mexico – Tax Measures on Soft Drinks and Other Beverages*, WT/DS308/R, adopted 24 March 2006, as modified by Appellate Body Report WT/DS308/AB/R (*Mexico – Soft Drinks*).

Appellate Body Report, *Brazil – Measures Affecting Imports of Retreaded Tyres*, WT/DS332/AB/R, adopted 17 December 2007 (*Brazil Tyres*).

Panel Report, *Brazil – Measures Affecting Imports of Retreaded Tyres*, WT/DS332/R, adopted 17 December 2007 (*Brazil Tyres*).

Appellate Body Report, *United States – Final Anti-Dumping Measures on Stainless Steel from Mexico*, WT/DS344/AB/R, adopted 20 May 2008 (US – *Stainless Steel (Mexico)*).

Appellate Body Report, *China – Measures Affecting Trading Rights and Distribution Services for Certain Publications and Audiovisual Entertainment Products*, WT/DS363/AB/R, adopted 19 January 2010 (*China Audiovisuals*).

Panel Report, *China – Measures Affecting Trading Rights and Distribution Services for Certain Publications and Audiovisual Entertainment Products*, WT/DS363/R and Corr.1, adopted 19 January 2010, as modified by Appellate Body Report WT/DS363/AB/R (*China Audiovisuals*).

Panel Report, *Colombia – Indicative Prices and Restrictions on Ports of Entry*, WT/DS366/R and Corr.1, adopted 20 May 2009 (*Colombia – Ports of Entry*).

Appellate Body Report, Thailand – Customs and Fiscal Measures on Cigarettes from the Philippines, WT/DS371/AB/R, adopted 15 July 2011 (*Thailand – Cigarettes (Philippines)*).

Appellate Body Report, *United States – Measures Affecting the Production and Sale of Clove Cigarettes*, WT/DS406/AB/R, adopted 24 April 2012 (*Clove Cigarettes*).

Panel Report, *United States – Measures Affecting the Production and Sale of Clove Cigarettes*, WT/DS406/R, adopted 24 April 2012 (*Clove Cigarettes*).

Appellate Body Reports, *European Communities – Measures Prohibiting the Importation and Marketing of Seal Products*, WT/DS400/AB/R / WT/DS401/AB/R, adopted 18 June 2014 (*Seals*).

Panel Reports, *European Communities – Measures Prohibiting the Importation and Marketing of Seal Products*, WT/DS400/R and Add.1 / WT/DS401/R and Add.1, adopted 18 June 2014 (*Seals*).

Appellate Body Reports, *China – Measures Related to the Exportation of Rare Earths, Tungsten, and Molybdenum*, WT/DS431/AB/R / WT/DS432/AB/R / WT/DS433/AB/R, adopted 29 August 2014 (*Rare Earths*).

Appellate Body Report, *India – Certain Measures Relating to Solar Cells and Solar Modules*, WT/DS456/AB/R and Add.1, adopted 14 October 2016 (*Solar Cells*).

Appellate Body Report, *Colombia – Measures Relating to the Importation of Textiles, Apparel and Footwear*, WT/DS461/AB/R and Add.1, adopted 22 June 2016 (*Colombia Textiles*).

Panel Report, *Indonesia – Measures Concerning the Importation of Chicken Meat and Chicken Products*, WT/DS484/R and Add.1, adopted 22 November 2017 (*Indonesia – Chicken*).

Panel Reports, *Brazil – Certain Measures Concerning Taxation and Charges*, WT/DS472/R, Add.1 and Corr.1, WT/DS497/R, Add.1 and Corr.1, adopted 11 January 2019, as modified by Appellate Body Reports WT/DS472/AB/R, WT/DS497/AB/R (*Brazil Taxation*).

Panel Report, European Union and its member States – Certain Measures Relating to the Energy Sector, WT/DS476/R and Add.1, circulated to WTO Members 10 August 2018 [appealed by the European Union 21 September 2018 – the Division suspended its work on 10 December 2019] (*EU – Energy Supply*).

TABLE OF TREATIES AND INTERNATIONAL INSTRUMENTS

Final Act Embodying the Results of the Uruguay Round of Multilateral Trade Negotiations (adopted 15 April 1994, entered into force 1 January 1995) 1867 UNTS 14.

Marrakesh Agreement Establishing the World Trade Organization (adopted 15 April 1994, entered into force 1 January 1995) 1867 UNTS 154 (“Marrakesh Agreement” or “WTO Agreement”).

General Agreement on Tariffs and Trade 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1A (adopted 15 April 1994, entered into force 1 January 1995) 1867 UNTS 187 (“GATT”).

Agreement on Subsidies and Countervailing Measures, Marrakesh Agreement Establishing the World Trade Organization, Annex 1A (adopted 15 April 1994, entered into force 1 January 1995) 1869 UNTS 14 (“Subsidies Agreement”).

Agreement on the Implementation of Article VI of GATT 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1A (adopted 15 April 1994, entered into force 1 January 1995) 1868 UNTS 201 (“Anti-Dumping Agreement”).

Agreement on Technical Barriers to Trade, Marrakesh Agreement Establishing the World Trade Organization, Annex 1A (adopted 15 April 1994, entered into force 1 January 1995) 1868 UNTS 20 (“TBT Agreement”).

Agreement on the Application of Sanitary and Phytosanitary Measures, Marrakesh Agreement Establishing the World Trade Organization, Annex 1A (adopted 15 April 1994, entered into force 1 January 1995) 1867 UNTS 493 (“SPS Agreement”).

General Agreement on Trade in Services, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1B (adopted 15 April 1994, entered into force 1 January 1995) 1869 UNTS 183 (“GATS”).

Dispute Settlement Rules: Understanding on Rules and Procedures Governing the Settlement of Disputes, Marrakesh Agreement Establishing the World Trade Organization, Annex 2 (adopted 15 April 1994, entered into force 1 January 1995) 1869 UNTS 401 (“DSU”).

Protocol amending the Marrakesh Agreement establishing the World Trade Organization (adopted 27 November 2014, entered into force 22 February 2017) WTO Treaty Series No. 47, WT/Let/1030, WT/L/940 (Trade Facilitation Agreement).

Vienna Convention on the Law of Treaties (entered into force 27 January 1980) 1155 UNTS 331 (“VCLT”).

WHO Framework Convention on Tobacco Control (opened for signature 16 June 2003, entered into force 27 February 2005) 2302 UNTS 166 (“WHO FCTC”).

BIBLIOGRAPHY

Articles and Books

Atkinson S, Velarde A and Algiers B, 'Assessment of Stun Quality at Commercial Slaughter in Cattle Shot with Captive Bolt' (2013) 22 *Animal Welfare* 473.

Bacchus J and Lester S, 'Of Precedent and Persuasion: The Crucial Role of an Appeals Court in WTO Disputes' (12 September 2019) 74 *Free Trade Bulletin* 1.

Balkin J, 'Deconstructive Practice and Legal Theory' (1987) 96 *Yale Law Journal* 743.

Bartels L, 'The Chapeau of the General Exceptions in the WTO GATT and GATS Agreements: A Reconstruction' (2015) 109(1) *American Journal of International Law* 95.

Bentham J, *An Introduction to the Principles of Morals and Legislation* (first published 1789, 1823 edition).

Bhagwati J, 'Free Trade: What Now?' (Keynote Address on the occasion of the International Management Symposium at which the 1998 Freedom Prize of the Max Schmidheiny Foundation was awarded, St Gallen, 25 May 1998).

Bix B, *Law, Language, and Legal Determinacy* (OUP 1995).

Bown C and Trachtman J, 'Brazil – Measures Affecting the Import of Retreaded Tyres: A Balancing Act' (2009) 8 *World Trade Review* 85.

Bratanova B, Loughnan S and Bastian B, 'The Effect of Categorization as Food on the Perceived Moral Standing of Animals' (2011) 57(1) *Appetite* 193.

Charnovitz S, 'The Moral Exception in Trade Policy' (1998) 38(4) *Virginia Journal of International Law* 689.

Charnovitz S, 'The Law of Environmental "PPMs" in the WTO: Debunking the Myth of Illegality' (2002) 27 *Yale Journal of International Law* 59.

Conconi P and Voon T, 'EC Seal Products: The Tension between Public Morals and International Trade Agreements' (2016) 15(2) *World Trade Review* 211.

De Ville F, 'Explaining the Genesis of a Trade Dispute the European Union's Seal Trade Ban' (2012) 34 *Journal of European integration* 37.

Diebold N, 'The Morals and Order Exceptions in WTO Law: Balancing the Toothless Tiger and the Undermining Mole' (2007) 11(1) *Journal of International Economic Law* 43.

Doyle C, 'Gimme Shelter: The Necessary Element of GATT Article XX in the Context of the Audiovisuals Products Case (2011) 29 *Boston University International Law Journal* 143.

Dworkin R, *A Matter of Principle* (Harvard University Press 1985).

Ehlermann, C-D, 'Tensions between the dispute settlement process and the diplomatic and treaty-making activities of the WTO' (2002) 1(3) World Trade Review 301.

Elsig M, Hoekman B and Pauwelyn J, 'Thinking about the performance of the World Trade Organization: A discussion across disciplines' (Robert Schuman Centre for Advanced Studies 2016, EUI Working Paper RSCAS 2016/13).

Fairclough N, *Critical Discourse Analysis: the Critical Study of Language* (London Longman 1995).

Feddersen C, 'Focusing on Substantive Law in International Economic Relations: The Public Morals of GATT's Article XX(a) and "Conventional" Rules of Interpretation', (1998) 7 Minnesota Journal of Global Trade 75.

Fontanelli F, 'Necessity Killed the GATT - Art XX GATT and the Misleading Rhetoric about Weighing and Balancing' (2012) 5 European Journal of Legal Studies 39.

Francione G, 'Animal Welfare and the Moral Value of Nonhuman Animals (2010) 6(1) Law, Culture and the Humanities 24.

Frank J, 'Are Judges Human? Part One: The Effect on Legal Thinking of the Assumption That Judges Behave like Human Beings' (1931) 80 University of Pennsylvania Law Review and American Law Register 17.

Frank J, 'Are Judges Human? Part Two: As Through a Glass Darkly' (1931) 80 University of Pennsylvania Law Review and American Law Register 233.

Gonzalez M, 'Trade and Morality: Preserving "Public Morals" Without Sacrificing the Global Economy' (2006) 39 Vanderbilt Journal of International Law 939.

Grossman G and Horn H, 'Why the WTO? An Introduction to the Economics of Trade Agreements' (Research Institute of Industrial Economics 2012, Stockholm, IFN Working Paper No 916).

Grossman G, Horn H and Mavroidis P, 'Legal and Economic Principles of World Trade Law: National Treatment' (Research Institute of Industrial Economics 2012, Stockholm, IFN Working Paper No 917).

Haidt J, 'The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment' (2001) 108(4) Psychological Review 814.

Halliday S and Schmidt P, *Conducting Law and Society Research: Reflections on Methods and Practices* (CUP 2009)

Hart HLA, *The Concept of Law* (2nd edition, OUP 1994).

Herwig A, 'Regulation of Seal Animal Welfare Risk, Public Morals and Inuit Culture under WTO Law: Between Techne, Oikos and Praxis' (2015) 6 European Journal of Risk Regulation 382.

Herwig A, 'Too much Zeal on Seals? Animal Welfare, Public Morals, and Consumer Ethics at the Bar of the WTO' (2016) 15(1) World Trade Review 109.

Herzog H and Burghardt G, 'The Next Frontier: Moral Heuristics and the Treatment of Animals' (2005) 28 Behavioural and Brain Sciences 554.

Howse R and Langille J, 'Permitting Pluralism: The Seal Products Dispute and Why the WTO Should Accept Trade Restrictions Justified by Noninstrumental Moral Values' (2012) 37(2) Yale Journal of International Law 367.

Howse R, Langille J and Sykes K, 'Pluralism in Practice: Moral Legislation and the Law of the WTO after Seal Products' (2015) 48 George Washington International Law Review 81.

Hudec R, 'Circumventing Democracy: The Political Morality of Trade Negotiations' (1993) 25 New York University Journal of International Law and Politics 311.

Hudec R, 'GATT/WTO Constraints on National Regulation – "Aims and Effects" Test' (1998) 32(3) The International Lawyer 619.

Hutcheson J, 'The Judgment Intuitive: The Function of the Hunch in Judicial Decision' (1929) 14(3) Cornell Law Review 274.

Irwin D, Mavroidis P and Sykes A, *The Genesis of the GATT* (CUP 2008).

Jackson J, 'Power and Rules in the Changing Economic Order: The Case of the World Trade Organization' (2008) 84(3) International Affairs 437.

Jackson J, 'The WTO 'constitution' and proposed reforms: Seven 'Mantras' revisited' (2001) 4(1) Journal of International Economic Law 67.

Joy M, *Why We Love Dogs, Eat Pigs, and Wear Cows: An Introduction to Carnism* (Conari Press 2009).

Kahneman D, Knetsch J and Thaler R, 'Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias' (1991) 5(1) The Journal of Economic Perspectives 193.

Kahneman D, *Thinking Fast and Slow* (Farrar, Straus and Giroux, 2011).

Kaplow L, 'Rules versus Standards: An Economic Analysis' (1992-1993) 42 Duke Law Journal 557.

Kaplow L, 'General Characteristics of Rules' in Bouckaert B and De Geest G (eds), *Encyclopedia of Law and Economics* (Edward Elgar, 2000) 502.

Kapterian G, 'A Critique of the WTO Jurisprudence on 'Necessity' (2010) 59 International & Comparative Law Quarterly 89.

Krugman P, 'What Should Trade Negotiators Negotiate About?' (1997) 35(1) Journal of Economic Literature 113.

Leiter B, 'Legal Realism and Legal Doctrine' (2015) 163 University of Pennsylvania Law Review 1975.

Lester S, 'The WTO Seal Products Dispute: A Preview of the Key Legal Issues' *ASIL Insights* 14(2) (13 January 2010).

Lester S, 'There's No "Consistency" Requirement for Animal Welfare and Public Morals' (International Economic Law and Policy Blog, 22 May 2014)

<<https://ielp.worldtradelaw.net/2014/05/theres-no-consistency-requirement-for-animal-welfare.html>> accessed 15 January 2020.

Lieb D, '13 States Launch New Legal Challenge to California Egg Law' (US News, 4 December 2017) <www.usnews.com/news/best-states/california/articles/2017-12-04/12-states-launch-new-legal-challenge-to-california-egg-law> accessed 3 February 2020.

Llewellyn K, 'Remarks on the Theory of Appellate Decision and the Rules or Canons about How Statutes Are to Be Construed' (1950) 3 Vanderbilt Law Review 395.

Lydgate E, 'Sorting out Mixed Messages under the WTO National Treatment Principle: A Proposed Approach' (2016) 15(3) World Trade Review 423.

Lydgate E, 'Is it Rational and Consistent? The WTO's Surprising Role in Shaping Domestic Public Policy' (2017) 20 Journal of International Economic Law 561.

Marceau G, 'A Comment on the Appellate Body Report in EC-Seal Products in the Context of the Trade and Environment Debate' (2014) 23(3) Review of European Community & International Environmental Law 319.

Marwell J, 'Trade and Morality: the WTO Public Morals Exception after Gambling' (2006) 81 New York University Law Review 816.

Mavroidis P, 'Sealed with a Doubt: EU, Seals, and the WTO' (2015) 6 European Journal of Risk Regulation 388.

McConville M and Chui W H, *Research Methods for Law* (Edinburgh University Press 2007).

McGrady B, 'Necessity Exceptions in WTO Law: Retreaded Tyres, Regulatory Purpose and Cumulative Regulatory Measures' (2008) 12 Journal of International Economic Law 153.

Ming Du M, 'Autonomy in Setting Appropriate Level of Protection under the WTO Law: Rhetoric or Reality?' (2010) 13 Journal of International Economic Law 1077.

Mitchell A, Heaton D and Henckels C, *Non-Discrimination and the Role of Regulatory Purpose in International Trade and Investment Law* (Elgar 2016).

Nachmani T, 'To Each His Own: The Case for Unilateral Determination of Public Morality under Article XX(a) of the GATT' (2013) 71 University of Toronto Faculty of Law Review 31.

Nagy C, 'Clash of Trade and National Public Interest in WTO Law: The Illusion of "Weighing and Balancing" and the Theory of Reservation' (2020) 23(1) Journal of International Economic Law 143.

Nuzzo S, 'Tackling Diversity inside WTO: GATT Moral Clause after Colombia-Textiles' (2017) 10 European Journal of Legal Studies 267.

Ortino F, 'GATT' in Daniel Bethlehem, Isabelle Van Damme, Donald McRae, and Rodney Neufeld (eds), *The Oxford Handbook of International Trade Law* (OUP 2009).

Piazza J, Ruby M, Loughnan S, Luong M, Kulik J, Watkins H and Seigerman M, 'Rationalizing Meat Consumption: the 4Ns' (2015) 91 *Appetite* 114.

Perišin T, 'Is the EU Seal Products Regulation a Sealed Deal? EU and WTO Challenges' (2013) 62(2) *International & Comparative Law Quarterly* 373.

Pirker B, *Proportionality Analysis and Models of Judicial Review: A Theoretical and Comparative Study* (Europa Law Publishing 2013).

Posner R, 'Animal Rights: Legal, Philosophical, and Pragmatic Perspectives' in Cass Sunstein and Martha Nussbaum (eds), *Animal Rights: Current Debates and New Directions* (OUP 2006).

Pound R, 'Law in Books and Law in Action' (1910) 44 *American Law Review* 12.

Radin M, 'Theory of Judicial Decision Or How Judges Think' (1925) 11 *ABA Journal* 357.

Rachels J, 'Drawing Lines' in Cass Sunstein and Martha Nussbaum (eds), *Animal Rights: Current Debates and New Directions* (OUP 2006).

Regan D, 'Regulatory Purpose and "Like Products" in Article III:4 of the GATT (with Additional Remarks on Article III:2)' (2002) 36(3) *Journal of World Trade* 443.

Regan D, 'Further Thoughts on the Role of Regulatory Purpose Under Article III of the General Agreement on Tariffs and Trade: A Tribute to Bob Hudec' (2003) 37(4) *Journal of World Trade* 737.

Regan D, 'What Are Trade Agreements For? – Two Conflicting Stories Told by Economists, With a Lesson for Lawyers' (2006) 9 *JIEL* 951.

Regan D, 'The Meaning of 'Necessary' in GATT Article XX and GATS Article XIV: The Myth of Cost-Benefit Balancing' (2007) 6 *World Trade Review* 347.

Regan D, 'International Adjudication: A Response to Paulus – Courts, Custom, Treaties, Regimes, and the WTO' in S Besson and J Tasioulas (eds), *The Philosophy of International Law* (Oxford University Press New York 2010).

Regan D, 'Explaining Trade Agreements: The Practitioners' Story and the Standard Model' (Robert Schuman Centre for Advanced Studies 2014, EUI Working Paper RSCAS 2014/113).

Regan T, *Animal Rights, Human Wrongs: An Introduction to Moral Philosophy* (Rowman & Littlefield 2003).

Rothgerber H, 'Efforts to Overcome Vegetarian-Induced Dissonance among Meat Eaters' (2014) 79 *Appetite* 32.

Ruggie J, 'International Regimes, Transactions, and Change: Embedded Liberalism in the Postwar Economic Order' (1982) 36(2) *International Organization* 379.

Russell J S, 'The Critical Legal Studies Challenge to Contemporary Mainstream Legal Philosophy' (1986) 18(1) *Ottawa Law Review* 1.

Sellheim N, 'Policies and Influence: Tracing and Locating the EU Seal Products Trade Regulation' (2015) 17 International Community Law Review 3.

Sellheim N, 'The Legal Question of Morality: Seal Hunting and the European Moral Standard' (2016) 25(2) Social and Legal Studies 141.

Shaffer G and Pabian D, 'European Communities—Measures Prohibiting the Importation and Marketing of Seal Products' (2015) 109(1) AJIL 154.

Shermer M, *The Moral Arc* (Henry Holt 2015).

Simon D, *Meatonomics* (Conari Press 2013).

Simon H, *Models of Man: Social and Rational* (Wiley 1957).

Singer P, *Animal Liberation* (1st edition, HarperCollins 1975).

Singer P, *Animal Liberation* (3rd edition, HarperCollins 2002).

Singer P, 'Ethics beyond Species and beyond Instincts: A Response to Richard Posner' in Cass Sunstein and Martha Nussbaum (eds), *Animal Rights: Current Debates and New Directions* (OUP 2006).

Singer P, 'Speciesism and Moral Status' (2009) 40(3)(4) Metaphilosophy 567.

Singer P, *Practical Ethics* (3rd edition, CUP 2011).

Singer P and Regan T, 'Dog in the Lifeboat: An Exchange' *New York Review of Books* (25 April 1985).

Smith T, 'Much Needed Reform in the Realm of Public Morals: A Proposed Addition to the GATT Article XX(a) 'Public Morals' Framework, Resulting from China-Audiovisual' (2011) 19:3 Cardozo Journal of International and Comparative Law 733.

Sowery, K 'Sentient Beings and Tradable Products: The Curious Constitutional Status of Animals under Union Law' (2018) 55(1) Common Market Law Review 1.

Steger D, 'The Culture of the WTO: Why It Needs to Change' (2007) 10(3) JIEL 483.

Stewart K and Cole M, 'The Conceptual Separation of Food and Animals in Childhood' (2009) 12(4) Food, Culture & Society 457.

Sunstein C, 'Moral Heuristics and Moral Framing Lecture' (2003) 88 Minnesota Law Review 1556.

Sunstein C, 'Moral heuristics' (2005) 28 Behavioural and Brain Sciences 531.

Sunstein C, 'Introduction: What Are Animal Rights?' in Cass Sunstein and Martha Nussbaum (eds), *Animal Rights: Current Debates and New Directions* (OUP 2006).

Sykes A, 'The Least Restrictive Means' (2003) 70 University of Chicago Law Review 403.

Sykes A, 'Regulatory Consistency Requirements in International Trade' (2017) 49 Arizona State Law Journal 821.

Sykes K, 'Sealing Animal Welfare into the GATT Exceptions: the International Dimension of Animal Welfare in WTO disputes' (2014) 13(3) World Trade Review 471.

Sykes K, 'Globalization and the Animal Turn: How International Trade Law Contributes to Global Norms of Animal Protection' (2016) 5(1) Transnational Environmental Law 55.

Thaler R, *The Winner's Curse: Paradoxes and Anomalies of Economic Life* (Simon & Schuster 1992).

Trachtman J, 'Trade and... Problems, Cost-Benefit Analysis and Subsidiarity' (1998) 9 European Journal of International Law 32.

Tushnet M, 'Following the Rules Laid down: A Critique of Interpretivism and Neutral Principles' (1983) 96(4) Harvard Law Review 781.

Tushnet M, 'Critical Legal Studies and Constitutional Law: An Essay in Deconstruction' (1984) 36 Stanford Law Review 623.

Voon T, 'Exploring the Meaning of Trade-Restrictiveness in the WTO' (2015) 14(3) World Trade Review 451.

Weiler JHH, 'The Rule of Lawyers and the Ethos of Diplomats: Reflections on the Internal and External Legitimacy of WTO Dispute Settlement' (Harvard Law School 2000, Harvard Jean Monnet Working Paper 9/00).

Weiler JHH, 'Law, Culture, and Values in the WTO – Gazing into the Crystal Ball' in Daniel Bethlehem, Isabelle Van Damme, Donald McRae, and Rodney Neufeld (eds), *The Oxford Handbook of International Trade Law* (OUP 2009).

Weiler JHH, 'Comment on Brazil - Measures Affecting Imports of Retreaded Tyres' (2009) 8 World Trade Review 137.

Wu M, 'Free Trade and the Protection of Public Morals: An Analysis of the Newly Emerging Public Moral Clause Doctrine' (2008) 33 The Yale Journal of International Law 215.

Zamir E and Teichman D (eds), *The Oxford Handbook of Behavioral Economics and the Law* (OUP, 2014).

Zamir E and Teichman D, 'Judicial Decision-Making: A Behavioral Perspective' in Zamir E and Teichman D (eds), *The Oxford Handbook of Behavioral Economics and the Law* (OUP, 2014) 664.

Other Official Documents

California Proposition 12, ‘Farm Animal Confinement Initiative’ (approved on 6 November 2018) <[https://ballotpedia.org/California_Proposition_12,_Farm_Animal_Confinement_Initiative_\(2018\)](https://ballotpedia.org/California_Proposition_12,_Farm_Animal_Confinement_Initiative_(2018))> accessed 5 April 2020 (“California Proposition 12”).

Commission Regulation (EU) No 737/2010 of 10 August 2010 laying down detailed rules for the implementation of Regulation (EC) No 1007/2009 of the European Parliament and of the Council on trade in seal products (2010) OJ L 216/1.

Council Regulation (EC) No 1099/2009 of 24 September 2009 on the protection of animals at the time of killing (2009) OJ L 303/1.

EFSA, ‘Opinion of the Scientific Panel on Animal and Welfare on a Request from the Commission Related to Welfare Aspects of the Main Systems of Stunning and Killing the Main Commercial Species of Animals’ (2004) 45 The EFSA Journal 1.

EFSA, ‘Scientific Opinion of the Panel on Animal Health and Welfare on a Request from the Commission on the Animal Welfare Aspects of the Killing and Skinning of Seals’ (2007) 610 The EFSA Journal 1.

EU, ‘First Written Submission by the European Union’ (*EC – Seals*, 21 December 2012) <https://trade.ec.europa.eu/doclib/docs/2012/december/tradoc_150190.pdf> (consulted 5 April 2020) (“EU First Panel Submission”).

EU, ‘Second Written Submission by the European Union’ (*EC – Seals*, 27 March 2013) <https://trade.ec.europa.eu/doclib/docs/2013/april/tradoc_150812.pdf> (consulted 5 April 2020) (“EU Second Panel Submission”).

Howse R, Langille J and Sykes K, ‘Written Submission of Non-Party Amici Curiae’ (*EC – Seals*, 11 February 2013) <<http://www.worldtradelaw.net/amicus/howsesealsamicus.pdf.download>> (consulted 5 April 2020) (“Amici Panel Submission”).

Howse R, Langille J and Sykes K, ‘Written Submission of Non-Party Amici Curiae’ (*EC – Seals*, 17 March 2014) <<http://stage.worldtradelaw.gvpi.net/amicus/HowseSealsAmicusAB.pdf.download>> (consulted 5 April 2020) (“Amici Appellate Submission”).

Regulation (EC) No 1165/2008 of the European Parliament and of the Council of 19 November 2008 concerning livestock and meat statistics (2008) OJ L 321/1.

Regulation (EC) No 1007/2009 of the European Parliament and of the Council of 16 September 2009 on trade in seal products (2009) OJ L286/36.

World Organisation for Animal Health, *Terrestrial Animal Health Code* (20th edition, 2011) (OIE Guidelines).