‘MANIFEST FAILING’: INVESTIGATING THE SUBSTANTIVE THRESHOLD FOR COLLECTIVE INTERNATIONAL ACTION IN RESPONSE TO MASS ATROCITY CRIMES

Blagovesta Dimova Tacheva

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

School of Politics and International Studies

March 2021

INTELLECTUAL PROPERTY AND PUBLICATIONS STATEMENTS

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

ACKNOWLEDGEMENTS

Five years after choosing the concept of 'genocide' as the topic of my first undergraduate assignment at the University of Sheffield, I returned to study the same notion in my PhD thesis, as one of the four mass atrocity crimes it examines. First, I would like to thank the ESRC, the Department of Politics at the University of Sheffield and the School of Politics and International Studies at the University of Leeds for granting me the PhD studentship, without which this would not have been possible. Having originally planned to finish my studies after my first university degree, the fact that I embarked on a PhD course, for which I received ESCR funding, is a fairly low-probability turn of events. One person deserves substantial credit for this dramatic change of plans.

Professor Garrett Wallace Brown has been a trustworthy primary supervisor, mentor and I dare say, friend, in the past 10 years. I met him during my undergraduate degree at the University of Sheffield, while working as a SURE Scheme Research Assistant under his supervision on a project at the Department of Politics. Following my BA degree, during which he acted as my third-year dissertation supervisor, he encouraged me to continue my education to Master's level and later to apply for a PhD. He has supported me in my PhD journey from the start at the University of Sheffield through to the end of my tenure as PhD student at the University of Leeds. In this time, he has provided invaluable advice to every problem that has affected my PhD work and has gone above and beyond to help me develop my ideas, put them in writing and revise them. I am greatly indebted to him for his patience and help throughout my PhD triumphs, trials and tribulations.

One of the factors that played a role in my decision to transfer to the University of Leeds was the opportunity to benefit from the advice of Dr. Adrian Gallagher, who has acted as my second supervisor in the last three years. Reading his articles, outlining a new research agenda in the study of the R2P at the time when I was trying to identify a topic that could hold my interest for the duration of a PhD, inspired me to pursue the questions that remained to be answered with regards to the 'manifest failing' threshold for international action in response to atrocity crimes, which is the subject of this thesis. Eventually, following my transfer to the University of Leeds, I had the opportunity to benefit from Adrian's astute advice, expertise on the subject and extensive knowledge of R2P cases in the most critical years of my PhD. I believe my work was greatly improved as a result.

In addition to my supervisors, I would also like to thank Professor Alex Bellamy at the University of Queensland, who generously hosted me at the Asia-Pacific Centre for the Responsibility to Protect during my visit over the summer of 2018. My conversations with him gave me confidence that the research agenda I was pursuing in my PhD thesis was worthwhile. I am also grateful to Professor Bellamy for kindly introducing me to leading R2P academics and advocates.

Besides those who have supported me in my academic endeavours, several members of staff at the University of Sheffield and the University of Leeds have been simply outstanding in helping me along the way. I would like to thank in particular Sarah Cooke, a former colleague at the University of Sheffield and a friend, who has supported me since my time as an MA student applying for PhD funding, through my Confirmation Review and personal struggles during my time at Sheffield. I am also grateful to Caroline Wise (my first point of contact when I transferred to the University of Leeds and always a friendly face in the crowd) and Matthew Wilkinson, who have responded to my emails as quick as I could send them and whose knowledge and assistance 'behind-the-scenes' has helped me to overcome any hurdles I faced during my time as PhD student at the University of Leeds.

My colleagues and friends at the University of Sheffield, whom I will not mention by name as I am sure they will recognise themselves in these lines, have become close friends and I hope they would remain such for the rest of my life. This list would be incomplete without mentioning my brilliant loyal friends – Elena and Diana, who have been faithfully involved in my long journey through higher education from the very first day of my BA degree to the day of my PhD submission and who thoroughly deserve to never hear about this thesis again.

Last but not least, this is not a feat I could have accomplished without the unreserved support of my family, to whom I am eternally grateful for, among a myriad of other things, doing literally everything they could so that I can have the opportunities they never did. I dedicate this work to my family – my Mum, Dad, Grandfather, who have been incredibly supportive through this journey, and my Grandmother, who would have been proud to see me complete my BA, MA and PhD degrees. Each of them have contributed in their own way to any success I have had or may have in the future and made numerous sacrifices along the way. Thank you!

ABSTRACT

This thesis provides the first comprehensive empirical and normative investigation of the substantive threshold for international action in response to situations in which civilian populations are subjected to grave human rights violations which may constitute mass atrocity crimes, defined as genocide, crimes against humanity, war crimes and ethnic cleansing. The most recent iteration of this threshold is the 'manifest failing' requirement, found in paragraph 139 of the 2005 World Summit Outcome Document, detailing the agreement of world leaders on the Responsibility to Protect (R2P) concept. This thesis synthesises current understanding of 'manifest failing' and explains how this threshold requirement came into existence, what it represents and what its function within the R2P framework is. Crucially, it draws attention to the 'substantive indeterminacy' of the 'manifest failing' threshold and the negative implications thereof.

In a nutshell, this thesis argues that the ambiguity surrounding the substance of the 'manifest failure' test needs to be addressed, because in its current indeterminate form, it neither places meaningful substantive constraints on coercive action, nor does it offer an adequate basis for the international community to justify the need for such action with reference to this threshold. In order to remedy these and other negative implications of the substantive indeterminacy of the 'manifest failing' threshold for responding to instances in which mass atrocity crimes may be taking place, this thesis investigates existing contributions to clarifying the substance of the 'manifest failing' test, qualitative and quantitative tools for the analysis of atrocity crimes; traces the historical lineage of the 'manifest failing' test and draws inspiration from analogous threshold requirements it originates from, and; studies states' understanding of the substantive threshold for collective international action in the context of the 2011 atrocity crisis in Libya. Taking into account the relevant insights from these three lines of investigation, this thesis proposes the first holistic normative framework comprising substantive determinants and associated indicators that should inform 'manifest failing' inquiries. It then applies this framework to the ongoing 'war on drugs' in the Philippines to showcase how it can provide effective guidance for determining whether a national government is 'manifestly failing' to protect its populations from the four atrocity crimes.

Table of Contents

Intellectual Property and Publications Statements 1

Acknowledgements 2

Abstract 4

List of Tables and Illustrative Material 9

Abbreviations 10

Introduction: The 'Manifest Failing' Threshold 13

R2P and the 2005 World Summit 14

Research agenda and basic argument: Addressing the substantive indeterminacy of the 'manifest failing' threshold requirement 18

The 'manifest failing' threshold as a substantive licence for and a leash against coercive action 20

Addressing objections to the study of the 'manifest failing' threshold 27

Research objectives 32

A framework for the interpretation (not the application) of the R2P 35

A common framework consistent with the case-by-case approach to atrocity crises response 36

Contribution 37

Chapter outline 40

Chapter I: Contributions to Clarifying the Substance of the 'Manifest Failing' threshold and Relevant Tools for the Analysis of Atrocity Crimes 43

I. The UN 'Framework of Analysis for Atrocity Crimes' 44

II. The determinants of a 'manifest failing' 48

III. The indicators of a 'manifest failing' 59

IV. The Study of 'Manifest Failing' in the Libyan Context 75

Conclusion 79

Chapter II: Research Strategy 83

Triangulation 84

Extending the 'unable or unwilling' test 90

Deriving ‘manifest failing’ factors from the views of states 96
The case study method 104
Political Discourse Analysis 117
Employing discourse analysis to identify the factors of the ‘manifest failing’ threshold 121
Conclusion 129
Chapter III: The Origins of ‘Manifest Failing’ Threshold – the ‘Unable or Unwilling’ Test in International Law 130
I. The notion of a ‘manifestly failing’ state and its predecessor the ‘unable or unwilling’ state 131
II. The ‘unable or unwilling’ test in the context of self-defence against non-state actors 138
The origins of the ‘unable or unwilling’ test 139
Is the ‘unable or unwilling’ test part of customary international law? 143
III. The factors of the ‘unable or unwilling’ test 149
The Principles of International Law on the Use of Force in Self-Defence 150
Principles Relevant to the Scope of a State’s Right of Self-Defense Against an Imminent or Actual Armed Attack by Nonstate Actors 152
The factors of the ‘unwilling or unable’ test 155
Conclusion 174
Chapter IV: ‘Manifest Failing’ in Libya 178
The sources analysed 180
The Libyan crisis and international response 182
Part I: The debate over threshold in Libya 193
Deliberating the threshold for collective action in the language of ‘manifest failing’ and ‘unable or unwilling’ 193
Deliberating the threshold for collective action in the absence of the ‘manifest failing’/’unable or unwilling’ language 201
Conclusion 209

Part II: The Determinants and Indicators of Libya's 'manifest failing' 212
FRANCE 212
UK 218
US 228
SECURITY COUNCIL 241
Conclusion 256
Chapter V: A Normative Framework for 'Manifest Failing' Inquiries 262
I. Framework composition and utility 262
II. The Proposed Framework 265
Determinant 1: Human rights violations associated with atrocity crimes 270
Determinant 2: The gravity of atrocity-related human rights violations 270
Determinant 3: The imminence of the threat of atrocities 271
Determinant 4: Host state intent 275
Determinant 5: The host state's territorial control and capacity to supress the threat of atrocity crimes 276
Determinant 6: Measures taken by the host state to mitigate the threat of atrocities 278
Determinant 7: Host state cooperation 279
Determinant 8: The host state's response to international requests to address the threat 280
Determinant 9: The host state's track record of grave human rights violations and relevant interaction with the international community 282
III. The Philippines' 'War on Drugs' 284
Basic context 285
Human rights violations associated with atrocity crimes 287
The gravity of atrocity-related human rights violations 288
The imminence of the threat of atrocities 290
Government intent 292

The Philippines' territorial control and capacity 295
The measures taken by the Philippines to mitigate the risk of atrocities 297
The Philippines' willingness to cooperate with international organisations 299
The Philippines' response to international requests to address the threat 301
The Philippines' track record of violence and gross human rights violations 303
The Philippines' 'manifest failing' 304
Conclusion 307
Conclusion 309
Annex I: Definitions of mass atrocity crimes 316
Genocide 316
Crimes against humanity 317
War crimes 318
Ethnic cleansing 323
Bibliography 325

LIST OF TABLES AND ILLUSTRATIVE MATERIAL

Box 1 Rosenberg and Strauss' list of human rights violations relevant to past cases of mass atrocities……..55

Box 2 Rosenberg and Strauss' risk-factors for determining the level of gravity of prospective human rights violations considering the number of victims or extent of irreparable harm……...56

Box 3 Rosenberg and Strauss' list of steps / measures to mitigate the risk of mass atrocities……...……58

Table 1.1 'Manifest Failing' Determinants and Indicators Derived from Contributions to the Study of 'Manifest Failing' and Mass Atrocities…….80

Figure 1 Data Triangulation……....87

Table 2.1 Triangulation Protocol……90

Table 3.1 'Manifest Failing' Determinants Derived from Analogous 'Unable or Unwilling' Factors…….……175

Table 4.1 List of UNSC meetings which saw contributions from member states between 22 February 2011 and 22 December…….…….182

Table 4.2 The Determinants and Indicators of Libya's 'Manifest Failing' Cited by States…….…….259

Table 5.1 A Normative Framework for 'Manifest Failing' Inquiries……265

Table 5.2 The Philippines' 'Manifest Failing' in the context of Duterte's 'war on drugs'…….305

Box 4 Legal definition of genocide……...316

Box 5 Legal definition of crimes against humanity……317

Box 6 Legal definition of war crimes……318

ABBREVIATIONS

A/RES	Resolution of the UN General Assembly
AFP	Armed Forces of the Philippines
APR2P	Asia-Pacific Centre for the Responsibility to Protect
ATA	Anti-Terrorism Act (Philippines)
AU	African Union
ASEAN	Association of Southeast Asian Nations
BARMM	Bangsamoro Autonomous Region in Muslim Mindanao – an autonomous region located in the Southern Philippines
BRICS	Brazil, Russia, India, China and South Africa
CAR	Central African Republic
CHR	Commission on Human Rights of the Philippines
E10	Elected 10 - non-permanent members of the UN Security Council
EU	European Union
FSI	Fragile State Index, formerly known as he Failed States Index
GCR2P	Global Centre for the Responsibility to Protect
GCC	Gulf Cooperation Council
HLP	High-level Panel on Threats, Challenges and Change of the United Nations
HRC	Human Rights Council of the United Nations

ICC	International Criminal Court
ICISS	International Commission on Intervention and State Sovereignty
ISIL	Islamic State of Iraq and the Levant
LAS	League of Arab States or Arab League
MILF	Moro Islamic Liberation Front
NATO	North Atlantic Treaty Organisation
NGO	Non-governmental organisation
NPA	New People's Army
NSA	Non-state actor
OHCHR	Office of the United Nations High Commissioner for Human Rights
OIC	Organisation of Islamic Cooperation, formerly the Organisation of the Islamic Conference
OTP	Office of the Prosecutor of the International Criminal Court
P2	Permanent Two members of the UN Security Council - China and Russia
P3	Permanent Three members of the UN Security Council - France, UK and USA
P5	Permanent Five members of the UN Security Council - China, France, Russia, UK and US
PDEA	Philippine Drug Enforcement Agency
PNP	Philippine National Police
R2P	Responsibility to Protect (as specified in the 2005 WSOD unless otherwise stated)

UNSCR	Security Council Report
S/PV	Meeting record of the Security Council
S/RES	Security Council Resolution
TMK	Targeted Mass Killing
UK	United Kingdom of Great Britain and Northern Ireland
UN	United Nations
UNGA	United Nations General Assembly
UNSC	United Nations Security Council
UNSG	United Nations Secretary General
UoU	Unwilling or Unable
US	United States of America
WoD	‘War on Drugs’ in the Philippines
WSOD	World Summit Outcome Document (2005)
WGI	World Governance Indicators (WGI)

INTRODUCTION: THE 'MANIFEST FAILING' THRESHOLD

Twenty years ago, Nicholas Wheeler wrote: 'The dilemma of what to do about strangers who are subjected to appalling cruelty by their governments has remained with us throughout the post-1945 world'.[1] While the language and focus of the debate have changed since then, the inaugural question of Wheeler's oft-cited book 'Saving Strangers' remains fundamentally the same. So does the rudimentary context in which it arises. National governments face the scrutiny of the international community and international and non-governmental organisations over their domestic conduct, especially with regards to their legal obligations towards their own populations put in place by the UN family.[2] Nevertheless, states routinely fail to respect the resulting limits to how they can treat individuals and groups within their territory. When states do not live up to these obligations under international law in the face of the imminent or ongoing mass violence against their people and severe deprivation of their fundamental rights, the following question arises: *Under what circumstances should the international community consider responding coercively to the threat posed to vulnerable populations by the (ir)responsible host state?*. As the question implies, a certain threshold must be met for such actions to be warranted and justifiable. This thesis will focus precisely on this threshold for coercive international action to protect populations at risk from the so-called mass atrocity crimes – genocide, ethnic cleansing, crimes against humanity and war crimes.[3] The most widely accepted expression of this notion is the 'manifest failing' threshold requirement found in the Responsibility to Protect (R2P) framework, agreed upon by over 150 heads of state and governments at the 2005 World Summit.

In a nutshell, this thesis starts from the premise that the ambiguity surrounding the substance of the 'manifest failure' threshold needs to be addressed, because in its current indeterminate form, it neither places meaningful substantive constraints on coercive action, nor does it offer an adequate basis for the international community to justify the

[1] Nicholas Wheeler, *Saving Strangers: Humanitarian Intervention in International Society* (New York: Oxford University Press, 2000), p. 1.

[2] Ibid.

[3] For the legal definitions of genocide, crimes against humanity and war crimes, and an operational definition of ethnic cleansing (which has not been recognised as an independent international crime under international law) please consult Annex I.

need for such action with reference to this threshold. As a result, we are left with an undertheorised and underspecified threshold requirement, whose ability to effectively guide, legitimate or delegitimise decisions on 'timely and decisive' action in response to atrocities is undermined. Thus, in its most general formulation, the principal research question at the heart of this thesis is: What determines a 'manifest failing'?. More specifically, the aim of this thesis is to rigorously investigate what the qualitative and quantitative criteria relevant to determining whether a state is 'manifestly failing' in its primary responsibility to protect its populations are, in a bid to devise a normative framework of factors that will help guide inquiries into whether a host state is 'manifestly failing' in its primary responsibility to protect. To do so, this thesis draws insights from existing research on mass atrocity crimes and the 'manifest failing' threshold, from analogous areas of international law, and from case study research.

R2P AND THE 2005 WORLD SUMMIT

Fifteen years on from its adoption, the R2P has become a notable, yet controversial, fixture in international debates about the protection of innocent populations from the four atrocity crimes. Some have even argued that 'the R2P has become the decision-making framework in which to deliberate – and take action' – to protect populations from such violence.[4] This thesis will not go as far as to defend this or any specific claims about the influence of the R2P framework on decision-making. However, it will operate on its premises, because it has provided the conceptual basis for recent state deliberations and scholarly debates over the threshold for international action in response to atrocity situations, which have been increasingly couched in R2P-terms.

The R2P framework was specified in Paragraphs 138-140 of the 2005 World Summit Outcome Document (WSOD), which stipulates two complementary responsibilities that states have with regards to protecting populations from the four crimes – the primary responsibility of individual states to protect their populations and a subsidiary responsibility of the international community 'to do what it can to ensure that populations are protected from atrocity crimes'.[5] The former is firmly established in international law

[4] Tim Dunne and Katharine Gelber, 'Arguing Matters: The Responsibility to Protect and the Case of Libya', 6 (3) *Global Responsibility to Protect* (2014): 326-349, p. 329.

[5] United Nations General Assembly, 2005 World Summit Outcome, A/RES/60/1, 24 October 2005, paras. 138-140.; Luke Glanville, 'Does R2P matter? Interpreting the impact of a norm', 51 (2) *Cooperation and Conflict* (2016): 184-199, p. 186.

and commonly understood as uncontroversial.[6] The latter is an assemblage of intersecting extralegal obligations. Paragraph 138 instructs that 'the international community *should*, as appropriate, *encourage and help* States to exercise this responsibility and support the United Nations in establishing an early warning capability' (emphasis added).[7]

In addition to this obligation to assist the host state in its primary protection responsibility, Paragraph 139 stipulates a two-tier responsibility of the international community to take collective action in response to atrocity crises.[8] Bellamy aptly distinguishes between the two tiers of collective action, respectively specified in the first and second sentence of Paragraph 139, as the 'first response' and 'last resort' component of R2P's reactive dimension.[9] The opening sentence of this paragraph specifies the responsibility of 'the international community, through the United Nations … to use appropriate diplomatic, humanitarian and other peaceful means, in accordance with Chapters VI and VIII of the Charter, to help protect populations from' the four crimes. It is understood 'to be an ongoing, generic responsibility that employs the kind of peaceful, pacific measures specified in Chapter VI and in Article 52, Chapter VIII'.[10] Article 33 (1) of Chapter VI details a wide range of pacific measures that '[t]he parties to any dispute, the continuance of which is likely to endanger the maintenance of international peace and security' can use to resolve it, namely 'negotiation, enquiry, mediation, conciliation, arbitration, judicial settlement, resort to regional agencies or arrangements, or other peaceful means of their own choice'.[11] Article 52 of Chapter VIII specifies the role of regional arrangements in the peaceful settlement of disputes.[12]

The second sentence of Paragraph 139 stipulates the adoption of a wider range of measures in the following statement: '*we are prepared to take collective action*, in a timely and decisive manner, *through the Security Council*, in accordance with the Charter, *including Chapter VII*, on a case-by-case basis and in cooperation with relevant regional

[6] Ibid.

[7] A/RES/60/1, para. 138.

[8] Similarly, Bellamy suggests that 'The wording agreed by states in 2005 suggests that Pillar III comprises two steps.': Alex J. Bellamy, 'The Three Pillars of the Responsibility to Protect', 41 (20) *Pensamiento Propio* (2015): 35-64, p. 49.

[9] Alex J. Bellamy, 'The First Response: Peaceful Means in the Third Pillar of the Responsibility to Protect', Policy analysis (Muscatine, IA: The Stanley Foundation, December 2015), p. 11.

[10] Ban Ki-moon, 'Implementing the Responsibility to Protect', Report of the Secretary-General, A/63/677 (12 January 2009), para. 49.

[11] Article 33, UN Charter.

[12] Article 52, UN Charter.

organizations as appropriate, *should peaceful means be inadequate and national authorities manifestly fail to protect their populations* from genocide, war crimes, ethnic cleansing and crimes against humanity' (emphases added).[13] To elaborate, Chapter VII of the UN Charter 'provides the framework within which the Security Council may take enforcement action' to respond to threats to 'international peace and security'.[14] Within this framework for coercive action, the use of force specified in Article 42 is the pinnacle of a wide spectrum of less intrusive forcible measures. Article 41 provides for a range of coercive measures short of the use of force, comprising 'complete or partial interruption of economic relations and of rail, sea, air, postal, telegraphic, radio, and other means of communication, and the severance of diplomatic relations'.[15] Historically, sanctions, be they diplomatic or in the form of arms embargoes, asset freezes, travel bans, and commodity interdictions, have been imposed on designated states, non-state entities and individuals to dissuade the perpetrators of atrocity crimes.[16] Likewise, Article 40 specifies that '[i]n order to prevent an aggravation of the situation, the Security Council may, before making the recommendations or deciding upon the measures provided for in Article 39 [(i.e. measures in accordance with Articles 41 and 42)], call upon the parties concerned to comply with… provisional measures'.[17] Although the latter are not explicitly listed in the Charter, they 'are distinct from recommendations made under Chapter VI [and] include a withdrawal of armed forces, a cessation of hostilities, a conclusion or observance of a ceasefire or a creation of the conditions necessary for unimpeded delivery of humanitarian assistance'.[18]

It is important to note that, unlike the primary responsibility of host states to protect their populations, the responsibilities of the international community set out in Paragraph 139 of the WSOD neither correspond to, nor create new obligations for state action under international law. As Michael Doyle reminds: 'A General Assembly resolution [such as the 2005 World Summit Outcome] is recommendatory; it does not make binding

[13] A/RES/60/1. para. 139.

[14] United Nations Security Council, 'Actions with Respect to Threats to the Peace, Breaches of the Peace, and Acts of Aggression' [Accessed 12 May 2020] Available at: https://www.un.org/securitycouncil/content/repertoire/actions

[15] Article 41, UN Charter.

[16] Security Council Report, 'Special Research Report: UN Sanctions', No 3, 25 November 2013, pp. 8-12. [Accessed 21 January 2021] Available at: https://www.securitycouncilreport.org/atf/cf/%7B65BFCF9B-6D27-4E9C-8CD3-CF6E4FF96FF9%7D/special_research_report_sanctions_2013.pdf

[17] Article 40, UN Charter.

[18] UNCS, 'Actions with Respect to Threats to Peace'.

international law'.[19] Even so, as Jennifer Welsh argues, Paragraph 139 of the WSOD did not openly specify 'an *international* responsibility to protect, which would be automatically activated if the state's primary responsibility to protect its population from mass atrocity crimes is not fulfilled'.[20] Instead, the phrase 'prepared to take collective action' in the second sentence of Paragraph 139 was deliberately chosen in the final stages of formulating the WSOD draft, so as to water down the international obligations stipulated in the R2P and, along with the element of case-by-case decision-making, to ensure that it 'would not go beyond the prevailing mechanisms – including the political negotiations among the P5 [(the five permanent members)] in the Council – in responding to mass atrocity situations'.[21]

Nonetheless, this coercive aspect of the R2P remains the most complex and conditional responsibility recognised in the WSOD. In addition to Security Council authorisation, the Outcome Document specifies two conditions upon which sanctioning coercive measures under Chapter VII is premised, namely 1) 'should peaceful means be inadequate' (that is, the peaceful measures in accordance with Chapters IV and VIII specified in the first sentence of Paragraph 139); and 2) 'national authorities manifestly fail to protect their populations from the four crimes'.[22] Together, they constitute the threshold for coercive action in mass atrocity situations under the umbrella of the R2P framework. With regard to the first requirement, Bellamy explains that 'the language used by paragraph 139 of the World Summit document echoes that of Article 42 of the UN Charter, which grants to the UN Security Council authority to take military or other types of action in situations where peaceful measures "would be inadequate or have proved to be inadequate"'.[23] While this first requirement is fairly straightforward, what the second requirement – the 'manifest failing' of the host state – entails, remains unclear from the text of the Outcome

[19] Michael W. Doyle, 'The Politics of Global Humanitarianism: The Responsibility to Protect before and after Libya', 53 (1) *International Politics: Responsibility to Protect and Humanity* (2016): 14-31, p. 19; A/RES/60/1 (2005).

[20] Jennifer M. Welsh, 'Norm Contestation and the Responsibility to Protect', 5 (4) *Global Responsibility to Protect* (2013): 365–396, p. 377.

[21] Ibid.

[22] This interpretation of the text of the WSOD is consistent with Alex Bellamy's detailed work on clarifying the distinction between the responsibilities specified in Paragraph 139 of the WSOD.: Bellamy, 'The First Response', p. 11; Bellamy, 'The Three Pillars of the Responsibility to Protect', pp. 49, 55.

[23] Bellamy, 'The First Response', p. 11. Contrary to claims 'that paragraph 139 sets a higher bar than Article 42 by requiring that peaceful means be proven inadequate', Bellamy makes a convincing case that 'it is difficult to find support for this judgement either in the paragraph's text ("should peaceful means be inadequate" does not specify that proof of inadequacy is needed, only that the council make a judgment that they are inadequate) or the travaux préparatoires ("preparatory works") of the World Summit, which indicate no intention to amend the requirements of Article 42 for application with R2P'.: Ibid. pp. 11-12.

Document. This has resulted in a profound lack of understanding as to what constitutes a 'manifest failing' and a number of negative implications that stem from the substantive indeterminacy of this threshold requirement.[24] Therefore, it is precisely this requirement, henceforth referred to as the 'manifest failing' threshold or the 'manifest failing' test, which will be the focus of this thesis.

RESEARCH AGENDA AND BASIC ARGUMENT: ADDRESSING THE SUBSTANTIVE INDETERMINACY OF THE 'MANIFEST FAILING' THRESHOLD REQUIREMENT

By placing the requirement of a 'manifest failing' at the heart of the international community's responsibility to act collectively to protect populations at risk without offering guidance as to what it entails or what type of evidence is needed to determine when it has been met, the drafters of the WSOD fashioned a rather peculiar threshold for R2P's sharp edge. On the one hand, this has led states, scholars, R2P advocates, and international organisations to cite it frequently without discussing its meaning, in a way that suggests that it is firmly established. As Chapter IV illustrates, the 'manifest failing' terminology and alternative formulations also featured in UN member states' discussion of the Libyan crisis. Notably, over the past decade 'the UN secretary-general, his special advisers on genocide prevention and R2P, and the UN high commissioner for human rights [have not only used] the R2P in their public advocacy',[25] but have also specifically invoked the 'manifest failing' threshold in a bid to persuade the UN Security Council to take timely and decisive action in Libya (see Chapter IV). The Global Centre for the Responsibility to Protect (GCR2P), whose mission is to 'sav[e] lives by mobilizing the international community to act in situations where populations are at risk of mass atrocity crimes', consistently refers to the 'manifest failing' of national governments in an attempt to advocate for the need for international action in situations at high risk of atrocities.[26] According to the latest publications of the 'R2P monitor', the GCR2P has determined that

[24] The term 'substantive indeterminacy' was originally used in a similar way by Ashley Deeks with regards to the 'unable or unwilling' test in the context of self-defence against non-state actors, where she argues that the test's lack of substance undermines its legitimacy and the legitimacy of potential uses of force.: Ashley Deeks, '"Unwilling or Unable": Toward a Normative Framework for Extraterritorial Self-Defense', 52 (3) *Virginia Journal of International Law* (2012): 483-550, p. 503.

[25] Bellamy, 'The First Response', p. 33.

[26] GCR2P website, 'About us'. [Accessed 08 August 2020]. Available at: https://www.globalr2p.org/about/

the national governments of Syria, Myanmar, Cameroon, Venezuela, China, and Ethiopia are currently failing to fulfil their responsibility to protect their populations.[27]

On the other hand, the reality is that the substance of the 'manifest failing' requirement remains far from clear or agreed upon. This can be attributed to the fact that in order to facilitate agreement on paragraphs relevant to the R2P, 'ambiguity was preferred over clarity' with regards to specifying the responsibilities of the international community in the WSOD.[28] According to Widmaier and Glanville, this ambiguity can be seen in states' recognition of the collective responsibility of the international community to protect populations… through a variety of means, spanning efforts to 'encourage and help' states, or to pressure them to protect their populations through 'appropriate diplomatic, humanitarian and other peaceful means'.[29] Likewise, the two authors underscore that 'the responsibility to take military action was expressed in equally ambiguous terms: the international community was 'prepared' to take collective action and only when states were 'manifestly failing' in efforts to protect their populations, and the decision to act through the Security Council was to be undertaken 'on a case-by-case basis'.[30] While Widmaier and Glanville point out that constructive ambiguity can provide 'a basis for wider consensus' among agents with various predispositions and sustain it by allowing them 'to adapt to evolving interpretations of the norm', they also acknowledge that 'too much ambiguity can impede norm development'.[31] In the case of the 'manifest failing' threshold, while constructive ambiguity has elicited weak normative buy-in, the lack of continued dialogue to incrementally clarify its substance over time has rendered it inert, which has created confusion in the long term. The sustained substantive indeterminacy of the 'manifest failing' threshold has a number of negative practical implications, which impede the consistent interpretation of situations in which mass atrocity crimes may arise with a view to determining the need for collective action. Crucially, in its current

[27] GCR2P, 'R2P Monitor', Issue 55, 15 January 2021. [Accessed 31 January 2021]. Available at: https://www.globalr2p.org/publications/r2p-monitor-issue-55-15-january-2021/; GCR2P, 'R2P Monitor', Issue 56, 15 March 2021. [Accessed 20 March 2021]. Available at: https://www.globalr2p.org/publications/r2p-monitor-issue-56-15-march-2021/

[28] Wesley W. Widmaier and Luke Glanville, 'The benefits of norm ambiguity: constructing the responsibility to protect across Rwanda, Iraq and Libya', 21 (4) *Contemporary Politics* (2015): 367-383, p. 375.

[29] Ibid.

[30] Ibid.

[31] Ibid., pp. 368, 370, 371.

indeterminate state, the 'manifest failing' requirement suffers to fulfil its function as a substantive threshold for international action.

The 'manifest failing' threshold as a substantive licence for and a leash against coercive action

It is well known that international law ordinarily prohibits coercive action against a sovereign state. At the heart of the UN Charter lies the prohibition on the use of force 'against the territorial integrity or political independence of any State' laid down in Article 2 (4), whereas Article 2 (7) stipulates that '[n]othing contained in the… Charter shall authorize the United Nations to intervene in matters which are essentially within the domestic jurisdiction of any state'. Nonetheless, Article 39 (Chapter VII) of the Charter carves out an exception to these injunctions, allowing the UN Security Council to mandate the application of coercive measures specified in Articles 41 and 42 when 'necessary to maintain or restore international peace and security'. Although the traditional body of rules governing the use of force and coercive action does not hold provisions consistent with the basis for such action under R2P, the heads of state and government at the 2005 World Summit reaffirmed 'that the relevant provisions of the Charter are sufficient to address the full range of threats to international peace and security' in Paragraph 79 of the WSOD.[32] That this recognition was followed by tying the implementation of R2P's coercive aspect to the UNSC's authority to mandate enforcement measures under Chapter VII (in Paragraph 139) suggested that world leaders relied on the Council to expand its interpretation of 'international' threats to include the protection of innocent populations from the threat of mass atrocities. The understanding 'that widespread human suffering can represent a threat to the maintenance of international peace and security and therefore can fall under the Security Council's responsibility under Chapter VII… is fairly widely accepted', despite reservations over mandating coercive action in instances where extensive human rights abuses occur in the absence of such clear 'international' threat.[33] In view of Security Council practice, Bellamy and Williams observe that 'the regular use of Chapter VII to mandate peace operations to protect civilians [(albeit in the presence of some form of government

[32] A/RES/60/1, para. 79.

[33] Edward Newman, 'Review Article – Humanitarian Intervention Legality and Legitimacy', 6 (4) *The International Journal of Human Rights* (2002): 102-120, p. 104.

consent)] represent[s] an important development in the Council's thinking'.[34] In addition, as Chapter IV of this thesis will evidence, states who voted in favour of the Libyan intervention in 2011 expressed a firm belief that a Security Council mandate provided a clear legal basis for the use of force for civilian protection. It is in this context that the WSOD requirement for UN Security Council authorisation under Chapter VII is understood as providing a legal basis for sanctioning coercive action against a sovereign state to protect vulnerable populations from the crimes of genocide, ethnic cleansing, war crimes and crimes against humanity.[35]

However, if the legality of coercive action in mass atrocity situations was the only thing that mattered, why did the WSOD set two additional threshold requirements (i.e. that peaceful means be inadequate and national authorities 'manifestly failing')? The answer to this question is convincingly articulated by the UN High-level Panel (HLP) on Threats Challenges and Change in a 2004 report titled 'A More Secure World: Our Shared Responsibility':

> 'The effectiveness of the global collective security system, as with any other legal order, depends ultimately not only on the legality of decisions but also on the common perception of their legitimacy — their being made on solid evidentiary grounds, and for the right reasons, morally as well as legally'.[36]

Taking the Libyan intervention as an instance of the application of R2P's coercive aspect, Michael Doyle explains how it 'joined legality (Security Council approval) to legitimacy (the cause of protecting civilians)'.[37] In general terms, Doyle argues that even though R2P does not impose binding legal obligation on the international community, it was successful in broadening the scope of 'legitimate armed intervention by licensing some (protective) interventions but only because it was seen as a [procedural] leash against other (exploitative) interventions' owing to the requirement for Security Council authorisation.[38] Yet he also implies that the Security Council's role as the 'gatekeeper of

---

[34] Alex J. Bellamy and Paul D. Williams, 'The new politics of protection? Côte d'Ivoire, Libya and the responsibility to protect', 87 (4) *International Affairs* (2011): 825-850, p. 828.

[35] On the legal basis for humanitarian intervention see: Simon Chesterman, 'Legality versus legitimacy: humanitarian intervention, the Security Council, and the rule of law', 33 (3) *Security Dialogue* (2002): 293-307, p. 295.; See also Newman, 'Review Article', pp. 103-104.

[36] High-level Panel on Threats, Challenges and Change, 'A More Secure World: Our Shared Responsibility', A/59/565, 2 December 2004, para. 204.

[37] Doyle, 'The Politics of Global Humanitarianism', p. 23.

[38] Ibid. pp.15, 26-27.

legitimacy'[39] alone is inadequate with the apt reminder that the 'Security Council practice of intervention in the 1990s, when anything that could muster the right votes passed' was a more permissive and less 'credible standard for intervention' than the one set by the R2P.[40] Despite arguing that the 'substantive license' for human protection in cases of atrocity crimes envisaged by the R2P was more permissive than Chapter VII of the Charter ('international' threats as traditionally defined) and that it 'still strained against the letter-of-the-law role that the Charter assigns the Security Council' with regards to 'defin[ing] an "international issue"', Doyle recognises the importance of substantive legitimacy.[41] Whilst he is not mistaken to identify the cause of civilian protection from the four crimes as a source of legitimacy for forcible intervention, according to the WSOD the latter is not sufficient grounds to legitimise coercive action in and of itself.

As Gallagher puts it, if 'proof of one of the four crimes being carried out is enough to warrant a pillar three response, then the requirement of a "manifest failing" would not be needed'.[42] He goes on to clarify that 'from a legal perspective, the killing of just one person or a small group of people, for example a group of hostages, could constitute genocide, but this would not mean that the host state had "manifestly failed" in its R2P'.[43] Gallagher's logic reinforces what is already clear from the Outcome Document, namely that the mere presence of atrocities is not enough to legitimise international action, if a specific threshold – the 'manifest failing' of the host state – is not met. In other words, it is the 'manifest failing' threshold requirement that draws the line between legitimate and illegitimate coercive action against a functioning sovereign state.

This means that it represents a substantive normative constraint on unwarranted coercive interference (on the part of the international community) in the domestic affairs of a state where mass atrocities are purportedly taking place. As such, it protects the equities of the host state. At the same time, it is also a trigger for the international community's responsibility to take 'timely and decisive' collective action under Chapter VII of the UN Charter to put an end to the violence against innocent populations. Therefore, the

[39] As described in: Stephen P. Marks & Nicholas Cooper, 'The Responsibility to Protect: Watershed or Old Wine in a New Bottle?', 2 (1) *Jindal Global Law Review* (2010): 86-130, p. 118.

[40] Ibid. p. 26.

[41] Ibid. pp. 23, 26.

[42] Adrian Gallagher, 'Syria and the indicators of a "manifest failing"', 18 (1) *The International Journal of Human Rights* (2014): 1-19, p. 2.

[43] Ibid.

'manifest failing' threshold itself has a dual function: it serves both as a license to and a leash against coercive action.

However, these functions are compromised by its lack of content. In its current nebulous form, the 'manifest failing' threshold neither places effective constraints on coercive action under the guise of humanitarianism, nor does it facilitate collective action in response to the plight of populations in need of protection. With regards to the former, the lack of clarity allows states to invoke the 'manifest failing' threshold and the R2P concept in situations where there is little evidence that a national government is not fulfilling its protection responsibilities and without having to provide robust justifications, conduct a satisfactory inquiry into a state's willingness and ability to address the atrocity crimes taking place or set reasonable expectations for them to do so. This allows states to seek cover for ulterior motives under the banner of the R2P. As Deeks remarks with regards to the 'unable or unwilling' threshold for extraterritorial self-defence against non-state actors, 'where the test is not clear, an [intervening] state's claim that a territorial state is unwilling or unable to act [against a threatening non-state actor on its territory] is easy to make, relatively hard to disprove, and at least superficially useful in concealing an incursion based on other motivations'.[44] In this sense, in 'the absence of a better-defined test', states' actions are not tempered by any standards they need consider to justify their decisions to take action against another state.[45]

Russia's attempt to justify its August 2008 military intervention in Georgia by invoking the R2P is a stark example of the potential for abusing a rule in the absence of clarity. Gareth Evans argued that 'even if the R2P norm were applicable' in this situation, Russia had failed to make a convincing case that the threat Georgia's attacks on South Ossetia's main city posed to population there 'was of a nature and scale as to make necessary or legitimate the use by it of military force', because the 'manifest failing' requirement had not been satisfied.[46] Specifically, he claimed that while Georgia's actions could be construed as an unwarranted excessive response to the provocations against Georgian villages it cited, 'the available evidence [was] not of the weight or clarity that is needed to justify a conclusion that it was "manifestly failing" to protect its population from these

[44] Deeks, '"Unwilling or Unable"', p. 507.

[45] Ibid. pp. 503, 512.

[46] Gareth Evans, 'Russia, Georgia and the Responsibility to Protect', 1 (2) *Amsterdam Law Forum* (2009): 25-28, p. 26.

atrocity crimes, in a way that would prima facie justify the use of coercive military action by others in response'.[47] On the one hand, the lack of clarity surrounding the 'manifest failing' threshold has made it too easy for Russia to completely disregard it. On the other hand, it undermines claims, such as those made by Evans, that the threat to civilian populations is not serious or imminent enough to suggest that the threshold for coercive action has been crossed. In light of this example, it is not surprising that within R2P circles, some have voiced their concerns over the potential for (Western) states to take advantage of the ambiguity surrounding the 'manifest failing' test and employ it in a way that serves their own interests, which can lead to unwarranted coercive action.[48] A great example of this is in a statement by the Vice President the Chinese foreign ministry's think tank, Ruan Zongze:

> 'There are so far no objective criterions to judge if "a country is unwilling or unable to execute responsibility to protect" and therefore it is likely that it would become another excuse for some countries to impose armed intervention in the internal affairs of other countries'.[49]

On the other side of the coin, the uncertainty about what constitutes a 'manifest failing' can lead to omissions to consider the need for such action in the face of atrocities for two reasons. First, it makes it more difficult for both state and non-state actors to justify the need to take collective action by invoking the 'manifest failing' threshold and the R2P in the hope to encourage the UNSC and regional actors to consider responding to a perceived atrocity crisis. An apt illustration of this is the failing of ASEAN and the UN to adequately respond to the continuous threat of atrocities in Myanmar and to heed international calls voicing the need to address the atrocities being committed against the Rohingya minority.[50] In a report published in November 2019, the Asia-Pacific Centre for the

---

[47] Ibid.

[48] See the discussion in Adrian Gallagher, 'What constitutes a "Manifest Failing"? Ambiguous and inconsistent terminology and the Responsibility to Protect', 28 (4) *International Relations* (2014): 428-444, pp. 429, 435-437, 439.

[49] Ruan Zongze, 'Responsible Protection: Building a Safer World', China Institute of International Studies (15 June 2012). [Accessed 20 June 2012] Available at: http://www.ciis.org.cn/eng- lish/2012-06/15/content_5090912.htm. Quoted in Gallagher, 'What constitutes a "Manifest Failing"?', p. 435.

[50] For a discussion as to why ASEAN's strategy of appeasement falls short of what is needed see: Kilian Spandler, 'Lessons from ASEAN's Rakhine response', 5 February 2020 [Accessed 26 October 2020] Available at: https://www.eastasiaforum.org/2020/02/05/lessons-from-aseans-rakhine-response/. For a critique of UN's '"systemic failure" in the… handling of the crisis' see: Al Jazeera, 'Why the UN failed to save the Rohingya' [Accessed 26 October 2020]Available at: https://www.aljazeera.com/news/2019/6/28/why-the-un-failed-to-save-the-rohingya; Hannah Ellis-Petersen and Emanuel Stoakes, 'UN report condemns its conduct in Myanmar as systemic failure', 17

Responsibility to Protect (APR2P) 'conclude[d] that two years after the military's 'clearance operations' against Rohingya Muslims in Rakhine State, Myanmar still harbours the conditions conducive to the commission or incitement of atrocity crimes'.[51] The enduring threat of atrocities in Myanmar has prompted the GCR2P to repeatedly suggest that: 'The *government of Myanmar has manifestly failed* to uphold its responsibility to protect the Rohingya and other minority populations, and bears responsibility for the commission of war crimes, crimes against humanity and genocide' (emphasis added), with the wording recently changed to 'Myanmar's military has manifestly failed', following the recent military coup.[52] Although, the GCR2P has repeatedly recommended that 'the UNSC should… refer the situation in Myanmar to the ICC and impose an arms embargo' amongst other measures,[53] '[t]he only formal response by the UN Security Council… to the genocide against the Rohingya was the adoption of a Presidential Statement on 6 November 2017 that stressed the "primary responsibility of the Myanmar government to protect its population."'[54] Even though some ASEAN countries, such as Malaysia and Indonesia have voiced their concerns over the treatment of the Rohingya in the past, with the Malaysian government 'condemn[ing] the Myanmar government for the massacre of Rohingya in Rakhine', other members of the block were intent on remaining silent over the matter of its responsibility for the ethnic cleansing campaign carried out by its military.[55]

June 2019 [Accessed 26 October 2020] Available at: https://www.theguardian.com/world/2019/jun/17/un-report-myanmar-rohingya-systemic-failure; Al Jazeera, 'UN fails to take action on order against Myanmar on Rohingya', 5 February 2020 [Accessed 26 October 2020] Available at: https://www.aljazeera.com/news/2020/2/5/un-fails-to-take-action-on-order-against-myanmar-on-rohingya

[51] APR2P, 'November 2019: Risk Assessment Series Volume 9: Myanmar', November 2019. [Accessed 21 October 2020]. Available at: https://r2pasiapacific.org/files/4243/Risk_Assessment_myanmar_vol9_november20192.pdf

[52] GCR2P, 'Myanmar (Burma): Populations at risk', 15 September 2020. [Accessed 21 October 2020]. Available at: https://www.globalr2p.org/countries/myanmar-burma/; GCR2P, 'R2P Monitor', Issue 55.

[53] Ibid.

[54] Ibid.

[55] CNA, 'Mahathir at UN condemns Myanmar's treatment of Rohingya', 29 September 2018. [Accessed 31 January 2021]. Available at: https://www.channelnewsasia.com/news/asia/mahathir-at-un-condemns-myanmar-s-treatment-of-rohingya-10770568; Reuters, 'Malaysia's dissent on Myanmar statement reveals cracks in ASEAN façade', 25 September 2017. [Accessed 31 January 2021]. Available at: https://www.reuters.com/article/us-myanmar-rohingya-malaysia/malaysias-dissent-on-myanmar-statement-reveals-cracks-in-asean-facade-idUSKCN1C0124; The Jakarta Post, 'Indonesia condemns Rakhine clashes', 29 August 2017. [Accessed 31 January 2021]. Available at: https://www.thejakartapost.com/news/2017/08/29/indonesia-condemns-rakhine-clashes.html; Human Rights Watch, 'ASEAN: Don't Whitewash Atrocities Against Rohingya', 19 June 2019. [Accessed 31 January 2021]. Available at: https://www.hrw.org/news/2019/06/19/asean-dont-whitewash-atrocities-against-rohingya?search=asean+divided+over++Rohingya

Second, it makes it very easy for the intended audience to object to and reject claims that a host state is failing in its responsibility. A highly indeterminate threshold means that Council members may easily disregard such calls as a legitimate basis for action, or else claim that threshold has not been met, or that the host state's failing cannot be determined to a sufficient degree of certainty. As a result, the ambiguity as to what the test requires may result in omissions to acknowledge the validity of claims that a state is 'manifestly failing' and in turn to omissions to protect, ultimately increasing the risk of vulnerable populations being systematically under-protected against the threat of atrocities. This argument is supported by Marks and Copper who find that in the absence of guidelines for making 'manifest failing' assessments, the world leaders' 'commitment at the World Summit allows the permanent members to continue to prioritize their self-interest behind the fig leaf of the World Summit Outcome, and remain unaccountable to the intended beneficiaries of the responsibility to protect'.[56] Hence, the absence of guidance as to what the 'manifest failing' test entails allows states not only to force an intervention but to wash their hands of responsibility for the protection of innocent populations, as their best interests dictate. This comes to show that addressing the substantive indeterminacy of the 'manifest failing' threshold is of critical import to consistent and adequate response to the threat of atrocity crimes on the part of the international community.

The substantive indeterminacy and lack of understanding surrounding the 'manifest failing' notion also makes it more challenging to establish the need for coercive action. It may be true that 'coercive intervention to protect populations is a low-probability, high-risk phenomenon'.[57] Notwithstanding, the need to contemplate such action arises far more often than it actually takes place. The last decade saw an atrocity crisis slip through all existing early warning mechanisms, which was deemed to merit a coercive action response, namely the 2011 Libyan intervention. Even though international action in Libya was not anticipated, it is possible that an occasion to seriously deliberate a strong international response to atrocities arises in the future sooner than we expect, even though such debates do not seem so prominent in the UN Security Council. We have also seen cases which have arguably warranted such response on account of the 'manifest failing' of the host state, but nevertheless failed to command the required consensus amongst UNSC members on timely and decisive action, such as that of the Syrian government in

[56] Marks and Cooper, 'The Responsibility to Protect', p. 120.
[57] Bellamy, 'The First Response', p. 11.

the context of the conflict within the country, which was triggered by the popular uprising against the regime of President Bashar al-Assad in mid-March 2011.[58] Relatedly, even when the need for strong collective action arises, it may be precluded owing to a range of geopolitical factors and prudential consideration (e.g. when the adoption of forcible measures cannot achieve civilian protection at an acceptable cost).[59] However, such considerations are separate from and subsequent to the matter of establishing the need for such action in the first place. Consequently, relevant stakeholders ought to be able to identify cases is which the host state may be 'manifestly failing' in its responsibility to protect, so that the international community stands a chance of being prepared to respond collectively, in an effective manner, to a variety of situations which may fall within the remit of the R2P.

Addressing objections to the study of the 'manifest failing' threshold

This research agenda and its premise will certainly not resonate with all. To begin with, it will be rejected by R2P critics of a pacifist disposition. Coining the notion of a 'responsibility to peace', Mary Ellen O'Connell argues for a return to 'building the norm of peace', because the 'flourishing of human rights' will be better supported 'through the promotion of peace and the rule of law in the world', as opposed to advocating military intervention under the R2P.[60] Similarly, Robin Dunford and Michael Neu argue that in order to move towards the goal of a 'less violent world' that is 'less conducive' to atrocities, we need to 'step aside from the R2P' in favour of supporting non-militarist alternatives.[61] Their stance is premised on the argument that 'far from limiting, preventing and appropriately responding to mass atrocities, the R2P gives a veneer of legitimacy to military interventions performed by states that are already engaged in damaging forms of

[58] For a detailed discussion of Syria's 'manifest failing' see Gallagher, 'Syria and the indicators of a "manifest failing"'.

[59] As the International Commission on Intervention and State Sovereignty (ICISS) specifies, 'military action can only be justified if it stands a reasonable chance of success, that is, halting or averting the atrocities or suffering that triggered the intervention in the first place'.: International Commission on Intervention and State Sovereignty, 'The Responsibility to Protect: Report of the International Commission on Intervention and State Sovereignty' (Ottawa, ON, Canada: International Development Research Centre, 2001), para. 4.41. Conversely, it is not justifiable 'if actual protection cannot be achieved, or if the consequences of embarking upon the intervention are likely to be worse than if there is no action at all [, for instance if such action] triggers a larger [regional] conflict…, involving major military powers.: Ibid.

[60] Mary Ellen O'Connell, 'Responsibility to Peace', in Philip Cunliffe (ed.) *Critical Perspectives on the Responsibility to Protect* (London: Routledge, 2011): 71-83, pp. 71, 80.

[61] Robin Dunford and Michael Neu, 'The Responsibility to Protect in a world of already existing intervention', 25 (4) *European Journal of International Relations* (2019): 1080–1102, p. 1097.

political and economic intervention', i.e. 'it serves to legitimise a moralistic form of militarism'.[62] As much as any R2P proponent would be delighted if the goal of protecting vulnerable populations from the threat of atrocities could be attained through peaceful means only, the existence of an alternative to take action in extreme scenarios, such as those where national governments are massacring their own populations, is not only necessary to prevent and halt atrocities, but also acts as a deterrent for regimes who moderate their behaviour toward their populations out of fear for international retribution. For these reasons, it is worth attempting to capitalise on the agreement reached at the 2005 World Summit and endeavour to reform the R2P to address its shortcomings, such as the negative implications stemming from the lack of clarity surrounding the 'manifest failing' threshold.

Other critics highlight the numerous impediments to the effective implementation of the R2P, including the difficulty of generating the political will to take collective action, logistical obstacles, and 'moral hazard', namely 'the unintended consequence of encouraging rebellion by members of vulnerable sub-state groups, prompting states to retaliate with genocidal violence before intervention can stop it'.[63] While these are valid concerns, they fall outside the scope of this thesis, which is concerned with addressing the problems with the conceptualisation and consistent interpretation of the 'manifest failing' threshold, rather than the operationalisation of R2P's collective action aspect.

Still others, who are not opposed to humanitarian intervention in principle, are critical of the R2P's failure to fulfil its aspiration to prevent mass atrocity crimes, owing to the fact that it does not make a difference to state behaviour and international relations. For instance, Aidan Hehir argues that the R2P 'does not alter the existing structure of international law regarding sovereign responsibility, the authority to use force or the thresholds for intervention, and is ultimately based on a highly idealistic belief in the capacity of moral pressure to alter the disposition of the world's states'.[64] Relatedly, Gallagher cautions that some may reject the 'manifest failing' research agenda because it has no bearing on the decision-making process. Namely, 'that it does not matter what phrase is used, the decision to react is a political choice based on things other than a

[62] Ibid. For a thorough critique of the R2P, see: Robin Dunford and Michael Neu, *Just War and the Responsibility to Protect: A Critique* (London: Zed Books, 2019).
[63] Alan J. Kuperman, 'Rethinking the Responsibility to Protect', 10 (1) *The Whitehead Journal of Diplomacy and International Relations* (Winter/Spring 2009): 33–43, p. 40.
[64] Aidan Hehir, 'R2P and international law', in Philip Cunliffe (ed.) *Critical Perspectives on the Responsibility to Protect* (London: Routledge, 2011): 84-99, p. 97.

'manifest failing', such as the national interest, sovereignty and the complexities of intervention'.[65] Following this line of argumentation, Justin Morris makes a forceful proposition to divorce R2P's coercive dimension from other aspects of the concept:

> 'R2P's international standing can be best preserved through the excision of its most coercive elements; R2P should be reconstituted as a standard of acceptable sovereign behaviour and a mechanism geared towards the provisions of international guidance and support, while decisions over coercive military intervention, inevitably infused with considerations of strategic interest, should be made outside the R2P framework.'[66]

My response to such critiques draws on Gallagher's defence of the 'manifest failing' research agenda. I have no intention to understate the role of national interest and geo-politics in decisions over the use of force (especially on humanitarian grounds),[67] nor claim that 'UN Security Council talks break down because the permanent five Member States cannot agree on whether a "manifest failing" is taking place or that it creates a new limit on the UN Charter'.[68] However, recognizing that there are other important factors that inform decisions to take coercive action in mass atrocity situations, by no means leads to the conclusion that it does not matter whether a state is deemed to be 'manifestly failing' in the responsibility to protect its citizens or not. Whereas considerations of national interest constitute an invariable part of the discourse and decisions surrounding R2P crises, so does the debate over threshold, as exemplified in analysis of international discourses on the Libyan crisis in Chapter IV of this thesis. According to Gallagher, the debate over threshold can also be seen in 'the narrative that surrounds any R2P crisis', even in the absence of the 'manifest failing' terminology:

> '[I]t is important to bear in mind that within the R2P framework (as set out in the WSOD and many UN Reports since), the 'manifest failing' requirement represents threshold. Whether this is framed in terms of illegitimacy,

---

[65] Gallagher, 'What constitutes a "Manifest Failing"?', p. 436. Adrian Gallagher, 'A Clash of Responsibilities: Engaging with Realist Critiques of the R2P', 4 (3) *Global Responsibility to Protect* (2012): 334–57.

[66] Justin Morris, 'Libya and Syria: R2P and the spectre of the swinging pendulum', 89 (5) *International Affairs* (2013): 1265-1283, p. 1266.

[67] Adrian Gallagher, 'Syria and the indicators of a "manifest failing"', p.5; Gallagher, 'What constitutes a "Manifest Failing"?', p. 436.

[68] Gallagher, 'What constitutes a "Manifest Failing"?', p. 430.

> irresponsibility, an escalation in violence, 'unable or unwilling' or 'manifest failing', the underlying logic is that of threshold: a line has been crossed in that what was tolerated yesterday cannot be today and action has to be taken'.[69]

Since 'it is clear that the UN Security Council appeals to some sort of threshold when judging whether [a host state] is failing to fulfil its R2P' and that the 'manifest failing' threshold is the most widely accepted representation of this threshold (as noted earlier), claims that the substance of this threshold is unimportant are misguided.[70] As Gallagher reminds, because UN member states representatives at the 2005 World Summit agreed to place the 'manifest failing' test 'at the heart of any pillar three decision-making process', '[n]ormatively [states] should agree that a "manifest failing" is taking place' prior to pursuing coercive action.[71] In light of the prospect that 'the lives of potentially millions of people may depend on' the decision to trigger the coercive aspect of the R2P, the significance of clarifying the substance cannot be denied.[72]

Having considered the negative implications of the 'manifest failing' test's substantive indeterminacy and highlighted the importance of addressing this problem, it is puzzling as to why its meaning is rarely discussed by scholars and policy-makers alike. Perhaps, this lack of attention can be attributed to the seemingly straightforward nature of the inquiry as to what the 'manifest failing' test entails. As Gallagher remarks, it could be argued 'that the phrase is so transparent there is no need for clarity; after all, the word manifest means "evident to the eye, mind, or judgement; obvious"'.[73] However, such inquiries are unlikely to be so straightforward in practice.

For instance, Gallagher's analysis of the Syrian crisis convincingly illustrates the Syrian government's 'manifestly failing' to protect its populations.[74] Nonetheless, China and Russia – the Permanent Two (P2) members of the UNSC – notably blocked three UN Security Council draft resolutions on Syria, prompted by the escalation of atrocities

[69] Ibid. p. 436. See also: Gallagher, 'Syria and the indicators of a "manifest failing"', p. 5.

[70] Gallagher, 'What constitutes a "Manifest Failing"?', p.430. Gallagher, 'Syria and the indicators of a "manifest failing", p. 5.

[71] Gallagher, 'What constitutes a "Manifest Failing"?', p. 430.

[72] Gallagher, 'Syria and the indicators of a "manifest failing"', p. 13.

[73] Gallagher, 'What constitutes a "Manifest Failing"?', p. 430.

[74] Gallagher, 'Syria and the indicators of a "manifest failing"'.

against civilian populations, over the course of ten months (October 2011 to July 2012).[75] China and Russia consistently opposed UN draft resolutions for being prejudiced against the Syrian state, thus actively refusing to interpret the situation as one in which a host state is failing in its responsibility to protect its populations. A prime example of this is China's justification for blocking the third draft resolution on Syria, namely that 'the draft resolution is seriously flawed, and its unbalanced content seeks to put pressure on only one party'.[76] Likewise, China was critical of the first draft resolution on Syria because 'it focuses solely on exerting pressure on Syria, even threatening to impose sanctions'.[77] At the same meeting, Russia expressed a similar sentiment, arguing that 'the continuation of [the Syrian] tragedy [could not] be blamed only on the harsh actions of the authorities [since non-state armed groups were] killing and perpetrating atrocities against people who comply with the law enforcement authorities'.[78] Russia maintained its position at the vote on subsequent draft resolutions, suggesting that the draft texts 'did not adequately reflect the true state of affairs in Syria and sent a biased signal to the Syrian sides', with threats and accusations levelled exclusively at the Government of Syria.[79]

The views of the P2 stand out in stark contrast with Gallagher's argument that while Syria is 'undoubtedly a multifaceted crisis', in which both sides were responsible for international crimes, the war crimes committed by 'armed groups did not…reach the intensity and scale of those committed by Government forces and affiliated militia'.[80] According to him, this and the fact that only the Syrian government committed both war crimes and crimes against humanity evidences 'a profound imbalance [between the two sides of the conflict] that should not be misconceptualised'.[81] While the failing of the Syrian state was 'obvious' to most, the P2 resisted framing the Syrian crisis as a 'manifest failing' of the host state and promulgated an alternative understanding of the situation as a civil war. This shows how the absence of clarity as to what constitutes a 'manifest failing' leaves loopholes for justifying inaction in the face of blatant failures on the part of national governments to protect their populations. A better understanding of what

[75] UNSC, S/2011/612, 4 October 2011; UNSC, S/2012/77, 4 February 2012; UNSC, S/2012/538, 19 July 2012.

[76] UNSC, S/PV.6810, 19 July 2012, p. 13

[77] UNSC, S/PV.6627, 4 October 2011, p. 5.

[78] Ibid. p. 4.

[79] UNSC, S/PV.6711, 4 February 2012, p. 9. S/PV.6810 (2012), p. 9.

[80] Gallagher, 'Syria and the indicators of a "manifest failing"', p. 7.

[81] Ibid.

determines a 'manifest failing' can help to address the difficult questions surrounding the actions of national governments in contested situations, with a view to ascertaining whether they are fulfilling their primary protection responsibility towards their populations and whether a decisive international response is warranted or not.

What is more, inquiries into a state's 'manifest failing' tend to be complicated, because host states neither tend to profess their unsavoury intentions, as will be discussed at length in Chapter I, nor openly refuse to comply with the demands of the international community.[82] The resulting uncertainty as to the genuine intentions and actions of states raises some difficult questions that have a bearing on how states do or should interpret the threshold for coercive action. For instance, what if the host state requires more time to deal with the threat only to see its murderous campaign through? Likewise, what if the host state promises to put an end to all violence against innocent populations but does nothing or not enough to supress the threat? As things stand, we are not only a long way away from possessing the forensic tools to obtain fully-fledged incontrovertible evidence of a state's involvement in mass atrocities and the precise nature of these crimes, but we are also dependant on the consent and cooperation of the host state to acquire such evidence. For these reasons, it will rarely be clear, at least at the outset, whether the host state is 'manifestly failing' to protect its populations. Once again, this underlines the need to address the confusion surrounding the threshold for coercive action in a way that facilitates identifying a 'manifestly failing' state in practice.

RESEARCH OBJECTIVES

As mentioned at the start, in order to address the negative implications arising from the substantive indeterminacy of the 'manifest failing' threshold, this thesis will develop a normative framework which can aid relevant stakeholders in determining whether a state is 'manifestly failing' in its individual 'responsibility to protect' or not. A similar proposition for establishing guidelines was made in the 2004 HLP report, which emphasised that 'it is critical that [the UNSC's] most important and influential decisions, those with large-scale life-and-death impact, be better made, better substantiated and better communicated'.[83] To enhance the legitimacy of such decisions the HLP recommended the adoption of a set of guidelines, namely 'five criteria of legitimacy' –

[82] On the problems of assessing government intentions see: Gallagher, 'Syria and the indicators of a "manifest failing"', p. 5.

[83] A/59/565 (2004), para. 205.

'seriousness of threat, proper purpose, last resort, proportional means and balance of consequences'.[84] According to the Panel, the 'Security Council (and anyone else involved in these decisions)' should 'always address [these criteria] — whatever other considerations [decision-makers] may take into account' when contemplating the authorisation or implementation of military force.[85] Although this thesis does not go as far as to insist that guidelines for 'manifest failing' determinations be formally endorsed by the UN Security Council, it remains similar to the HLP in the important sense that it proposes a set of guidelines 'going directly not to whether force can legally be used but whether, as a matter of good conscience and good sense, it should be'.[86]

Undoubtedly, some would question the utility of such an informal framework. After all, '[e]ven the most successful legal test in the use of force area cannot excise considerations of politics and diplomacy from statements made by members of the Security Council'.[87] However, the framework to be devised in this thesis proposes to serve no such ambitious (if not unattainable) objective. Instead, what the abovementioned set of informal guidelines for 'manifest failing' assessments can provide is a common ground for debates over the appropriateness of collective action, which is currently lacking.

The value and practical utility of such frameworks is widely acknowledged within the broader scholarship on the use of force. For instance, by building on the logic of Just War theory and Wheeler's suggestion for creating 'common reference within which argumentation can take place' as to what constitutes a legitimate humanitarian intervention, Gallagher rationalises establishing a general framework for judging a host state's 'manifest failing'.[88] Deeks also proposes a normative framework to guide assessments as to whether a state is 'unable or unwilling' to supress the threat of attacks perpetrated by a non-state actor on its territory against another state, although it is worth noting that she is concerned with the legality as well as the legitimacy of the use of force. Specifically, she argues that 'a clearer and more detailed "unwilling or unable" test would [not only] provide a common vocabulary for all states to use in discussing and evaluating a …state's use of force' in self-defence, but also contribute to 'more coherent analysis' concerning the legality of use of force in self-defence by states and international bodies

[84] Ibid. p. 13.
[85] Ibid. p. 53.
[86] Ibid. para. 205.
[87] Deeks, '"Unwilling or Unable"', p. 514.
[88] Ibid.; Wheeler, *Saving Strangers,* p. 33.

who are not involved in the decision to intervene.[89] In addition, Deeks maintains that 'giving players a common script against which to measure the facts' will force states to practise 'more measured decision-making' and potentially dissuade them from using force, because the prospect of having to justify their actions against a specific test will incentivise them to consider it carefully.[90] Likewise, a rudimentary touchstone against which to analyse the information relevant to 'manifest failing' determinations will encourage decision-makers to justify their judgements concerning R2P-spirited international action with care and limit the reasonable arguments they can make to support or undermine a particular course of action (see discussion in Chapter II). Thus, as Rosenberg and Strauss suggest a common standard of assessment of R2P cases will have a legitimising effect (if applied successfully), because 'while inevitably open to interpretation by all parties, [it] will at the very least begin to require parties to explain their reasoning from a common reference point'.[91] Relatedly, according to Deeks, if a specific threshold standard is used by some actors to make a case for or against the use of force, then the rest will also be 'forced' to use it to frame their arguments.[92]

As the above discussion suggests, the potential benefits of addressing the lack of clarify surrounding the threshold for coercive action through establishing a common reference framework for making 'manifest failing' determinations are numerous. Most importantly, however, they feed directly into enhancing the functions of this threshold requirement as licence for and a leash against coercive action. A clearer 'manifest failing' test will make it more challenging for powerful states to justify a forcible intervention in another state where there is no good cause for it. In this sense, it will help to delegitimise coercive interventions should arguments that the host state was 'manifestly failing' be found unconvincing. Equally, more clarity as to the parameters of the 'manifest failing' requirement would facilitate making the case that a state is not fulfilling its responsibility to protect its populations, in the presence of compelling evidence of the latter. In such

[89] Deeks, '"Unwilling or Unable"', pp. 511, 514.

[90] Ibid. pp. 512, 514.

[91] Sheri Rosenberg and Ekkehard Strauss, 'A Common Approach to the Application of the Responsibility to Protect', in Daniel Fiott, Robert Zuber and Joachim Koops (ed), *Operationalizing the Responsibility to Protect: A Contribution to the Third Pillar Approach* (Brussels: The Madariaga – College of Europe Foundation, Global Action to Prevent War, the Global Governance Institute and the International Coalition for the Responsibility to Protect, 2012): 55-72, p. 56.

[92] Deeks, '"Unwilling or Unable"', p. 514.

circumstances, decision-makers will face increased international scrutiny for failures to protect.

A framework for the interpretation (not the application) of the R2P

In order to avoid misunderstanding as to the precise objective for establishing such a framework, it is important to note that, 'consistency in the application and implementation of the RtoP and consistency in the interpretation of the concept' are two separate matters.[93] By developing a framework to guide assessments of what constitutes a 'manifest failing' this thesis seeks to make a specific contribution to the question of interpretation, whilst remaining consistent with and complementing the UN approach to implementation. In so doing, it shares Gallagher's view that 'our interpretation of the [R2P] concept should be consistent', if it is to be sustained.[94] When it comes to the question of application, it is widely acknowledged that '[a]s each situation is different, it would be counterproductive to try to make the application of [R2P's principles] appear identical in all situations', because whether collective action takes place (especially when enforcement measures are concerned) depends on a range of other factors, including 'the circumstances on the ground and [an] informed judgment of the likely consequences'.[95] The 'case-by-case' approach to collective action specified in the 2005 Outcome Document ensures that international response to atrocity crises is always tailored to the circumstances of the incident under consideration. As Reike observes, the WSOD's stipulation that decisions regarding coercive action are to be made on a 'case-by-case basis' weaved 'an element of inconsistency' into the fabric of R2P's reactive dimension.[96] As things stand, even though the ambiguity surrounding the 'manifest failing' test may mean more flexibility, it also means facilitating 'great power manipulation' and 'the ability of states to distance themselves from their responsibilities'.[97] The risks associated with maintaining this flexibility are hard to justify, since relying on the 'substantive

[93] Alex J. Bellamy, 'Making RtoP a Living Reality: Reflections on the 2012 General Assembly Dialogue on Timely and Decisive Response', 5 (1) *Global Responsibility to Protect* (2013): 109–25, p. 119. Quoted in: Gallagher, 'Syria and the indicators of a "manifest failing"', p. 4.

[94] Gallagher, 'Syria and the indicators of a "manifest failing"', p. 4.

[95] Ban Ki-moon 'Responsibility to Protect: Timely and Decisive Response', Report of the Secretary-General, A/66/874-S/2012/578, 25 July 2012, p. 6; Welsh, 'Norm Contestation', p. 368.

[96] Ruben Reike, 'Libya and the Responsibility to Protect: Lessons for "Pillar Three"', in Daniel Fiott, Robert Zuber and Joachim Koops (ed.) *Operationalizing the Responsibility to Protect: A Contribution to the Third Pillar Approach* (Brussels: The Madariaga – College of Europe Foundation, Global Action to Prevent War, the Global Governance Institute and the International Coalition for the Responsibility to Protect, 2012): 73-79, p. 73.

[97] Ibid.

indeterminacy' of the 'manifest failing' test is hardly necessary to ensure sufficient wiggle room for decision-marking. After all, the WSOD already provides ample space of political manoeuvring, not least because of the 'case-by-case' approach to collective action, the absence of an explicitly specified international 'responsibility to protect' and the fact that this diluted commitment to 'timely and decisive' action is not legally binding.

A common framework consistent with the case-by-case approach to atrocity crises response

Nonetheless, pushback against a more specific threshold for coercive action often seems to arise from the perceived contradiction between establishing guidelines for determining whether a state is 'manifestly failing' in its responsibility or not and the assessment of each situation on a case-by-case basis.[98] However, the supposition that these goals are mutually exclusive is false. 'Guidelines', as the term suggests, are meant to *guide* assessments, not trump, replace, invalidate or compromise tailored context-specific treatment of individual cases. What is more, as Marks and Cooper note, '[w]hat a case-by-case approach does *not* specify is that the response must be taken in the absence of a set of basic standards, as asserted by the responsibility to protect'.[99] This view is backed by Rosenberg and Strauss who advocate that '[a] coherent, common standard of assessment that can be utilized on a case-by-case basis provides clear boundaries to discussions over when the RtoP applies'.[100] Gallagher provides further support for this by drawing on the similarity between his endeavour to develop indicators of 'manifest failing' and the principles for the justifiable use of force established in the Just War tradition. According to him, whilst recognising that decisions over the use of force are complex and cannot simply follow from a checklist of triggers, Just War theorists have developed a framework for 'identifying a just war'.[101] This gives weight to the idea that even though decisions over forcible action are heavily dependent on the context of individual cases, states still need a substantive reference as to what makes such action justifiable in order to determine and rationalise its legitimacy. Therefore, it is plain that

[98] Marks and Cooper, 'The Responsibility to Protect', p. 120.

[99] Ibid.

[100] Rosenberg and Strauss, 'A Common Approach', p. 55. Gallagher, 'What constitutes a "Manifest Failing"?', p. 439.

[101] Gallagher, 'What constitutes a "Manifest Failing"?', p. 439.

even on a case-by-case basis decision-makers need to refer to something to provide justifications that the threshold for coercive action has been crossed.[102]

Put simply, it important to understand that the framework to be developed in this thesis is not intended to 'produce agreed conclusions with push-button predictability', but to facilitate decision-making on the appropriateness of coercive action in response to mass atrocities.[103] This objective mirrors Gallagher's motivation for establishing threshold 'indicators [that] may help aid decision-makers in making their assessment of a "manifest failing"', namely to supplement 'the UN approach by helping them to fulfil their commitment to respond in a "timely and decisive" manner'.[104] In addition to providing guidance to those involved in decisions on coercive R2P action, developing a normative framework for 'manifest failing' assessments will provide a common reference point for arbitrating between competing claims made by various actor as to whether a state is 'manifestly failing' in its protection responsibility and judge the merits of collective action based on these claims.

CONTRIBUTION

The foremost original contribution made in this thesis is developing a fully-fledged framework that will aid relevant stakeholders in determining whether a host state is 'manifestly failing' to protect its populations or not, because none exists presently. To do so, it develops a research strategy, which builds on, but goes beyond, narrative case-study analysis, to identify key factors signalling that a host state is failing in its responsibility to protect. Similarly to Rosenberg and Strauss' standard for the implementation of R2P, the normative framework for identifying a 'manifest failing' to be developed in this thesis can 'be used by governments, bodies of regional and international organizations, and civil society, which are all called upon to make assessments as to the risks of mass atrocities occurring in a particular situation'.[105] Specifically, by offering guidelines as to the principled assessment of a host state' 'manifest failing' this framework can aid decision-makers' assessments of their conduct during atrocity crisis and in doing so to boost their 'preparedness' to respond in a 'timely and decisive manner'. In addition, by providing a common reference 'against which political (in)action can be judged' this framework can

[102] Ibid. p. 437.

[103] Rosenberg and Strauss, 'A Common Approach', p. 53.

[104] Gallagher, 'What constitutes a "Manifest Failing"?', pp. 437-438.

[105] Rosenberg and Strauss, 'A Common Approach', p. 56.

'help those outside government to hold decision-makers to account', thus making it more difficult for the international community to shun their commitments under the R2P.[106] Eventually, it may also provide an incentive for host states to uphold their responsibilities towards their populations and readily address the threat of atrocities should it arise, so as to avoid the scrutiny of the international community and prevent incursions on their sovereignty.

In addition, this thesis makes several theoretical contributions. As noted above, the scholarship on mass violence and the R2P has barely scratched the surface of what the pursuit of a research agenda addressing the 'substantive indeterminacy' of the 'manifest failing' threshold might cover and offer. Hence, this research fills a lacuna the scholarly literature on R2P and mass atrocities by contributing to the better understanding of the terminological and substantive developments concerning the international community's collective action responsibility. However, although this thesis sets out to explore the threshold for coercive action for the purpose of protecting populations subjected to mass violence through the lens of the R2P framework, it is not just about the R2P concept and its normative development. As the opening paragraph states, it is, first and foremost, a study of the threshold for coercive international action in the context of mass atrocities. In this sense, the focus on the R2P is instrumental in fleshing out the contemporary debate over this threshold, because it provides the conceptual framework and language employed in these discussions and is the closest thing there is to an internationally agreed upon framework for responding to atrocity crimes. For the same reasons, the terminologies used to represent this threshold, be it 'manifestly failing', 'unable or unwilling' or anything in-between, are not the primary object of study in and of themselves. They matter because of what they stand for – the issue of threshold – and are used as vehicles to gain an insight into the understating of the threshold of coercive action of those who employ them. Understood in this broader sense, this endeavour to illuminate the substance of the 'manifestly failing' threshold requirement also feeds into a key debate over 'which humanitarian crises justify international moral action and which do not?',[107] with notable contributions from Wheeler, Bellamy and Pape.[108] By the same token, it is also adds to

[106] Gallagher, 'Syria and the indicators of a "manifest failing"', p. 4.

[107] Ibid. p. 2.

[108] Wheeler, *Saving Strangers*. Alex J. Bellamy, 'Mass Atrocities and Armed Conflict: Links, Distinctions, and Implications for the Responsibility to Prevent' (Muscatine, IA: The Stanley Foundation, February 2011) pp. 1-20. Robert A. Pape, 'When Duty Calls: A Pragmatic Standard for Humanitarian Intervention', 37 (1) *International Security* (2012): 41-80.

the broader debate on the legitimate use of force and coercive action short of force in international relations.

Besides theoretical contributions, this thesis also makes two empirical contributions to the study of the threshold for international action. In Chapter IV, it examines one of the most high-profile cases discussed in the context of the R2P framework, namely the 2011 atrocity crisis in Libya, with a focus on the threshold for coercive action. Despite the prominence of the Libyan case, attentive examination of Libya's 'manifest failing' has remained beyond the scope of scholarly attention. The only exception is an 800-word section of a book chapter by Ingvild Bode, which neither focuses on clarifying the substance of the 'manifest failing' test, nor does it aim to identify any specific factors relevant to the threshold for coercive action.[109] Hence, my in-depth analysis of the Libyan crisis specifically dedicated to learning more about the substantive content of the 'manifest failing' test, as understood by UN member states, is an innovative contribution to the study of this otherwise widely discussed case. The second empirical contribution ensues from the application of the 'manifest failing' framework developed in this thesis in the context of the 'war on drugs' pursued by the Government of the Philippines since 1 July 2016, according to the International Criminal Court (ICC).[110] The Philippines was identified as a 'manifestly failing' state by Gallagher et al. in a 2019 publication and as 'remaining at a very high-risk for atrocities' by the APR2P in February 2020.[111] In Chapter V of this thesis, the 'manifest failing' of the Philippine government will be re-appraised in the context of current events in the country, enhanced international scrutiny and the Office of the United Nations High Commissioner for Human Right's (OHCHR) written report on 'the situation of human rights in the Philippines' presented to the Human Rights Council (HRC) in June 2020.[112]

---

[109] Ingvild Bode, '"Manifestly Failing" and "Unwilling or Unable" as Intervention Formulas: A Critical Assessment', in Aiden Warren and Damian Grenfell (eds.) *Rethinking Humanitarian Intervention in the 21st Century* (Edinburgh: Edinburgh University Press, 2017): 164-191.

[110] International Criminal Court website, 'Preliminary examination: Republic of the Philippines', (no date). [Accessed 14 January 2021]. Available at: https://www.icc-cpi.int/philippines

[111] Adrian Gallagher, Euan Raffle and Zain Maulana, 'Failing to fulfil the responsibility to protect: the war on drugs as crimes against humanity in the Philippines', 33 (2) *The Pacific Review* (2019): 247-277; APR2P, 'The Philippines: Asia Pacific Regional Outlook February 2021', 20 February 2021. [Accessed 24 March 2021]. Available at: https://r2pasiapacific.org/files/6262/APO%20The%20Philippines%20Feb%202021%20PDF.pdf

[112] United Nations General Assembly, 'Promotion and protection of human rights in the Philippines', Resolution adopted by the Human Rights Council on 11 July 2019, A/HRC/RES/41/2, 17 July 2019.

CHAPTER OUTLINE

This thesis comprises five chapters, plus the present introductory chapter and a Conclusion. The first chapter focuses on the literature on mass atrocity crimes and examines existing contributions to clarifying the substance of the 'manifest failing' threshold, with a view to revisiting and synthesising key determinants and indicators that others have identified as relevant to determining whether a host state is 'manifestly failing' in its protection responsibilities, as well as highlights lacunae in existing propositions for clarifying the 'manifest failing' test that the framework proposed in this thesis ought to address. In addition, this chapter draws on a wide range of sources, including the broader literature on mass atrocity crimes and the definitions of the four crimes, in order to supplement previous attempts to pinpoint the substantive aspects of the 'manifest failing' test and arrive at a continuum of determinants and associated indicators that will inform the normative framework for 'manifest failing' determinations presented in Chapter V. In so doing, this chapter not only positions the present endeavour within existing scholarship on mass atrocity crimes and 'manifest failing', but takes the first step towards specifying a framework that can assist relevant stakeholders in judging whether national governments are 'manifestly failing' to protect their populations.

Chapter II presents the research design underpinning this thesis. First, it advocates the rationale behind the adoption of a triangulation research strategy, which allows the integration of insights from multiple perspectives to aid the development of a comprehensive framework for the assessment of mass atrocity situations, with a view to determining whether a national government is 'manifestly failing' in its 'responsibility to protect'. Specifically, it involves triangulating 1) the determinants and indicators of 'manifest failing' identified from the literature on R2P and mass atrocities in Chapter I; 2) newly formulated 'manifest failing' determinants, premised on the study of an analogous threshold requirement – the 'unable or unwilling' test in the context of self-defence against non-state actors in Chapter III, and 3) determinants and indicators of 'manifest failing' derived from the aspects of the situation that states underscored in their justifications for pursing collective action in Libya in 2011. After elaborating on the key methodological points in the overall research strategy, the second chapter moves to consider the concrete research steps and accompanying methods.

Chapter III traces the historical parentage of the 'manifest failing' threshold and investigates the substance of its predecessor – the 'unable or unwilling' test in the context

of self-defence against non-state actors. After expanding on the basis for the analogy that supports the move to extend the 'unable or unwilling' test to the R2P, the chapter engages with the progress of legal scholars in detailing what a more specific 'unable or unwilling' test should entail. Specifically, it conducts a detailed investigation into the concrete factors that legal scholars have identified as relevant to making 'unable or unwilling' determinations. On the basis of this analysis the chapter formulates a continuum of analogous determinants and indicators that will inform the normative framework for making 'manifest failing' determinations in the context of situations that fall under the remit of the R2P, articulated in Chapter V.

Chapter IV constitutes a detailed analysis of the 2011 Libya crisis with a focus on the threshold for coercive action. It is divided in two parts, corresponding to the two lines of inquiry, specified above. The first part illustrates the broader debate over threshold, with a focus on the language used by states to represent the threshold for coercive action in the discourse on collective action in Libya. Drawing on the instances where states defended their position on the Libyan crisis, identified through the examination of the above debate, the second part of the chapter focuses on the concrete determinants and indicators that states tend to invoke in their justifications for forceful action to protect the Libyan populations from the threat of atrocity crimes. The determinants and indicators derived from the analysis of state discourses over international action in Libya, together with those identified in Chapter I and III, comprise the basis for the normative framework for 'manifest failing' determinations proposed in Chapter V. As 2021 marks the 10th anniversary of the Libyan intervention, now is an apt time to revisit the case and shed light on an important aspect of the situation, which has received little scholarly attention.

Chapter V presents the normative framework for 'manifest failing' inquiries proposed in this thesis by combining the insights into the determinants and indicators that should inform 'manifest failing' judgements, derived from the investigations in Chapters I, III and IV. It then demonstrates how the proposed 'manifest failing' framework can be applied to a country-specific case, namely the ongoing violent anti-drug campaign in the Philippines. As mentioned above, what makes this case particularly suitable for illustrating the utility of the proposed framework is the disparity in the way President Duterte's 'war on drugs' has been framed by 'the West' and ASEAN. While the actions of the Philippines' government have been met with condemnation by the UN, EU and ICC, which have raised concerns over the government's responsibility for the commission

of crimes against humanity, ASEAN states view 'the war on drugs through an anti-drug lens rather than a mass atrocity lens'.[113] By applying the framework developed in this thesis in light of the recently published results of UN and ICC investigations of the situation in the Philippines, it will be argued that Duterte's government is 'manifestly failing' to protect its populations and bears primary responsibility for acts of atrocities committed in the country.

The Conclusion highlights the key findings and contributions resultant from this research, reiterates the importance of pursuing the 'manifest failing' research agenda and sketches the ways in which it can be furthered through future research.

[113] Gallagher et al., 'Failing to fulfil the responsibility to protect', p. 264.

CHAPTER I: CONTRIBUTIONS TO CLARIFYING THE SUBSTANCE OF THE 'MANIFEST FAILING' THRESHOLD AND RELEVANT TOOLS FOR THE ANALYSIS OF ATROCITY CRIMES

As evidenced in the Introduction, the abstract formulation of the threshold for coercive action, when applied to the conduct of host states, raises questions about what the term 'manifestly failing' means.[114] The term 'manifestly failing' appears frequently in R2P-related texts and talks produced by UN officials, NGOs and even states. However, these mentions are rarely more than citations of the text of Paragraph 139 of the WSOD and almost never dwell on what makes a national government a 'manifestly failing' state. Against this backdrop, this chapter turns to the work of R2P and legal scholars who have contributed to the study of what constitutes a 'manifest failing' to shed light on what determines whether this threshold for collective action has been crossed. Aside from fulfilling the standard preliminary requirement of informing the reader of existing scholarly contributions, the aim here is to scrutinise a host of 'manifest failing' determinants and indicators that others have identified in order to assess their utility for informing the framework for 'manifest failing' inquiries proposed in Chapter V.

The chapter proceeds in four sections. The first section, explains why despite significant developments in the tools for the assessment of the risk of atrocities, such as the UN 'Framework of Analysis for Atrocity Crimes', existing framework for the analysis of atrocity crimes are insufficient to support inquiries into whether a national government is fulfilling its responsibility to protect. The second and third sections, respectively explore existing contributions to specifying the determinants and indicators of a 'manifest failing', with a view to singling out concrete determinants and indicators that will inform the 'manifest failing' framework to be proposed in Chapter V. The final section will revisit the most recent scholarly endeavour that explicitly tackles the issue of 'manifest failing' states in the context of the Libyan crisis and demonstrate why it falls short of contributing to our understanding of the 'manifest failing' threshold in the Libyan context. The analysis in this chapter will highlight the strengths of existing attempts to

[114] Gallagher, 'What constitutes a "Manifest Failing"?', pp. 437-438.; Ludmila Pavlova, 'The Concept "Responsibility to Protect": Analysis and Legal Evaluation', 4 (67) *Journal of International Law and International Relations* (2013): 3-8, p. 6.

clarify the criteria relevant to 'manifest failing' determinations, which will underpin the framework to be developed in this thesis, as well as locate existing lacunae in scholarly conceptualisations of what determines a 'manifest failing', which will be addressed over the course of this thesis in a way that contributes to the development of a comprehensive 'manifest failing' framework in Chapter V.

I. THE UN 'FRAMEWORK OF ANALYSIS FOR ATROCITY CRIMES'

The last decade saw the development of perhaps the most widely recognized tool for the assessment of atrocity crimes. In 2014 the UN published a document titled 'Framework of Analysis for Atrocity Crimes: A tool for prevention'. It was introduced to replace a 'framework of analysis to support the assessment of the risk of the crime of genocide from an early warning perspective' developed by the then Office of the Special Adviser on the Prevention of Genocide in 2009.[115] This revision emerged in response to the need for a tool that 'provides an integrated analysis and risk assessment tool for atrocity crimes', in light of the extension of the Office's responsibilities to assist in the work of the Special Adviser on the Responsibility to Protect and the need to 'reflect recent developments and new research into the processes that lead to those crimes'.[116]

The logic behind this framework is that since 'atrocity crimes are processes, it is possible to identify warning signs or indicators that they might occur'.[117] Therefore,

> 'If we **understand the root causes and precursors** of these crimes, and can **identify risk factors** that can lead to or enable their commission, it follows that we can also **identify measures that can be taken by States and the international community to prevent these crimes**'.[118]

The professed objective of this framework is 'to provid[e] a set of elements to help monitors or analysts to make **qualitative and systematic assessments of the risk of atrocity crimes** in specific situations'.[119] The UN framework is intended to 'serv[e] as a working tool for the assessment of the risk of atrocity crimes in all parts of the world and

[115] United Nations, 'Framework of Analysis for Atrocity Crimes: A tool for prevention', 2014 [Accessed 27 July 2020] Available at: https://www.refworld.org/docid/548afd5f4.html, p. 5.
[116] Ibid.
[117] Ibid. p. 4.
[118] Ibid.
[119] Ibid. p. 7.

in identifying those countries most at risk'.[120] It comprises 14 broad risk factors, 8 of which are 'common' to all atrocity crimes and 6 are crime-specific (related to either genocide, crimes against humanity or war crimes) and a variable number of indicators attached to each risk factor. Ethnic cleansing does not have its own crime-specific risk factors, but 'has been integrated into the analysis of the risk factors' pertaining to other atrocity crimes, because 'it includes acts that can constitute [those] crimes or elements of them'.[121] All risk factors within the UN framework represent:

> 'conditions that increase the risk of or susceptibility to negative outcomes… includ[ing] behaviours, circumstances or elements that create an environment conducive to the commission of atrocity crimes, or indicate the potential, probability or risk of their occurrence'. [122]

The indicators falling under each risk factor constitute a non-exhaustive list 'drawn from past and current cases', defined as 'different manifestations of each risk factor, [which] therefore assist in determining the degree to which an individual risk factor is present'.[123] Together the risk factors and indicators, 'reflect definitions of the crimes in international law, case law from the work of international courts or tribunals, and empirical analysis of past and present situations'.[124] It is envisioned that in order 'to be functional, assessments require the systematic collection of accurate and reliable information based on the [specified] risk factors and indicators'.[125]

The UN framework links a wide range of elements that can manifest themselves in a given situations to the probability of atrocity crimes occurring by grouping them together under broader risk factors. This makes it a formidable tool for the assessment of the risk of atrocity crimes, which can help identify and monitor situations of concern. As such, it is a well-designed implement for the fulfilment of its intended function – to assist and enhance early warning and preventative efforts. Beyond this, however, it tells us little about the assessment of situations in which atrocities are imminent or already underway with a view to determining whether a situation may merit international collective action

[120] Ibid. p. 5.
[121] Ibid.
[122] Ibid.
[123] Ibid. p. 6.
[124] Ibid. p. 5.
[125] Ibid.

as envisioned in the 2005 agreement on the R2P. Specifically, the presence of risk factors of atrocities in a given situation or even establishing a very high risk of atrocities does not necessarily mean that the 'manifest failing' threshold requirement has been satisfied. As clarified in the Introduction, even if mass atrocities are deemed to be taking place in a given case, it does not automatically follow that the host state is 'manifestly failing' in its responsibility to protect. Therefore, assessing the risk or presence of atrocities and of a 'manifest failing' are not one and the same thing.

To put this into context, in viewing atrocities from an early warning perspective, the UN framework identifies a very broad spectrum of risk factors, a number of which 'will be present in many situations or societies around the world where atrocity crimes have not taken place'.[126] Risk Factor 4 from the UN framework suggests that 'from an early warning perspective, it is extremely important to be able to identify motivations, aims or drivers that could influence certain individuals or groups to resort to massive violence as a way to achieve goals, feed an ideology or respond to real or perceived threats', because it: 1) 'allows for a higher degree of prediction of the likelihood of those crimes' and 2) 'it opens the opportunity to develop prevention strategies aimed at neutralizing or curbing those motives or incentives'.[127] Given that the authors of the UN framework admit that 'no one specific motive or incentive will automatically lead to atrocity crimes'[128], such an assessment will tell us even less about whether the host state is 'manifestly failing'. Besides, as will be argued later in this chapter 'motives' are not a reliable indicator of the host state's intent to commit atrocity crimes. Similarly, Risk Factor 1 of the UN framework, 'situations of armed conflict or other forms of instability' (political, social or economic), may 'highly increase the likelihood' of atrocity crimes, but not 'necessarily lead to' their occurrence.[129] More importantly, the fact that a state is a party to an armed conflict does not render it 'unwilling or unable' to protect its populations. Hence, such a determinant is not very helpful for ascertaining whether the host state is fulfilling its protection responsibilities or not. Consequently, it is clear that many aspects of the situation that are very much relevant from an early warning perspective do not translate well into a framework that is concerned with more advanced steps in the prevention of

[126] Ibid. p. 7
[127] Ibid. p. 13.
[128] Ibid.
[129] Ibid. p. 10.

atrocity crimes, pertaining to an imminent or ongoing crisis of a gravity that may warrant an international response.

Likewise, factors that are irrelevant or of secondary importance for a tool for prevention such as the UN framework, may be crucial and of primary importance in a framework that provides guidelines for assessing situations which may require a strong international response. Such indicators are found in crime-specific factors 10 to 13 of the UN framework. The limited relevance of late-stage prevention factors in the context of the UN framework is recognised in the comments on the risk factors specific to war crimes included therein. Namely, because the latter international crime 'must always take place in the context of an armed conflict' (unlike genocide and crimes against humanity), 'indicators specific to war crimes surface at a late stage, when options for prevention are more limited'.[130] In this regard, the UN advises that '[f]or earlier preventive action, common risk factors should be considered first'.[131] The same surely applies to indicators of other atrocity crimes that manifest themselves at a later stage. By contrast, these aspects of the situation would be particularly pertinent to assessments of a host state's 'manifest failing'.

In light of the explanations and examples discussed above it is clear that the UN 'Framework of Analysis for Atrocity Crimes' for atrocity prevention cannot double as a tool that can effectively guide relevant actors in making 'manifest failing' determinations. However, this does not mean that there will be no common ground between the framework offered by the UN and the 'manifest failing' framework proposed in this thesis. It is expected that some of the indicators of the risk of atrocities will overlap with indicators attached to various risk factors of the UN framework, because in both cases the indicators represent aspects of a situation, which have come to be associated with the presence of atrocities. However, what makes each framework distinct and renders it effective in answering some questions and not others is not the corpus of indicators it comprises but the way in which these elements or aspects of a situation relevant to the analysis of mass atrocities are assembled under different determinants ('risk factors' in the UN framework), in order to help guide specific assessments. Therefore, the UN framework itself is not a suitable source for identifying or justifying the inclusion of new determinants and indicators of 'manifest failing' in the framework being developed in the

[130] Ibid. p. 23.
[131] Ibid.

proposed thesis. However, once a given determinant or a family of indicators have been identified as relevant to determining a 'manifest failing' through other sources/methods, this thesis will draw on the UN framework in a bid to expand the list of possible indicators that can manifest themselves in a given case.

Ultimately, the UN framework is an excellent tool for early warning and the prevention of atrocity crimes, but following it will not provide us with an answer to the question of whether a host state is failing in its responsibility to protect or not, because it was not designed to do so. On the contrary, it specifically focuses on the early prevention of atrocities because it is 'much less costly than intervening to halt these crimes' and 'the political cost and challenges of early engagement by the international community are also less than when the crimes are imminent or ongoing, by which time options for preventive action are much more limited'.[132] Nonetheless, as exemplified in the Introduction, situations of the latter variety still occur more often than anticipated. In this regard, Reike argues that '[i]t seems likely that… pillar three late-stage prevention will be the norm rather than the exception for as long as we have not developed better structural indicators to foresee mass atrocity risk early enough to engage in meaningful structural prevention'.[133] Hence, we also need to develop tools for the analysis of situations which may warrant an enquiry into whether the host state is 'manifestly failing' to protect its populations, in order to enhance international preparedness to respond adequately to the threat of imminent or ongoing atrocities. The next two sections of this chapter focus precisely on extant scholarly contributions to this research agenda.

II. THE DETERMINANTS OF A 'MANIFEST FAILING'

While there has been no conscious attempt to identify the determinants of 'manifest failing' within the existing literature on R2P and mass atrocities, there is one contribution to the study of the 'manifest failing' threshold that takes some intuitive first steps in this direction. That is the work of Professor Sheri Rosenberg and Ekkehard Strauss, an Expert Consultant on mass atrocities,[134] on a 'multi-staged research project', titled 'Assessing the Parameters for Identifying a "Manifest Failure" to Protect Populations under R2P', which was selected by the APR2P as one of 14 projects to receive funding from a AUD$2

[132] Ibid. p. 2.

[133] Reike, 'Libya and the Responsibility to Protect: Lessons for "Pillar Three"', p. 79.

[134] Sheri P. Rosenberg and Ekkehard Strauss, 'A Common Standard for Applying R2P', Policy Brief (Cardozo Law School, Holocaust, Genocide, and Human Rights Program, 2013): 1–8, p. 8.

fund established by the Australian government in support of R2P initiatives.[135] The main outcome of this project was the dissemination of Rosenberg and Strauss' common approach / standard for the application of R2P in several publications - a chapter in a collection of papers titled 'Operationalising the Responsibility to Protect: A Contribution to the Third Pillar Approach', an APR2P policy brief, and a 2013 Cardozo Law project report and policy brief.[136] However, Rosenberg and Strauss' work did not result in further outputs to match the expectations set by the project title, namely to shed more light on the parameters for identifying a 'manifest failing'. It is likely that the continuation of this research and the publication of additional findings were prevented by the untimely passing of the principal investigator in 2015. Notwithstanding, much can be learned from a broader framework for the assessment of 'situations for risks of mass atrocities', especially one that (despite its focus on prevention as 'the best form of protection') promises to offer 'principles for the determination of whether a state is "manifestly failing" to protect its population from one or more of the four RtoP acts' in keeping with the text of the 2005 WSOD.[137]

Rosenberg and Strauss helpfully position their endeavour to find the most appropriate standard for the application of the R2P within what they believe to be the four distinct 'elements encompassed in the analysis of a situation through a potential RtoP lens' – a 'substantive dimension', a 'gravity dimension', a 'temporal dimension' and 'the consequences of a situation falling within the RtoP'.[138] These four elements inform the guiding principles for considering a situation under the R2P, which include a determination of the host state's 'manifest failing'. Hence, they are worth unpacking with a view to establishing the key elements that should underpin a comprehensive 'manifest failing' framework.

The first element is the 'substantive dimension', which is concerned with the type of human rights violations associated with the four atrocity crimes that define the scope of

---

[135] Gallagher, 'What constitutes a "Manifest Failing"?', pp. 432-33. Noële Crossley, *Evaluating the Responsibility to Protect: Mass Atrocity Prevention as a Consolidating Norm in International Society* (Abingdon and New York: Routledge, 2016), pp. 108-109.

[136] Rosenberg and Strauss, 'A Common Approach', pp. 55-72; Rosenberg and Strauss, 'A Common Standard', Policy Brief (2013), pp. 1-8; Sheri Rosenberg and Ekkehard Strauss, 'A Common Standard for Applying R2P', APR2P Brief, 2 (2) (2012): 1-6; Sheri Rosenberg and Ekkehard Strauss, 'A Common Standard for Applying R2P' (Cardozo Law School, Holocaust, Genocide, and Human Rights Program, 2013): 1-148.

[137] Rosenberg and Strauss, 'A Common Approach', pp. 55-56.

[138] Ibid. p. 57.

the R2P and whose substantive content can be found in 'legal and policy documents'.[139] Pattison similarly claims that it is widely recognised 'that, to have just cause, humanitarian intervention needs to tackle the ongoing or impending mass violation of basic human rights', and coercive intervention under the R2P should 'be in response to the manifest failure to tackle ongoing or impending genocide, war crimes, crimes against humanity, and ethnic cleansing'.[140] Despite his noninterventionist stance, Puri acknowledges that when gross or persistent human rights violations occur, 'a determination would need to be made whether the requisite threshold (which warrants coercive measures on the part of the international community) has been reached'.[141]

Second, the 'gravity dimension' is a reflection of the fundamental understanding that the R2P applies 'at a specific level of gravity or seriousness of potential violations, with this level set at below that which would place an individual as criminally liable or a state internationally responsible'.[142] As Puri explains, even though the trigger point for international actions cannot be human rights violations, the UN Security Council does not have to wait 'till genocide actually occurs before it takes action [as it] happened in the case of Rwanda – in breach of' the 1948 Convention on Genocide.[143] On the one hand, 'a small-scale example of genocide, war crimes, crimes against humanity or ethnic cleansing may take place [from a legal standpoint], which would not meet the threshold required to merit a pillar three response'.[144] As Gallagher exemplifies, 'a war crime can take place if private property is targeted and [as mentioned in the Introduction] genocide could be deemed to have taken place if just one person or a group of hostages were killed, as long as it was proven that their murder was intended to destroy their national, racial, ethnic or religious group in whole or in part'.[145] This lends support to the understanding that the gravity of human rights violations has an important role to play in determining whether collective action is warranted, as part of the determination of the host state's 'manifest failing'.

---

[139] Ibid.

[140] James Pattison, 'Perilous Noninterventions? The Counterfactual Assessment of Libya and the Need to be a Responsible Power', 9 (2) *Global Responsibility to Protect* (2017): 219-228, p. 222.

[141] Hardeep Singh Puri, *Perilous Interventions: The Security Council and the Politics of Chaos* (London: HarperCollins, 2016), pp. 199-200.

[142] Rosenberg and Strauss, 'A Common Approach', p. 57.

[143] Puri, *Perilous Interventions*, p. 200

[144] Gallagher, 'What constitutes a "manifest failing"?', p. 438.

[145] Ibid.

The third 'temporal dimension', which is the main focus of Rosenberg and Strauss' work, pertains to 'the standard of proof that is applied to determine when that level of seriousness or gravity set out in the second element has been reached, and therefore when the RtoP framework (and its corresponding set of responsibilities or potential legal obligations) is applicable'.[146] According to authors, this standard, ranging from 'potentially applicable' to 'definitively proven', is necessary 'for all potential forms of state or collective action, whether coercive or otherwise'.[147] In other words, the 'temporal dimension' is concerned with the threshold of evidence against which to measure incoming information about relevant situations in order to determine when the R2P applies. Because of R2P's normative agenda 'to proactively attempt to prevent imminent or on-going forms of mass atrocities', if this standard was to only apply at the point when one of the four crimes 'has definitely occurred', the whole concept of a 'responsibility to protect' populations would be meaningless.[148] Hence, Rosenberg and Strauss argue that the R2P calls for a different kind of standard than that of establishing 'responsibility under international criminal law for an individual culprit', because '[t]he assessment of the likelihood of prospective conduct is by its nature a very different enquiry than the assessment of the evidence to determine whether a fact has been proven about a past event'.[149] More specifically, the standard should relate to the primary responsibility of states to protect their populations from the four atrocity crimes premised on existing international law 'which prevail[s] independently of the RtoP at all times' as well as take into account the 'future risk aspect of the prevention component of the RtoP and combine prospective and retrospective elements'.[150] Put simply, it 'requires the prospective assessment of future developments based on present facts and circumstances'.[151] There is little doubt that the same evidentiary standard should apply to 'manifest failing' determinations, even though the frequent use of the terms 'manifestly failed' and 'manifest failure' in the discourse over threshold would suggest otherwise.

Employing the phrases 'manifest failure' or 'manifestly failed' instead of 'manifest(ly) failing' is problematic precisely because it suggests that determining whether this

[146] Rosenberg and Strauss, 'A Common Approach', pp. 57-58.
[147] Ibid. p. 58.
[148] Ibid. pp. 55, 57.
[149] Ibid. pp. 58-59.
[150] Ibid. p. 70.
[151] Ibid. p. 55.

threshold has been crossed involves judging whether a fact about a past event has been established. However, as Bellamy argues, in the context of Paragraph 139 of the WSOD, '"Manifestly failing"… should be understood as an unfolding process ["failing"] rather than an end product ["failed"]' (square brackets used in original text).[152] Likewise, Gallagher reasons that it is important to recognise that 'the drafters [of the WSOD] chose the phrase "manifest failing" rather than "manifest failure" precisely because they wanted to highlight that this should be identified as an ongoing process (manifest failing), so that we do not have to wait until the point that the crime is indeed over (manifest failure)'.[153] This distinction is not to be overlooked as it could make the difference between 'timely and decisive' action by the international community to protect populations at risk of atrocities and the lack thereof. If the international community could only act when the state has failed to protect its populations from atrocity crimes, then the whole notion of a subsidiary international 'responsibility to protect' would be rendered meaningless. Consequently, Rosenberg and Strauss' suggestion that 'the enquiry involving the RtoP will often, perhaps always, have elements of both forward-looking and backward-looking investigations, assessing whether sufficient acts have occurred to fall within the RtoP and whether future atrocities are potentially to occur' is particularly pertinent to inquiries into the host state's manifest failing'.[154]

In order to arrive at such a standard for the application of the R2P, Rosenberg and Strauss conduct an in-depth analysis of diverse areas of national and international law in which 'similar challenges' with regards to ascertaining 'the risk of a violation of international legal obligations in the future based on present facts and circumstances' arise.[155] Their review demonstrates how international and national courts have addressed this problem by 'engag[ing] in balancing of the probability of an event occurring based on the evidence available at the time of the decision, with the level of harm that would occur if such a situation would develop'.[156] Based on this investigation the authors find that in the practice of different courts in different legal areas 'the risk that must be demonstrated ranges from the low end of "reasonable possibility" to the high end of "substantial

[152] Bellamy, 'The First Response', p.11.
[153] Gallagher, 'Syria and the indicators of a "manifest failing"', p. 8.
[154] Rosenberg and Strauss, 'A Common Approach', p. 58.
[155] Ibid. pp. 55, 69.
[156] Ibid. p. 69.

risk".'[157] By drawing parallels between these contexts and the R2P, they argue that the standard used 'to determine the prospect of serious crimes and seek to avoid such crimes prospectively' should lie in the middle of this spectrum, given that the R2P requires balancing between the need to prevent a level of harm that is 'exceptionally grave' and 'the prerogatives of sovereignty'.[158] This 'mid-level' evidentiary standard, referred to as the 'real-risk', caters to the need to 'individuali[se] risks and consider concrete scenarios', thus ensuring a 'reasonable and measured' response to the threat of atrocity crimes.[159] In essence, Rosenberg and Strauss' assert that a 'situation will be considered in the context of the RtoP, if the examination of the situation establishes a real risk that exceptionally grave human rights violations, as described in genocide, war crimes, crimes against humanity and ethnic cleansing, are occurring or could occur in the future'.[160] The aim of this assessment of available evidence is to ascertain the 'likelihood of future conduct' and may draw on information of 'past events as relevant'.[161] Likewise, determining whether a 'manifest failing' is underway would involve assessing the present conduct of states in relations to the threat of atrocity crimes, based on available facts about a given situation where atrocities are suspected to be taking place, and may involve considering past conduct regarding the commission of serious human rights violations.

The fourth element, 'the consequences of a situation falling within the RtoP' reflects the fact that the R2P does not stipulate which specific measures ought to be taken in a given situation for the implementation of the responsibilities specified therein but 'offers a methodology to address an exceptional set of situations more effectively through a continuum of steps'.[162] In this regard, the authors' objective is to propose a standard that is to be used 'on a case-by-case basis' with a view to 'assisting relevant actors to determine, whether a situation could benefit from applying the RtoP' by 'provid[ing] clear boundaries to discussions over when the RtoP applies'.[163] They envision that such a standard would at the very least serve to compel relevant actors to 'explain their reasoning from a common reference point'.[164] As outlined in the Introduction, the vision

[157] Ibid.
[158] Ibid.
[159] Ibid. pp. 69-70.
[160] Ibid. p. 71
[161] Ibid.
[162] Ibid.
[163] Ibid. pp. 55-56.
[164] Ibid.

for and rationale behind this thesis' proposal for a framework to aid determinations as to whether a host state is 'manifestly failing' to protect its populations is similar. Namely, to devise a comprehensive normative framework, comprising a continuum of determinants and indicators that can aid relevant stakeholders in 'manifest failing' inquiries on a case-by case basis and create a common reference within which argumentation and justification can take place.

These four elements of analysing an situation through an R2P lens are reflected in Rosenberg and Strauss' guiding principles, designed to 'offer an approach to assess whether a situation would benefit from applying the RtoP and suggests [sic.] a procedure how to identify required action'.[165] The four principles encompass: 1) 'determination of relevant human rights violations'; 2) 'determination of the level of gravity and seriousness of potential threats'; 3) 'application of RtoP'; and, 4) 'determination whether a State is "manifestly failing" to meet the RtoP'.[166] While only the last principle includes direction as to how to determine whether a 'manifest failing' is taking place, its definition suggests that its assessment is based on the first three determinants.

Namely, Rosenberg and Strauss specify that the determination of a 'manifest failing' should be premised on: 1) 'the information of relevant human rights violations'; 2) 'the state of implementation of measurable steps to mitigate risk factors, and its impact on the real risk that exceptional grave violations of human rights could occur in the future'.[167] Although the authors do not clarify what the 'information of relevant human rights violations' entails, in the context of their overall framework of guiding principles for the application of the R2P and the aforementioned four dimensions for assessing R2P situations, it is clear that it encompasses the determinations of the concrete human rights abuses relevant to atrocity crimes and their gravity, i.e. the first two of their guiding principles for the application of R2P (specified above). What is more, the above discussion of the substantive and gravity dimensions of 'manifest failing' assessments revealed that both determinations (of the nature and gravity of the threat of atrocities that vulnerable populations are faced with) are essential to the analysis of any situation with a view to determining the need for a collective action response or the precise form that this response should take. For these reasons, Rosenberg and Strauss' understanding of

[165] Ibid., p. 70.
[166] Ibid., pp. 71-72.
[167] Ibid.

these determinations will be reflected in the determinants of the 'manifest failing' framework presented in Chapter V of this thesis.

With regards to the 'determination of relevant human rights violations', Rosenberg and Strauss explain that by specifying atrocity-related human rights violations they aim to identify 'a common consciousness of the risks involved in any massive violation of human rights', rather than segregated 'legal categories of mass atrocity crimes on the one hand and other human rights violations on the other'.[168] For the purpose of this determination, they provide a list of 'human rights violations [that] have been of particular relevance in past cases of mass atrocities' (see Box 1).[169]

Box 1. Rosenberg and Strauss' list of human rights violations relevant to past cases of mass atrocities:

'killings, torture, mutilation, rape and sexual violence, abduction, forced population movement, expropriation, destruction of property, looting, lack of freedom of speech/ press/ assembly/ religion, destruction of subsistence food supply, denial of water or medical attention, man-made famine, redirection of aid supplies, discrimination in access to work and resources, political marginalization, restricted movement, discrimination in education and lack of access to justice and redress'.[170]

The second determination specified by Rosenberg and Strauss relates to the gravity of these human rights violations and proposes to assess their 'significance…in light of the number of potential victims of violence or level of irreparable harm that may be caused to potential victims'. To do so, this determination takes into account risk factors informed by the understanding that a central aspect of the extraordinary situations that fall within the remit of the R2P is '[t]he persecution of large parts of the population based on identities applied by the perpetrators'.[171] From the list of relevant identity criteria provided by Rosenberg and Strauss (see Box 2), it is clear that they are upholding a

[168] Ibid. p. 71.

[169] Ibid.

[170] Rosenberg and Strauss, 'A Common Approach', p. 71.

[171] Ibid. Rosenberg and Strauss further assert that the standard to be applied for 'fact-finding on present and past events' should be 'reasonable suspicion…, which is met when a reliable body of evidence [that must be corroborated] indicates the occurrence of a particular incident or event'.: Ibid.

broader definition of group identity than the archaic conception of this notion found in the Genocide Convention, which is only concerned with the 'intent to destroy, in whole or in part, a national, ethnical, racial or religious group'.[172]

Box 2. Rosenberg and Strauss' risk-factors for determining the level of gravity of prospective human rights violations considering the number of victims or extent of irreparable harm:

'identification of the victims based on identity criteria linked to race, colour, descent, religion, ethnic, or national origin, gender, sexual orientation or other ground and their association with a specific political opinion or group; public hate speech, incitement to violence, or humiliation of a group publicly or in the media; exclusionary ideologies that purport to justify discrimination; a past history of violence against perceived groups; and a climate of impunity in which these events unfold.'[173]

Contemporary scholars studying mass atrocities, including Gallagher and Leaning, have criticised and rejected the legal definition of genocide for its narrow definition of group identity and its concomitant failing to protect other identifiable groups.[174] To rectify this, scholars such as Gallagher and Shaw have suggested that the definition of genocide should be expanded to protect all groups.[175] In terms of specifying the parameters of group identity, Gallagher upholds the seminal conceptualisation of this notion found in Chalk and Jonassohn's definition of genocide as 'a form of one-sided mass killing in which a state or other authority intends to destroy a group, as that group and membership in it are defined by the perpetrator'.[176] The same contemporary understanding of group identity appears to be implied in Rosenberg and Strauss definition of 'identities applied by the

[172] UN General Assembly, *Convention on the Prevention and Punishment of the Crime of Genocide*, A/RES/260 (III), 9 December 1948.

[173] Ibid.

[174] Adrian Gallagher, *Genocide and its Threat to Contemporary International Order* (Basingstoke: Palgrave Macmillan, 2013), pp. 32-37. Jennifer Leaning, 'Enforced Displacement of Civilian Populations in War: A Potential New Element in Crimes against Humanity', 11 (3) *International Criminal Law Review* (2011): 445-462, p. 458.

[175] Gallagher, *Genocide and its Threat to Contemporary International Order*, p. 36. Martin Shaw, *What is Genocide?* (Cambridge: Polity Press, 2007), p. 78.

[176] Gallagher, *Genocide and its Threat to Contemporary International Order*, p. 26. Frank Chalk and Kurt Jonassohn, *The History and Sociology of Genocide, Analyses and Case Studies* (London: Yale University Press, 1990), p. 23.

perpetrator'. This conceptualisation will be explicitly reflected in the framework proposed in this thesis, in order to take into account the many kinds of groups that are targeted for grave harm 'as identified by the perpetrator' (see Table 1.1 at the end of this chapter).

Having proposed guidelines for determining the nature and gravity of the threat, Rosenberg and Strauss' third principle specifies concrete measures / measurable steps (see Box 3) that national governments and the international community can take 'to mitigate the real risk of mass atrocities', as the R2P encourages them to do, even though it does not instruct what these measures ought to be.[177] Based on such consecutive measures, the authors suggest that 'the compliance of national governments and the international community can be established' as part of the determination as to whether a 'manifest failing' is taking place.[178] Significantly, they specify that '[m]anifest failure occurs when foreseeable consequences have not been addressed and the risk level prevails or increases'.[179] For the purposes of this determination, the authors envision that '[t]he impact of measures taken by the government and their impact on the risk-level should be monitored on a permanent basis against the timelines and indicators of the concept'.[180] The authors further postulate that 'the nature and timeline of the steps [based on existing resources and strategies] depends on the gravity and urgency of the situation' and remind that the actions of international actors are auxiliary to those of the host state.[181]

Rosenberg and Strauss' proposals for how to determine whether a 'manifest failing' is taking place based on the outline of concrete measures to mitigate the threat of atrocities are sound and instructive. However, as far as determining the national government's 'manifest failing' is concerned, the focus should be on the measures taken by the host or the compliance of the host state with atrocity prevention measures taken by the international community, as explained in Chapters III and IV. Whether the international community takes measures to mitigate the threat of atrocities is a different matter.

[177] Rosenberg and Strauss, 'A Common Approach', p. 72.
[178] Ibid.
[179] Ibid.
[180] Ibid.
[181] Ibid.

Box 3. Rosenberg and Strauss' list of steps / measures to mitigate the risk of mass atrocities:

'public acknowledgement and condemnation of human rights violations; clear and public orders to military, police or security forces to respect international human rights and humanitarian law; immediate enforcement of accountability for the most relevant violations; ensuring humanitarian assistance and protection for victims of violence; and, in cooperation with relevant stakeholders, including potential victims and drawing an action plan with timelines for mitigating the most urgent risk factors'.[182]

Overall, Rosenberg and Strauss make a strong contribution to the evidentiary standards that should apply to R2P assessments, which is very informative, thorough and well-reasoned. Their examination of different areas of international law sheds light on the complexities of making prospective assessments based on present, and sometimes past, facts and circumstances. This is a challenge that ought to be taken into account in the present endeavour to specify the determinants and indicators of a 'manifest failing'. The guidelines (i.e. the determinations discussed above) accompanying the evidentiary standard proposed by Rosenberg and Strauss are instructive as to how to the determine whether a 'manifest failing' is taking place. However, they come as an afterthought in the last two pages of an article dedicated to determining the appropriate standard of proof to be applied to situations when they are being analysed through an R2P lens.

Although Rosenberg and Strauss identify several determinants relevant to the determination of the host state's 'manifest failing', their explanation as to how to judge whether this threshold has been crossed is rather abstract, mostly relying on descriptions of the process and definitions. For instance, it is not clear how to assess whether and how the measures taken by the host state and the international community impact the risk of atrocities occurring in the future, which is an impediment to determining to what extent the host state is failing in its responsibility to protect. The vagueness of this determination stands out particularly in comparison with the first three principles specified by Rosenberg and Strauss, with clearly defined concrete parameters / risk-factors attached to them, making them easy to follow and apply with regards to the assessment of a given

[182] Ibid.

situation. This can be attributed to the fact that Rosenberg and Strauss pay relatively little attention to the 'manifest failing' threshold (elaborated on in a 100 words) in the context of the presentation of their 'common approach for the application of R2P', which does not allow them to address such a complex problem as the substantive indeterminacy of this threshold requirement. Therefore, more research into the substance of the 'manifest failing' threshold is required in order to propose a continuum of determinants and indicators that can effectively aid stakeholders in their attempts to determine whether this substantive threshold for international action has been crossed.

All things considered, Rosenberg and Strauss offer some good foundations for the developments of similar types of frameworks designed to guide assessment of potential R2P scenarios. The following key determinants proposed by Rosenberg and Strauss, namely the information of relevant human rights violations, the gravity of these violations and the steps / measures to mitigate the risk of mass atrocities, as well as the indicators associated with them that were identified as key to 'manifest failing' determinations, will be taken forward in the specification of the framework to be proposed in this thesis. Where additional indicators relevant to these determinations surface in the discussion in the remainder of this chapter, they will be added to the lists of indicators provided by the two authors. This cumulative list of indicators associated with each determinant can be seen in a table at the end of this chapter, which combines all insights into determining a 'manifest failing' discussed here.

III. THE INDICATORS OF A 'MANIFEST FAILING'

Building on Rosenberg and Strauss' endeavour to help relevant stakeholders determine to what extent national governments are mitigating the risk of atrocity crimes, Gallagher makes 'the first tentative steps towards establishing a list of specific criteria… to try and identify the qualitative and quantitative indicators that actors invoke when attempting to make the case that the threshold of a 'manifest failing' has been reached'.[183] Specifically, he proposes a set of five indicators informed by 'interdisciplinary research into mass violence…and applies them to Syria: (i) government intentions, (ii) weapons used, (iii) death toll, (iv) number of people displaced, and (v) the intentional targeting of civilians, especially women, children and the elderly'.[184] Whilst recognising that there are a host of

[183] Gallagher, 'What constitutes a "Manifest Failing"?', p. 439.

[184] Gallagher, 'Syria and the indicators of a "manifest failing"', p. 1.

other factors which precluded an international consensus on Syria, such as 'Russia's relationship with Syria combined with its opposition of Western intervention', he contends 'that the (il)legitimacy and (ir)responsibility of the Syrian government is an important issue which, from within an R2P framework, asks us to question whether the regime has "manifestly failed"?'.[185] The choice of case 'does not aim to imply that there have been explicit United Nations (UN) Security Council debates over whether the Syrian regime is "manifestly failing", but rests on the premise 'that it is important to learn lessons from this "R2P failure of the first order", as it has been characterised by R2P scholars.[186] In this sense, Gallagher suggests that analysing the broader discourse surrounding Syrian case can serve to inform our understanding of 'what evidence we base judgments of a "manifest failing" on'.[187] The rest of this section will explore the content of the five indicators identified by Gallagher and expand on each by drawing on the broader literature on mass atrocity crimes, so as to address some important themes and questions pertinent to any substantive investigations of the 'manifest failing' threshold.

Government intentions

Gallagher argues that the role of the host state should be 'the starting point' when trying to determine whether a national government is 'manifestly failing' to protect its populations.[188] What makes ascertaining the role of state or non-state actors a challenging undertaking is 'the "Other Minds problem" of trying to assess the intentions of decision-makers when we simply do not know what is going on in their minds'.[189] Still, intent is notoriously difficult to prove, because it is much more common for state authorities intent on mass atrocities to conceal their true purpose and act covertly by 'hiring militia to perform mass killing, denying the commission of crimes, arguing that their victims were not civilians, or insisting that the crimes were committed by rogue elements'.[190] As Bellamy points out, Colonel Qaddafi's Rwanda-esque statements that openly communicated the regime's resolve to commit crimes against humanity are a testimony to Libyan exceptionalism, rather than the norm.[191] Therefore, except on the rare occasion

[185] Ibid. p. 6.

[186] Ibid. p. 2

[187] Ibid.

[188] Gallagher, 'Syria and the indicators of a "Manifest Failing"', p. 6.

[189] Ibid.

[190] Alex J. Bellamy, 'Libya and the Responsibility to Protect: The Exception and the Norm', 25 (3) *Ethics & International Affairs* (2011): 263-269, pp. 265-266.

[191] Ibid. p. 265.

that the perpetrator comes forward to announce their plan to obliterate a group targeted on the basis of their self-proclaimed or assigned identity, it is hard to ascertain the extent to which a host state bears responsibility for the commission of atrocities against its populations. To be clear, 'the R2P does not require the international community to prove intent in the same way that the 1948 Genocide Convention does',[192] not least because it is activated earlier than the point at which international state responsibility or individual liability for one of the four crimes can be established under international law, as discussed in the previous section. However, as Gallagher suggests, 'if it can be established that the government in question is deliberately facilitating and/or perpetrating any of the four crimes then this is a clear cut indicator that a "manifest failing" is taking place'.[193]

According to Gallagher, the most accomplished approach to resolving this evidentiary problem is the behaviour-based understanding of intent, which arose in the context of debates surrounding the requirement of intent found in the legal definition of genocide.[194] Specifically, the acts/crimes listed under Article II of the Convention on the Prevention and Punishment of the Crime of Genocide constitute the crime of genocide only if they are 'committed with intent to destroy, in whole or in part, a national, ethnical, racial or religious group'.[195] The behaviour-based approach, advocated by Herbert Hirsch and Helen Fein, proposes that 'instead of emphasizing an obscure and impossible-to-define psychological state of intent', it can be established by 'showing a pattern of purposeful action'.[196] According to Gallagher, this means that the focus should be on state policy, as reflected in the words of Hilary Clinton in regards to the Syrian crisis: '[w]e will judge Assad's sincerity and seriousness by what he does, not by what he says'.[197] On Gallagher's analysis, in the case of the Syrian government, a pattern of purposeful action indicative of intent can be inferred by the war crimes and 'the crimes against humanity of murder, torture, rape, enforced disappearance and other inhumane acts' committed by the Syrian military and affiliated militia.[198] Other examples of behaviour-based indicators of

[192] Gallagher, 'Syria and the indicators of a "manifest failing"', p. 6.

[193] Ibid.

[194] For a discussion of the alternative motives-based and knowledge-based approaches and their comparative disadvantages in relation to the behaviour-based approach see Gallagher, *Genocide and its Threat to Contemporary International Order*, pp. 19-23.

[195] For the list of the crimes that may constitute genocide if genocidal intent can be established, see Article II of the text of the Convention in Annex I.

[196] Gallagher, 'Syria and the indicators of a "manifest failing"', p. 6.

[197] Ibid.

[198] Ibid. p. 7.

intent include the abovementioned threats to massacre Libyan civilians made by the Qaddafi regime, as well as 'the historical policy record of Slobodan Milosevic', which was factored in the decision-making process on Kosovo.[199] Rosenberg and Strauss also acknowledge indicators of intent, namely 'public hate speech, incitement to violence, or humiliation of a group publicly or in the media', as part of their broader determination of gravity.

In addition to the examples provided by Gallagher, Rosenberg and Strauss, Butcher et. al identify a wide range of indicators of intent consistent with the behaviour-based approach. The requirement of intent that underpins their Targeted Mass Killing (TMK) dataset 'implies that people of a particular ethnic, religious, or political group are selectively killed with 'the intent to destroy to intimidate a group by creating the perception of an imminent threat to that group's survival'.[200] As part of their approach to determining intent, the authors propose 'two higher-order measures of genocidal intent…: (1) public statements and (2) systematic preparation.[201] The first measure is premised on the observation that '[s]pecific types of hate speech are, anecdotally, common precursors to genocidal violence'.[202] Hence, intent can be inferred by examining 'primary and secondary sources for public statements of intent to destroy a group or public statements that deadly violence was specifically directed toward a group (e.g., as "enemies of the state").[203] Likewise, the UN Framework of Analysis for Atrocity Crimes suggests that '[o]fficial documents, political manifests, media records, or any other documentation through which a direct intent, or incitement, to target a protected group is revealed, or can be inferred in a way that the implicit message could reasonably lead to acts of destruction against that group' are indicative of government intent.[204] Rosenberg et al. also assert that hate speech is rightly considered to be one of the most oft-cited 'drivers of atrocities', particularly when they involve the targeting of a group of people based on their identity', owing the crucial 'role political power plays in creating and manipulating human emotion

[199] Ibid.

[200] Charles Butcher, Benjamin E. Goldsmith, Sascha Nanlohy, Arcot Sowmya, and David Muchlinski, 'Introducing the Targeted Mass Killing Data Set for the Study and Forecasting of Mass Atrocities', 64 (7-8) *Journal of Conflict Resolution* (2020): 1524-1547, p. 1529. The authors characterise TMKs as 'a set of events that cross a minimum threshold of severity [(a low minimum death toll)] and intent for killing specific identity groups'.: Ibid. p. 1528.

[201] Ibid. p. 1529.

[202] Ibid. p. 1532.

[203] Ibid. pp. 1531-1532.

[204] UN, 'Framework of Analysis for Atrocity Crimes', p. 19.

and inciting people to carry out heinous and gruesome acts against their fellow human beings'.[205] Hence, public statements of intent, are key to examining the role of the host state for the atrocities crimes being committed on its territory and determining whether it is 'manifestly failing' to protect its populations.

Recognising that perpetrators may not declare their intentions, in cases where there is 'evidence of systematic political, logistical, or organizational preparation to facilitate large-scale killing of targeted groups' Butcher et al. turn to Verdeja's suggestion that 'genocidal intent can often be inferred from level of lethality, degree of coordination, and scope (portion of the victim group affected)'.[206] While the authors consider the elements of death toll and how it relates to the size of the affected victim group under a separate category of 'severity' of episodes of TMK, their assessment of intent draws on 'Verdeja's "degree of coordination" [, which] captures activities that may signal preparation for genocide'.[207] Specifically, the authors claim that:

> 'Clear territorial control by the perpetrator group in the affected area, separation of people on the basis of identity, destruction of cultural symbols, systematic use of sexual violence against a population, the pattern of refugee origin and internal displacement, and the clear development of organizational infrastructure for genocide potentially contribute to evidence of intent'.[208]

Other indicators of systematic preparations can be found under Risk Factor 10 of the UN Framework, 'Signs of an intent to destroy in whole or in part a protected group', including the the use of 'methods or practices of violence that are particularly harmful against or that dehumanize a protected group, that reveal an intention to cause humiliation, fear or terror to fragment the group, or that reveal an intention to change its identity', 'means of violence that are particularly harmful or prohibited under international law, including prohibited weapons, against a protected group' and 'widespread or systematic discriminatory practices' against a protected group.[209] Similarly, Rosenberg and Strauss cite 'exclusionary ideologies that purport to justify discrimination' under their

[205] Sheri Rosenberg, Tibi Galis, and Alex Zucker, 'Introduction', in Sheri Rosenberg, Tibi Galis, and Alex Zucker (eds.) *Reconstructing Atrocity Prevention* (New York: Cambridge University Press, 2016): 1-16, p. 10.

[206] Butcher et al., 'Introducing the Targeted Mass Killing Data Set', p. 1532.

[207] Ibid.

[208] Ibid.

[209] UN, 'Framework of Analysis for Atrocity Crimes', p. 19.

determination of gravity (see Box 2). Butcher et al. conclude that, 'the strongest evidence of intent exists when both categories are met: perpetrators declare they plan to kill the targeted group on a mass scale, and they make the observable preparations using the resources of the state or other resources at their command'.[210] One example of such compelling evidence of intent is the situation in the Central African Republic in 2013, where 'Christian self-defense militias calling themselves "anti-Balaka" were responsible for killings of Muslims'.[211]As Butcher et al. describe the situation: 'Anti-Balaka militias entered villages over which they established clear territorial control; selected out Muslim men, women, and children to be murdered; destroyed Mosques; and sometimes announced that they were going to "kill all Muslims" in the village'.[212] In this context 'organisational evidence of intent' was manifested in 'territorial control, separation of Muslims, and targeting of cultural symbols', whereas the threats to destroy Muslim populations and their habitat represented public statements of intent.[213]

The above discussion and relevant illustrations suggest that '[t]he role of the state has to play an integral part of any future assessment of 'manifest failing' and… that a behaviour-based approach [for establishing intent] provides a basis for assessing to what extent the host state has failed in its R2P'.[214]

Death toll

Gallagher argues that civilian death toll could indicate whether a 'manifest failing' is under way or not because '[t]he more people are killed, the more one would expect that the government is either unable or unwilling to prevent mass violence'.[215] In order to define this indicator, he subscribes to Robert Pape's view that the death toll threshold for a 'pragmatic humanitarian intervention' is crossed in the instance of 'a mass homicide campaign sponsored by the local government in which thousands have died and thousands more are likely to die', for several reasons, which render it well-suited to aiding 'manifest failing' determinations.[216] First, in terms of quantifying death toll, he finds that Pape's definition is preferable to highly determinate thresholds which set a rigid quantitative

[210] Butcher et al., 'Introducing the Targeted Mass Killing Data Set', p. 1532.
[211] Ibid.
[212] Ibid.
[213] Ibid.
[214] Ibid.
[215] Ibid.
[216] Pape, 'When Duty Calls', p. 43, see also p. 53.

threshold of 5, 000 (Bellamy) or 50, 000 deaths (Valentino), because if a case of mass violence fails to meet a specific quantitative measure even by one lost life (e.g. 999 instead of 1000 deaths), it would have to be disregarded.[217] Second and relatedly, Gallagher argues that upholding a death toll in the thousands is low enough a quantitative threshold, so as not to overlook the threat of mass atrocity attacks against smaller groups, e.g. 'genocidal massacres', which is an issue with definitions that set too high a number (exemplified in Valentino's proposal for a threshold of 50, 000 deaths).[218] Third, Gallagher highlights that Pape's proposition of 'thousands have died and thousands more are likely to die' 'helps us demonstrate that a systematic ongoing murderous process is taking place'.[219] In this regard, Pape's death toll indicator reflects the argument made by Rosenberg and Strauss that every R2P situation requires both prospective and retrospective assessments, which in this case is the number of deaths that have already occurred and the number of deaths likely to occur. Relatedly, according to Gallagher, Pape's definition captures the spirit of the 'manifest failing' threshold, understood as a process rather than an end state, because it implies that there should be an ongoing process of mass killing at play with an expected escalation in death toll.[220] This 'means that we do not have to wait until tens of thousands of bodies pile up before a pillar three response can be considered (which is precisely what happened in the context of Qaddafi)'.[221]

Within the framework proposed in this thesis, death toll will be incorporated under the determinant of 'gravity', similarly to Rosenberg and Strauss second principle and the TMK, which attach death toll to their measure of 'severity'. In other words, it will be a part of an assessment of the magnitude and scope of human rights violations associated with atrocity crimes, which takes into account the present and projected number of victims of a range of relevant human rights violations or acts of violence. This understanding is preferred because, within the assessment of gravity the number of people

[217] Gallagher, 'Syria and the indicators of a "manifest failing"', p. 7. Benjamin A. Valentino, *Final Solutions, Mass Killing and Genocide in the 20th Century* (New York: Cornell University Press, 2004), pp. 11–12; Alex J. Bellamy, 'Mass Atrocities and Armed Conflict: Links, Distinctions, and Implications for the Responsibility to Prevent', Appendices (Muscatine, IA: The Stanley Foundation, February 2011), p. 2.

[218] Gallagher, 'Syria and the indicators of a "manifest failing"', p. 8.

[219] Ibid.

[220] Ibid.

[221] Ibid.

killed/murdered is on an equal footing with other human rights violations that have been found to be relevant in past cases of atrocities.

Relatedly, as Jones underscores:

> 'In fact, *one does not need to kill anyone at all to commit genocide!'*. Inflicting "serious bodily or mental harm" qualifies, as does preventing births or transferring children between groups. It is fair to say, however, that from a legal perspective, genocide unaccompanied by mass killing is rare, and has stood little chance of being prosecuted.'[222]

What is more, Butcher et al. point out that numerous instances which are customarily 'characterized as genocides, such as Bosnia (1992–1995) and Darfur (2003–2011), did not involve the large-scale killing seen in other prominent cases, like Cambodia (1975–1979) or Rwanda (1994), in which substantial portions of the targeted groups were actually eliminated'.[223] The authors also remark that relying on a high death toll count would mean excluding episodes of attempted genocide that were thwarted, such as East Timor, Cote D'Ivoire and Libya.[224] As already clarified, the threshold and evidentiary standard for a collective action response is lower than that of proving that any atrocity crimes have actually occurred under international law, precisely because the aim is to prevent the occurrence of such international crimes. Hence, while the significance of death toll for legal proceedings explains why this it has received considerable, if not disproportionate, amount of attention in scholarly debates on the topic of atrocity crimes, this should not be the case in debates over whether a national government is 'manifestly failing' to protect its populations from the four crimes.

This conclusion is reinforced when one consider the fact that the emphasis on death toll in some scholarly contributions to the study of atrocity crimes can also be attributed to misconceptions and ambiguities surrounding the definition of genocide. For instance, Jones points to Irving Horowitz' observation that genocide is sometimes viewed as a

[222] Adam Jones, *Genocide: A Comprehensive Introduction*. 2nd Edition (New York: Routledge, 2011), pp. 13-14.

[223] Butcher et al., 'Introducing the Targeted Mass Killing Data Set', p. 1529.

[224] Ibid.

'synonym' for mass murder[225] or a 'special case of mass murder'.[226] Relatedly, Jones highlights the ambiguity of the phrase 'serious bodily or mental harm' in the Genocide Convention. In an attempt to clarify its meaning, he cites examples of its interpretations in the practice of international courts and tribunals (e.g. ICTR), including: 'bodily or mental torture, inhuman treatment, and persecution', 'acts of rape and mutilation', the 'enslavement, starvation, deportation and persecution of [populations] their detention in ghettos, transit camps and concentration camps in conditions which were designed to cause their degradation, deprivation of their rights as human beings, and to . . . cause them inhumane suffering and torture'.[227]

In addition, he draws attention to other acts that have increasingly been related to genocide, namely '[m]easures to prevent births [including] forced sterilization and separation of the sexes', as well as '[s]exual trauma and impregnation through gang rape'.[228] For instance, first-hand accounts of the Chinese government's treatment of the Uighur Muslim minority obtained by the BBC have revealed that '[w]omen in China's "re-education" camps for Uighurs have been systematically raped, sexually abused, and tortured'.[229] Earlier reports suggest the government's attempts to supress the population of the ethnic Uighur community through imposing forced sterilisation and birth control measures on Uighurs women.[230] Furthermore, Jones also concurs with the position of Ratner and Abrams that despite the fact that the drafters of the Genocide Convention rejected a proposal to include acts of ethnic cleansing in the final text of the Convention, 'several sources correctly take the view that mass deportations under inhumane conditions may constitute genocide if accompanied by the requisite intent'.[231] For instance, the UN Commission of Experts for the former Yugoslavia suggested that certain acts carried out as part of the ethnic cleansing campaign in Yugoslavia 'could fall under'

---

[225] Adam Jones, *Genocide: A Comprehensive Introduction* (New York: Routledge, 2006), p. 20.

[226] Irving Horowitz, *Taking Lives: Genocide and State Power*. 5th Edition (New Brunswick, NJ: Transaction Publishers, 2002), p. 360.

[227] Jones, *Genocide* (2011), p. 14.

[228] Ibid.

[229] Matthew Hill, David Campanale and Joel Gunter, 'Their goal is to destroy everyone': Uighur camp detainees allege systematic rape', 2 February 2021. [Accessed 20 March 2021]. Available at: https://www.bbc.com/news/world-asia-china-55794071

[230] BBC news, 'China forcing birth control on Uighurs to suppress population, report says', 29 June 2020 [Accessed 20 March 2021] Available at: https://www.bbc.com/news/world-asia-china-53220713

[231] Steven R. Ratner and Jason S. Abrams, *Accountability for Human Rights Atrocities: Beyond the Nuremberg Legacy*, 2nd Edition (Oxford: Oxford University Press, 2001), pp. 30, 32. Cited in Jones, *Genocide* (2011), p. 14.

the Genocide Convention.[232] Specifically, 'murder, torture, arbitrary arrest and detention, extra-judicial executions, rape and sexual assaults, confinement of civilian populations in ghetto areas, forcible removal, displacement and deportation of civilian populations, deliberate military attacks or threats of attacks on civilians and civilian areas, and wanton destruction of property'.[233] Whereas their inclusion under Article II of the Genocide Convention 'serious bodily or mental harm' is debatable, such acts most certainly fall under the umbrella of mass atrocities that this thesis is concerned with.

This understanding is reflected in Rosenberg and Strauss' proposal for determining the gravity of human rights violations relevant to atrocity crimes by considering the number of 'victims of violence or level of irreparable harm'. As seen from their guiding principles for the application of R2P, killing and murder are only two of the human rights violations listed under their 'determination of relevant human rights violations' (Principle 1), the potential victims of which should be quantified to determine the gravity of the human rights violations (Principle 2). Put simply, death toll is just one indicator of the seriousness of human rights violations associated with atrocity crimes, alongside the number of people tortured, internally displaced, subjected to sexual violence, etc. As such, it is one of many quantitative indicators of the gravity of human rights violations that ought to be considered when determining whether the host state is 'manifestly failing'. This means that the fact that there are many reasons why death toll is weighted heavily in comparison with other indicators of atrocity crimes, does not mean that they are particularly good reasons to downplay other human rights violations linked to atrocity crimes or qualitative indicators that atrocities may be taking place. The latter is accounted for in Bellamy's definition of atrocity crimes, which marries a quantitative measure of 5, 000 deaths with behaviour-based indicators of intent, i.e. 'demonstrated evidence of deliberate civilian-targeting' and 'deaths caused by induced famine'.[234]

On balance, while placing emphasis on death toll makes sense in the context of Gallagher's discussion of the Syrian case, where the number of civilian deaths (estimated at '59,648 unique killings') was high enough to meet even Valentino's threshold of 'at least' 50, 000 deaths, making it one of the foremost indicators of the government's

[232] Ratner and Abrams, *Accountability for Human Rights Atrocities,* p. 32.

[233] Ibid.

[234] Bellamy, 'Mass Atrocities and Armed Conflicts', p. 18; Alex J. Bellamy, 'Mass Atrocities and Armed Conflict', Appendices, p. 2.

'manifest failing',[235] it is both counter-intuitive and counter-productive for the purposes of a guiding framework for the assessment of instances where national governments may be failing to protect their populations from mass atrocities on a case-by-case basis and so is specifying an arbitrary numeric threshold (be it in absolute numbers or relative to the populations' size). Similarly to Wheeler, this thesis upholds the view that 'It is no good trying to define [a humanitarian] emergency in terms of the numbers killed or displaced, because this is too arbitrary'.[236] This is not to say that death toll is not an important indicator of the gravity of human rights violations that should not be considered when determining whether the host state is fulfilling its R2P. However, whether it plays a role and how important a role it plays in a given case is contingent upon the specific situation and the presence and prominence of other indicators of the gravity of the situation as well as qualitative indicators of the gravity and imminence of mass atrocities, as the rest of this thesis and the outline of the 'manifest failing' framework in Chapter V will elaborate.

Displacement of people

Even though 'the R2P was not set up to address the problem of internally displaced people (IDP) and refugees', Gallagher suggests that it is another strong indicator that can help us identify a potential 'manifest failing' on part of the host state in cases of mass violence, because 'mass IDP and refugee movements help demonstrate that the government is failing in its internal responsibility to protect the safety and welfare of citizens as well as its external responsibility as refugees destabilise regional order'.[237] According to him, the 'manifestly failing' of the Syrian government in relation to this indicator was clear because it 'not only failed to create a safe security environment domestically but has committed war crimes and crimes against humanity which have played a major role in escalating the IDP and refugee crisis'.[238]

Jennifer Leaning's work on human security and complex humanitarian emergencies also highlights the link between the mass movement of populations, atrocities and the protection of populations. Leaning and Arie point out that 'forced flight from home and temporary settlement in refugee camps or as IDPs' threatens the physical safety and well-being of 'individuals and groups' as it exposes them to 'great risk of fear, hunger,

[235] Gallagher, 'Syria and the indicators of a "manifest failing"', p. 7.
[236] Wheeler, *Saving Strangers*, p. 34.
[237] Gallagher, 'Syria and the indicators of a "manifest failing"', p. 9.
[238] Ibid.

exhaustion, dehydration, disease, attack, robbery, torture, sexual assault, kidnap, and death'.[239] In this regard, they introduce the concept of 'social dislocation', which occurs when 'people [are forced] to leave established social ties and familiar territory or travel long distances, often to wind up at a destination but never then to truly settle down', as one of three negative measures, or inverse indicators, of human security.[240]

Elsewhere, Leaning observes that population dislocation in its various forms (e.g. 'forced migration, refugee settlement, internal displacement, and local entrapment') is a key characteristic of the greater scale and intensity that defines contemporary humanitarian emergencies, including those laced with atrocity crimes.[241] For instance, in Bosnia and Rwanda the mass movement of populations led to the disappearance of entire villages.[242] Leaning also notes that some commentators attribute the atrocities against civilian populations in the context of contemporary intra-state wars to a 'deliberate disregard for… international humanitarian law, in service of a strategy that privileges assault on civilian populations'.[243] This strategy could spring from 'the need to cause populations to flee and empty the land, because without supply lines and communication systems [non-state] armed groups cannot secure and hold a populated area; or the direct animus of ethnic cleansing, where the aim is not the capture of territory but the expulsion of stigmatized people'.[244] 'Forced population movement' and 'deportation' also feature in Rosenberg and Strauss' list of human rights violations associated with atrocity crimes, whereas forcible removal, displacement and deportation of civilian population are among the coercive practices used to remove populations as part of ethnic cleansing campaigns, as defined in Annex I.

As the above discussion indicates, there are a number of reasons why forms of population dislocation constitute human rights violations that are indicative of the presence of atrocities, which means they are to be included in the determination of 'relevant human

---

[239] Jennifer Leaning and Sam Arie, 'Human Security: A Framework for Assessment In Conflict and Transition', Harvard Center for Population and Development Studies, 11 (8) Working Paper Series (September 2001): 1-70, p. 18.

[240] Ibid. p. 19.

[241] Jennifer Leaning, 'Introduction', in Jennifer Leaning, Susan M. Briggs and Lincoln C. Chen (eds.) *Humanitarian Crises: The Medical and Public Health Response* (Cambridge, MA and London: Harvard University Press, 1999): 1-15, pp. 5-6.

[242] Ibid. p. 6.

[243] Jennifer Leaning, 'Human Security and Conflict', in Marci Green (ed.) *Risking Human Security: Attachment and Public Life* (Abingdon and New York: Routledge, 2008): 125-150, p. 134.

[244] Ibid.

rights violations / acts of atrocity crimes' (see Table 1.1.) of the framework to be proposed in this thesis. This also means that the number of individuals subjected to these forms of forced population movement will be taken into account as quantitative indicators in the determinations of the gravity of the threat of atrocity crimes.

Weapons used

Gallagher argues that paying attention to the weapons used to carry out atrocities is integral to 'manifest failing' assessments as it can help us understand a state's behaviour and the extent of their involvement in the associated violations of human rights taking place. As Gallagher puts it, '[q]uite simply, if government weaponry is being used in a systematic manner (rather than a rebel group capturing and using some government artillery) then this implies that the government is an active participant in the mass violence and is thus obviously failing in its R2P'.[245] For Gallagher, the use of heavy weapons in attacks on the city of Homs in February 2012, as well as in subsequent heavy fights in towns north of Homs, evidenced a marked escalation in violence and 'unacceptable levels' of violence.[246] What is more, evidence that opposition groups' use of heavy weapons was 'limited in "quality and quantity"' reaffirmed Gallagher's determination that 'the overwhelming responsibility' for the atrocities that were being committed was borne by the Syrian government, 'precisely because it was government weaponry that was used'.[247]

On a further note, the discussion around the reported prospective use of chemical weapons, in breach of the 1993 Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons, was a headline feature of the Syrian case. Notably, President Obama characterised establishing 'the use of chemical weapons [as] a game changer', which, as Gallagher remarks, 'led to a heated debate over whether a 'red-line' had been crossed'.[248] This shows that [d]espite the fact that an estimated 100,000 people had been killed by conventional weapons, chemical weapons were interpreted as a "qualitative leap" in the violence'. Put simply, Obama's statement regarding the potential use of chemical weapons in Syria not only 'draws attention to the central idea of threshold' with the use of chemical weapons identified as a definitive

[245] Gallagher, 'Syria and the indicators of a "manifest failing"', p. 10.
[246] Ibid.
[247] Ibid.
[248] Ibid.

'tipping point', but also encapsulates the qualitative dimensions of the 'weapons used' indicator and evidences that this indicator has been considered in determination of Syria's 'manifest failing'.[249] As Chapter IV clarifies, the qualitative dimension of the 'weapons used' indicator, specifically the use of heavy weapons, was also factored into assessments of 'manifest failing' in Libya by a number of states and UN officials. Likewise, the UN framework specifies '[r]esort to means of violence that are particularly harmful or prohibited under international law, including prohibited weapons, against a protected group' as well as 'methods or practices of violence that are particularly harmful against or that dehumanize a protected group, that reveal an intention to cause humiliation, fear or terror to fragment the group, or that reveal an intention to change its identity' as indicators suggestive of 'intent, by action or omission, to destroy all or part of a protected group'.[250] While, it is important to acknowledge the use of heavy and prohibited weapons or other particularly harmful methods or practices of violence identified by the UN as key for determining the role of the government in the atrocities taking place, they are not the only signs of government involvement.

For instance, Gallagher warns that focusing on the use of heavy weapons 'as a comparative indicator is problematic', despite its importance in the Syrian case, because, historically, such weaponry has not been an invariable feature of atrocity crimes. A notable testimony to the fact that atrocities can be committed without the use of advanced or heavy weapons is the 1994 Rwandan genocide, where the use of light weapons, viz. machetes and clubs, resulted in an estimated 57.7% of the total death toll.[251] What is more, Gallagher underscores that the 'Rwandan government played an integral role in "arming the country" as [it invested US$725.669 in] importing new machetes in 1993 so that "there was an estimated one new machete for every third male in the country"'.[252] According to Gallagher, this comes to show that it is 'not just the type but also quantity [of weapons that] helps us to understand what the government is doing and what it is not doing'.[253]

In sum, the weapons used in the course of a mass atrocity situation has considerable purchase as a key indicator of the host state's 'manifest failing', which can be associated

[249] Ibid.
[250] UN, 'Framework of Analysis for Atrocity Crimes', p. 19.
[251] Gallagher, 'Syria and the indicators of a "manifest failing"', p.11.
[252] Ibid.
[253] Ibid.

with the escalation of violence or the 'level of violence', as suggested by Gallagher. However, it is important to keep in mind that it is a multifaceted indicator that has both quantitative and qualitative dimensions. In addition to the type and number of weapons used, their origin could also be a tell-tale sign of state-sponsored atrocities in cases where the weapons used by militia or other non-state actors could be traced back to the state. Although information regarding the supply of weaponry in the latter case, will most likely be heavily guarded and difficult to obtain, where applicable, it ought to be considered in assessments of the host state's manifest failing, alongside the type and quantity of weapons used to inflict human rights violations associated with the four international crimes. Put simply, if a host state supplies weapons to a non-state actor committing atrocities against its populations, it is clear indication of that government's 'manifest failing', on the grounds of being complicit in the attacks against its populations. Given that the means of violence are key to establishing government intent, the framework proposed in this thesis will list the indicators of 'weapons used' highlighted in this section under the determination of host state intent (see Table 1.1 below). That said, like most indicators, even if there is nothing remarkable about the 'choice of weapons' or 'means of violence' to suggest that host state is falling short of its protection responsibilities, it could still be deemed 'manifestly failing' on account of the manifestation of indicators.

Targeting civilians: especially women, children and the elderly

One of these indicators, according to Gallagher, is the targeting of civilians, especially vulnerable groups – women, the elderly and especially children, which was a prominent aspect of the atrocities taking place in Syria.[254] The systematic targeting of civilians is symptomatic of either a policy formulated by the host state or of its inability to prevent non-state actors from executing this tactic.[255] In this regard, Gallagher argues that although the systematic targeting of civilians provides sufficient basis to make the case that the host state is 'manifestly failing', this is most apparent when 'the host state fails to protect [vulnerable] groups, or if it is the perpetrator itself'.[256] In making this argument, Gallagher neither attempts to 'downplay attacks on other civilians', nor to imply that women, children and the elderly are 'passive actors' devoid of agency.[257] Instead he seeks

[254] For concrete evidence and examples of the targeting of vulnerable groups in Syria see: Ibid., pp. 11-12.

[255] Ibid., p. 11.

[256] Ibid.

[257] Ibid.

'to highlight that certain groups are more vulnerable in war than others' and 'that the gravity and character of the crime takes on another qualitative dimension when' these groups, especially children, are targeted.[258] Therefore, '[a]lthough the civilian focus is enough to warrant a 'manifest failing'…we should pay attention to the targeting of vulnerable groups when assessing the (ir)responsibility of the host state.[259] Moving forward, the targeting of civilians and vulnerable groups will be subsumed under the assessment of the gravity of relevant human rights violations, alongside the targeting of groups based on their identity, found in Rosenberg and Strauss' understanding of the gravity dimension.

Overall, Gallagher's indicators reinforce the argument that the mere presence of atrocities does not automatically render host state 'manifest failing'. As Bode suggests, 'lists such as [the list of 'manifest failing' indicators offered by Gallagher] are clearly a step forward in clarifying assessments on "manifestly failing" and therefore also increasing the accountability of decision-makers'.[260] Another distinct strength of Gallagher's contribution is the detailed discussion of each of the proposed indicators and their manifestation in the Syrian context, which not only explains why and how certain aspects of the situation are indicative of the host state's manifest failing to but provides guidance as to how they ought to be interpreted in practice. Yet, Gallagher himself recognises that these are the 'first tentative steps towards establishing a specific list of criteria',[261] which 'provide a framework to aid further debate and analysis', and 'hope[s]… that through further research a more comprehensive framework can be established which aids the assessment of other case studies'.[262] Specifically, in order to develop a comprehensive framework that could aid 'manifest failing' determinations in a wide range of cases, other determinants of host state conduct need to be explored, in a bid to cover cases where the role of the host state in the planning or commission of atrocities cannot be established directly through indicators such as intent and the use of government weaponry.

[258] Ibid. p.12.

[259] Ibid.

[260] Bode, '"Manifestly Failing" and "Unwilling or Unable"', p. 170.

[261] Gallagher, 'What constitutes a "Manifest Failing"?', p. 439.

[262] Gallagher, 'Syria and the indicators of a "manifest failing"', p. 13.

IV. THE STUDY OF 'MANIFEST FAILING' IN THE LIBYAN CONTEXT

Given that Chapter IV of this thesis will focus on the international intervention in Libya in a bid to derive determinants of 'manifest failing', the present review of relevant contributions will be incomplete without the inclusion of Bode's critical assessment of the 'unable or unwilling' and 'manifest failing' 'intervention formulas', in which she sets out to answer some challenging questions with regards to the threshold for collective action through an analysis of the 'manifest failing' of the Libyan authorities in the context of the 2011 atrocity crisis in the country. Specifically, she promises to showcase 'how the Security Council has determined a state to be "manifestly failing" in practice'.[263] Bode starts with a summary of the events that preceded the intervention in Libya, highlighting the following: 1) the commencement of violence against civilians in February 2011; 2) the international condemnation that followed and the adoption of Resolution 1970; 3) the subsequent continuation of Qaddafi's offensive and instances of hate speech, threatening innocent populations with further atrocities; 4) the response of the international community with the adoption of Resolution 1973, imposing a no-fly zone, and; 5) the launch of NATO-led operation Unified Protector after Qaddafi failed to meet the demands of the international community.[264] '[B]ased on this brief summary of events [, Bode argues], the Security Council clearly had determined the Libyan government's "manifest failure" to protect its own population'.[265]

In her defence of this claim, Bode attributes great importance to one specific event – the adoption of Resolution 1970, which condemned the violence against civilian populations and imposed sanction on the Libyan regime.[266] In particular, she draws attention to two statements of the resolution's text to substantiate her argument. First, the mention of the responsibility of the Libyan state to protect its populations.[267] Second, the affirmation that the violence against civilians taking place in Libya could amount to crimes against humanity, reinforced by the inclusion of a referral of the situation to the International Criminal Court (ICC) dealing exclusively with the investigation and prosecution of atrocity crimes.[268] However, these provisions of Resolution 1970, individually or

[263] Bode, '"Manifestly Failing" and "Unwilling or Unable"', p. 168.
[264] Ibid. pp.168-169.
[265] Ibid. p.169.
[266] UNSC, S/RES/1970, 26 February 2011.
[267] Bode, '"Manifestly Failing" and "Unwilling or Unable"', p. 169.
[268] Ibid.

together, do not provide sufficient evidence to suggest that the UN Security Council had deemed Libya to be 'manifestly failing' at the time. Whilst the resolution invoked the R2P concept, it only referred to Libya's primary responsibility to protect its populations, but not to the international community's subsidiary protection responsibility involving collective action, to which the 'manifest failing' threshold attaches. As already discussed, whilst evidence that one of the four crimes may be taking place is one of the requisite preconditions for a situation to fall under the remit of the R2P, this alone is insufficient to conclude that a state is 'manifestly failing' to protect its populations. In sum, Bode's analysis of the text of Resolution 1970 does not explain or provide evidence how or why UN Security Council member states inferred that the Libyan authorities were 'manifestly failing'.

The author proceeds by claiming that the above 'type of reasoning', she believes to evidence the UN Security Council's conviction that Libya was 'manifestly failing', also finds support in Security Council meeting records preceding Resolution 1970. To be precise, there was only one such publicly available meeting that took place the day before the resolution was adopted and saw no contributions from UN Security Council members.[269] And in this meeting, there was only one contribution made by a state representative – that of the Libyan permanent representative Mr. Shalgham, who made an emotive appeal to the Council to take forceful action against his own government ("Please, United Nations, save Libya. No to bloodshed. No to the killing of innocents").[270] As strong as the statement of the Libyan ambassador is and the symbolic value of his resignation that it gave effect to for highlighting the gravity of the situation in Libya,[271] it does no substantiate Bode's claim that the UN Secretary General 'and a range of member states kept alluding to the "manifest failure" of Libya and what this meant for the responsibility of the international community' nor the broader argument that 'determining Libya's manifest failure had already started in the context of Resolution 1970'.[272] Therefore, it remains unclear how the meeting preceding the adoption of Resolution 1970 provides any evidence that the UN Security Council had determined Libya's 'manifest

[269] UNSC, S/PV.6490, 25 February 2011.

[270] Ibid., p. 5.

[271] Susan Rice, *Tough Love: My Story of the Things Worth Fighting For* (New York: Simon and Schuster, 2019), p. 280.

[272] Bode, '"Manifest Failing" and "Unwilling or Unable"', p. 169.

failing' by the time the first resolution on Libya was passed. Bode's cursory analysis of the events that followed does not shed more light on this matter.

On the basis of her discussion of events and Security Council meetings in the run-up to the first two resolutions on the situation in Libya, Bode concludes that although it would seem that the UNSC 'followed a dual approach' to determine the Libyan government's 'manifest failing', it 'was mostly already judged at the time of SCR 1970'.[273] This understanding, which attributes a greater importance to the events surrounding Resolution 1970 for determining Libya's 'manifest failing' in comparison with the developments that led to the adoption of Resolution 1973, is problematic because it fails to recognise that 'manifest failing' is an ongoing process. Whilst Libya might have already been deemed to be 'manifestly failing' in the context of Resolution 1970 to an extent, the international community's response to the events that took place between the two resolutions indicated that their gradually increasing confidence in Libya's failing had a bearing on the adoption of stricter and more intrusive measures under Resolution 1973 (as Chapter IV of this thesis will evidence). What is more, R2P scholars have stressed that the events preceding the vote on Resolution 1973 were decisive in forging a consensus on the authorisation of 'all necessary measures' (the standard language employed by the UNSC to authorise the use of force).[274] Therefore, Bode's emphasis on the significance of the events surrounding the adoption of Resolution 1970 for judging Libya's 'manifest failing' and the concomitant downplaying of the context in which Resolution 1973 was adopted are misleading.

Perhaps the most obvious sign of Bode's misunderstanding as to what the 'manifest failing' threshold entails is the interchangeable use of the terms 'manifestly failing' and 'manifestly failed' in her chapter. Judging by the title of the chapter section dedicated to the Libyan crisis 'The Libyan Intervention and the Criteria of Manifestly Failing', it appears that the author is aware of the wording of Paragraph 139 of the WSOD. However, the subsequent references to 'manifest failure' for the rest of the section and in every other instance in Bode's chapter where she is not directly citing the Outcome Document are indicative of a limited understanding of R2P's threshold for coercive action, as envisioned by the drafters of the WSOD. At this point, it is clear that the author not only employs the

[273] Ibid., p. 170.

[274] Bellamy and Williams, 'The new politics of protection?', p. 841; Luke Glanville, 'Intervention in Libya: From Sovereign Consent to Regional Consent', 14 *International Studies Perspectives* (2013): 325-342, p. 334.

phrase incorrectly, but also that she does not see the 'manifestly failing' threshold for what it really is – a process.

Whereas Bode's initiative to discuss the 'manifest failing' threshold in the context of Libya for the first time is laudable, the arguments she makes and conclusions she reaches are questionable for several reasons. First, it is her limited and flawed understanding of the threshold at the heart of R2P's reactive dimension that renders her account unconvincing and leads her to some problematic conclusions. Second, Bode does not provide sufficient or appropriate evidence to substantiate her claims. Her brief account of Libya neither provides sufficient textual evidence from UN Resolutions and meetings, nor from secondary sources. This systematic lack of support to back up her arguments renders them feeble, at best. Third, Bode's presentation of key events is not accompanied by a robust analysis thereof to explain how these events evidenced that the UN Security Council had determined the incidence of a 'manifestly failing' in Libya. On account of this lack of sufficient argumentation and analysis in support of her assertions, the author fails to ground any of the claims she makes with regards to the 'manifest failing' threshold in the context of the Libyan crisis. This is not surprising, given that Bode dedicates just under 800 words to her investigation of the difficult questions surrounding of how the UN Security Council arguably made the determination that the Libyan state was 'manifest failing', half of which comprise a summary of events and a concluding paragraph highlighting the need to develop indicators of 'manifest failing'. This is not to say that there is no truth to Bode's suggestions. However, it means that to either support or reject them would require solid understanding of the R2P and the 'manifest failing' threshold, robust analysis and a deeper examination of the available textual evidence (including the records of UN Security Council meetings on Libya), which will be offered in Chapter IV of this thesis.

In sum, the limitations of Bode's account are too many and too serious to overlook. The problem of unreliability aside, the author does not make any suggestions to address the substantive indeterminacy of the 'manifest failing' test. This is not surprising, given that her focus is by no means on shedding light on the substance or specific indicators of 'manifest failing'. As a result, the relevance and utility of Bode's analysis of Libya's 'manifest failing' to this thesis is very limited. For these reasons, it will not provide the basis for the analysis of the Libyan government's 'manifest failing' in Chapter IV of this thesis nor underwrite the framework for determining 'manifest failing' in this thesis.

CONCLUSION

This chapter analysed various contributions to the assessment of mass atrocity crimes and 'manifest failing' in order to probe their utility for improving our understanding of the substantive threshold for collective action and for providing effective guidance to those who wish to determine whether a host state is 'manifestly failing' in its responsibility to protect. Some were found to be more relevant in this regard (the contributions by Gallagher, Rosenberg and Strauss) than others (Bode's analysis of Libya's 'manifest failing').

The propositions made by Rosenberg, Strauss and Gallagher point to different elements that the assessment of a 'manifest failing' ought to encompass and constitute important first steps in specifying what it means for a host state to be 'manifestly failing'. Rosenberg and Strauss outline broader determinants, accompanied by a list of their concrete manifestations, that could assist relevant stakeholders in judging whether the 'manifest failing' threshold has been crossed, namely relevant human rights violations, their gravity, and measures taken by the host state to mitigate the risk of atrocities. Gallagher adds to the study of the 'manifest failing' test by grounding this abstract idea of a threshold in specific indicators relevant to determining the role of the state in the planning and commission of atrocities and their real-life manifestation in the context of the Syrian crisis. Building on the work of these scholars by engaging with the literature on mass atrocities, the definitions of atrocity crimes, the UN 'Framework of Analysis for Atrocity Crimes' and the conceptual underpinnings of mass atrocity data sets, this chapter formulated a continuum of criteria relevant to the determination of a 'manifest failing' (see Table 1.1). These determinants and indicators constitute the first step towards putting together a fully-fledged framework for 'manifest failing' inquiries, to be presented in Chapter V of this thesis.

However, further research is needed in order to arrive at a core set of determinants of 'manifest failing' that will assist the analysis of a wider range of cases. Specifically, more work needs to be done to carefully consider 'manifest failing' scenarios, in which there are reasons to doubt that the host state possesses the capacity to address the threat of atrocities posed by a non-state group. Likewise, further work is required in order to explore aspects of host state conduct, other than the role of the national government in the planning and commission of atrocities, which constitute evidence that the host state

is ‘manifestly failing’ to protect its populations. These lacunae will be addressed in Chapter III of this thesis, which draws inspiration from the study of the ‘unwilling or unable’ test in the context of self-defence against non-state actors. In addition, the study of the ‘manifest failing’ threshold will benefit from exploring the way in which states understand the threshold for collective action, which can lend important insights for building a ‘manifest failing’ framework that aligns with the factors that decision-makers see as vital to such assessments (as explained in Chapter II). Such an investigation will be conducted in Chapter IV, which will identify the criteria that states associated with the threshold for collective action in their justifications for the intervention in Libya.

Table 1.1. ‘Manifest Failing’ Determinants and Indicators Derived from Contributions to the Study of ‘Manifest Failing’ and Mass Atrocities

DETERMINANTS	INDICATORS
Relevant human rights violations / acts of atrocity crimes	Killing, murder, deaths caused by induced famine, deaths caused by indiscriminate targeting, torture, inhuman treatment, mutilation, rape and sexual violence, enslavement, abduction, or otherwise deliberately causing great suffering, physical or mental harm, extra-judicial executions, arbitrary arrest and detention, deliberate military attacks or threats of attacks on civilians and civilian areas, expropriation, wanton destruction of property, looting, destruction of subsistence food supply, denial of water or medical attention, destruction or denial of medical supplies, man-made famine, redirection of aid supplies, forced population movement (e.g. forcible removal, displacement and deportation of civilian populations), and restricted movement, (e.g. confinement of civilian populations in ghetto areas), discrimination in access to work and resources, discrimination in education and lack of access to justice and redress, lack of freedom of speech/ press/ assembly/ religion, political marginalization.
The gravity of the human rights violations taking place	*Magnitude and scope of relevant human rights violations*: - the number of victims of violence or level of irreparable harm inflicted by the human rights violations specified under Determinant 1 (number of dead, internally displaced, severely injured, tortured, victims of sexual violence) - the portion of the victim populations/group affected by the violence

	The targeting of identifiable groups, where: victims are identified on the basis of identity criteria related to national and ethnic origin, religion, gender, race, colour, descent, place of birth, sexual orientation and other identity markers, or on grounds linked to their association with a political opinion or group, or any other group or collective as identified by the perpetrator. *The targeting of civilians and especially vulnerable groups*, such as women, children and the elderly.
Measures taken by the host state to mitigate the risk of atrocities	Public acknowledgement and condemnation of human rights violations; clear and public orders to military, police or security forces to respect international human rights and humanitarian law; immediate enforcement of accountability for the most relevant violations; ensuring humanitarian assistance and protection for victims of violence; and, in cooperation with relevant stakeholders, including potential victims and drawing an action plan with timelines for mitigating the most urgent risk factors.
Host state intent	*Public statements of intent:* the examination of primary and secondary sources (including political manifests, media records, official documents) for types of hate speech known to precede or accompany episodes of mass atrocities, including statements that atrocity-related violence was directed towards a specific group, or statements of intent to destroy a specific group or intimidate a group by creating the perception of an existential threat or incitement to violence. *Systematic preparation*: the selective killing of members of a group with the intent of communicating to other members that their survival may be threatened, the intentional targeting of vulnerable groups (including women, children and the elderly), impunity for or tolerance of acts of atrocities or associated serious human rights violations, evidence of systematic organisation, logistical or political preparation to commit or facilitate large-scale or systematic gross human rights violations, the pattern of refugee origin and internal displacement, systematic use of sexual violence against a population, measures to prevent births including forced sterilization and separation of the sexes, sexual trauma and impregnation through gang rape, humiliation of a group publicly or in the media, denial of the existence of recognised group or parts of their identity, exclusionary ideologies that purport to justify discrimination,

	destruction of cultural symbols, the separation of people on the basis of identity, clear territorial control by the perpetrator group in the affected area, support to or failure to condemn the actions of non-state groups accused of committing gross human rights violations / acts of violence associated with atrocities or failure to condemn their actions; the systematic use of government weaponry against civilian populations, the use of heavy weapons, the quantity of weapons (heavy or light) used, the use of weapons prohibited under international law, such as chemical and biological weapons, indiscriminate weapons, or weapons, methods and materials of warfare causing unnecessary harm or suffering, poisonous or poisoned weapons, the proliferation of all categories of weapons, including the illicit proliferation of small arms and light weapons, the recruitment of mercenaries (by the host state) charged with the task to intimidate or attack civilian populations.

CHAPTER II: RESEARCH STRATEGY

Having underscored the need to establish a framework for 'manifest failing' inquiries that offers relevant stakeholders a common script against which to assess the facts of a given situation, this chapter will develop a methodology to derive an independent set of factors (determinants and indicators) that will provide the skeleton for this normative framework. To respond to the challenge of developing a 'manifest failing' framework, which can serve as tool for the analysis of situations at risk of atrocities in practice, this thesis subscribes to what Sill and Katzenstein call 'analytic eclecticism' – a pragmatic research style which supports intellectual endeavours 'to complement, engage, and selectively utilize' concepts, theories, and methods embedded in a variety of 'research traditions to build complex arguments that bear on substantive problems of interest to both scholars and practitioners'.[275]

This approach is particularly well-suited to the task of building a framework for judging a 'manifest failing', because '[i]t is intended to generate diverse and flexible frameworks, each organized around a concrete problem, with the understanding that… it is the problem that drives the construction of the framework'.[276] Adopted by those 'motivated by problem-driven, rather than paradigm- or method-driven research',[277] analytic eclecticism is:

> 'consistent with an ethos of pragmatism in seeking engagement with the world of policy and practice, downplaying unresolvable metaphysical divides and presumptions of incommensurability and encouraging a conception of inquiry marked by practical engagement, inclusive dialogue, and a spirit of fallibilism'.[278]

As such, it is defined by a willingness to 'step beyond conventional disciplinary boundaries' in order to:

[275] Rudra Sill and Peter Katzenstein, 'Analytic Eclecticism in the Study of World Politics: Reconfiguring Problems and Mechanisms across Research Traditions', 8 (2*) Perspectives on Politics* (2010): 411-431, p. 411; Benjamin J. Cohen, 'The Way Forward', 14 (3) *New Political Economy* (2009): 395-400, p. 400.
[276] Sill and Peter Katzenstein, 'Analytic Eclecticism in the Study of World Politics', p. 411.
[277] Catherine Weaver, 'IPE's Split Brain', 12 (3) *New Political Economy* (2009): 337-346, p. 344.
[278] Sill and Peter Katzenstein, 'Analytic Eclecticism in the Study of World Politics', p. 411.

> ‘make better use of the innovative and creative scholarship produced within [different research] traditions in the process of recognising socially important problems and building interpretations… that… can be analytically coherent, intellectual interesting, and responsive to normative concerns and policy debates surrounding these problems’.[279]

The flexibility afforded by analytic eclecticism supports this thesis in crossing disciplinary boundaries, in order to make use of the contributions of International Relations scholars to the study of the ‘manifest failing’ test, gain insights from the factors states tend to offer when making the case that a state is ‘manifestly failing’ in its primary responsibility to protect its populations, and draw on International Law scholarship on the ‘unwilling or unable’ test, in a bid to develop a comprehensive framework for ‘manifest failing’ inquiries. In so doing, this thesis engages with legal scholarship on the ‘unable or unwilling’ test in the context of self-defence against non-state actors and debates as to whether it constitutes a legal norm. It also draws extensively on the constructivist understanding of social norms, justificatory discourses, and Quentin Skinner’s thesis that the behaviours and actions of states are constrained by the principles they profess.

Analytic eclecticism’s licence for combining different approaches also offers the benefit of enabling methodological triangulation. In particular, it allows for the operationalisation of a qualitative research design, which combines the legal method of reasoning by analogy to draw inspiration from the factors of the ‘unwilling or unable test’ with a detailed case study analysis of the 2011 Libyan intervention looking at the high-level justificatory discourses surrounding the threshold for collective action. The remainder of this chapter proceeds in three sections, which elaborate on aspects of the rationale, concrete research steps and accompanying methods underpinning this research design.

TRIANGULATION

There are two key challenges to devising a comprehensive ‘manifest failing’ framework. The first is the scarcity of cases where states have explicitly framed their collective response to atrocity crises in R2P-terms with reference to R2P’s reactive dimension and the ‘manifest failing’ threshold, and not just the primary responsibility of individual states to protect their citizens. The second difficulty, as illustrated in Chapter I, arises from the

[279] Weaver, ‘IPE’s Split Brain’, pp. 344-345.

paucity of research that makes a meaningful contribution to our understanding of the 'manifest failing' threshold, which means there is little existing knowledge to draw on. Both can, to some extent, be attributed to the circumstance that it was not until the 2005 World Summit that the 'manifest failing' conception was coined and the R2P principle officially recognised by world leaders. Arguably, this is not a long lifespan for the development of an emerging principle that intersects areas of international law as contentious and resilient as the use of force and deals with the most serious international crimes of all. Because of the relative novelty and constant evolution of the R2P, an approach sensitive to change over time was required. All of this meant that collecting suitable data for this research was no easy task. To deal with these challenges this thesis responds by implementing a triangulation research strategy.

Broadly speaking, the concept of triangulation as used in the social sciences 'refers to using multiple, different approaches to generate better understanding of a given theory or phenomenon'.[280] Its effectiveness was originally premised upon the argument that the weaknesses of one method will be neutralized and compensated for by exploiting 'the counter-balancing strengths of another'.[281] Having said that, investigators diverge in the emphasis they place on the purposes and uses of this approach. Those with a positivist and post-positivist disposition, view triangulation as vital to 'overcoming the limitations inherent in using only one approach to research' and as a means to 'establish corroborating evidence' and verify findings.[282] The idea of triangulation as a strategy of validation, 'which assumes one reality and one conception of the subject under study independent of the special methodical approach'[283] can be traced back to Denzin's oft-cited conception of this approach as 'the combination of methodologies in the study of the same phenomenon'.[284] The notion of validity as the main drive for employing triangulation has been subjected to much criticism from fellow qualitative researchers,[285]

---

[280] Scott F. Turner, Laura B. Cardinal, and Richard M. Burton, 'Research Design for Mixed Methods: A Triangulation-based Framework and Roadmap', 20 (2) *Organizational Research Methods* (2017): 243-267, p. 243.

[281] Todd D. Jick, 'Mixing Qualitative and Quantitative Methods: Triangulation in Action', 24 (4) *Administrative Science Quarterly. Qualitative Methodology* (1979): 602-611, p. 604.

[282] Sarah L. Hastings, 'Triangulation', in Neil J. Salkind (ed.) *Encyclopedia of Research Design. Vol I.* (Thousand Oaks, CA: SAGE, 2010): 1537-1540, pp. 1537-38.

[283] Uwe Flick, 'Triangulation Revisited – Strategy of Validation or Alternative', 22 (2) *Journal for the Theory of Social Behaviour* (1992): 175-197, p. 178.

[284] Norman Denzin, *The Research Act: A Theoretical Introduction to Sociological Methods* (New York: McGraw-Hill, 1978), p. 291.

[285] Flick, 'Triangulation Revisited', p. 178.

and was later abandoned by Denzin himself in favour of a more interpretivist research stance that establishes 'in-depth understanding, not validity' as the main purpose for the use of triangulation.[286] In line with this revised conceptualisation, those who view reality as socially constructed place emphasis on triangulation's 'potential to provide multiple lines of sight and multiple contexts to enrich the understanding of a research question'.[287]

This thesis relies precisely on the capacity of triangulation for obtaining multiple perspectives on the phenomenon of interest – 'manifest failing' – and in doing so to expand and deepen our understanding of it in a bid to produce a normative framework to guide its consideration. This involves employing the triangulation method as a tool to 'ad[d] range and depth to our understanding of the complexity of the phenomenon' under study,[288] to 'ad[d] breadth or depth to our analysis but not for the purpose of pursuing "objective" truth'.[289] As Kockeis-Stangl suggests, this approach is best understood as 'perspective triangulation' and must involve a preparedness to arrive at an outcome that is 'no uniform picture but rather one of a kaleidoscopic kind'.[290] This is a reasonable expectation, given that aside from convergence, triangulation can also result in inconsistency and contradiction.[291] This means that triangulation may not necessarily lead to a decrease in bias and an increase in validity. However, in the search for additional perspectives rather than validation of one single meaning,[292] conflicting information is not detrimental but can be constructive in helping render a more holistic picture of the phenomenon under study.[293] It can also lead to a more nuanced understanding, namely,

[286] Ibid. p. 179.

[287] Hastings, 'Triangulation', p. 1537.

[288] Flick, 'Triangulation Revisited', p. 189.

[289] Nigel G. Fielding & Jane L. Fielding, *Linking Data* (Beverley Hills: SAGE, 1986), p. 33.

[290] Eva *Köckeis-Stangl*, 'Methoden der Sozialisationsforschung'. In Hurrelmann, K. and Ulich, D. (eds.) *Handbuch der Sozialisationsforschung*. (Weinheim: Beltz, 1982): 321-370, p. 363. Quoted in Flick, 'Triangulation Revisited', p. 179.

[291] Matthew B. Miles, and A. Michael Huberman, *Qualitative Data Analysis*. 2nd Edition. (Thousand Oaks, CA: SAGE, 1994), p. 267.; Christian Erzberger and Gerald Prein, 'Triangulation: Validity and empirically based hypothesis construction. Quality and Quantity', 31 (2) *Quality and Quantity* (1997): 141-154, p. 147.

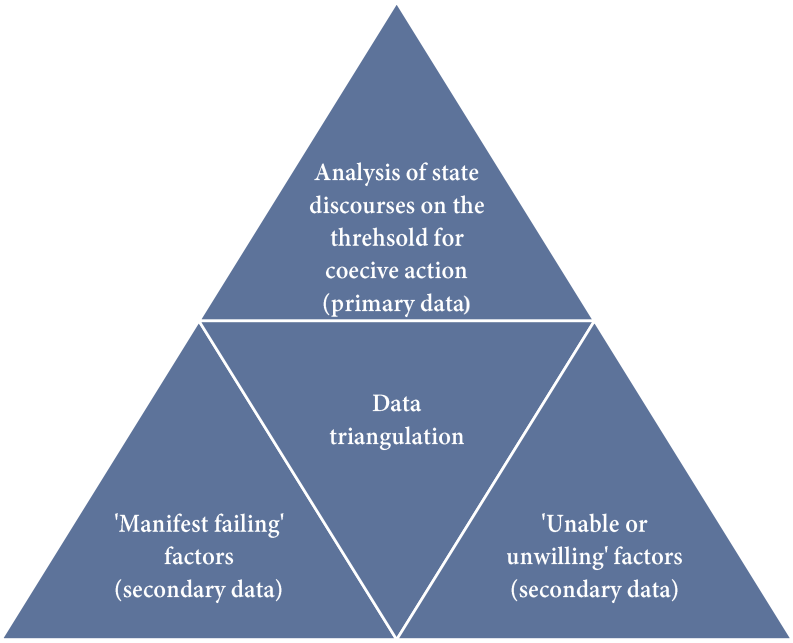
[292] Robert E. Stake, *The Art of Case Study Research* (Thousand Oaks, CA: SAGE, 1995), p. 114.

[293] Rogerio M. Pinto, 'Mixed Methods Design', in Neil J. Salkind (ed.) *Encyclopedia of Research Design. Vol I.* (Thousand Oaks, CA: SAGE, 2010): 812-818, pp. 816-17.

'a deeper appreciation for the multidimensionality of the research question and point to new avenues of inquiry'.[294]

In order to develop a 'manifest failing' framework informed by a variety of perspectives, this thesis relies on the triangulation of several different types and sources of data, also known as *data triangulation*.[295] The following triangulation approach couples primary data collected through documentary analysis of state discourses on the 'manifest failing' threshold with secondary data from existing publications on the 'manifest failing' threshold and the 'unable or unwilling' test (see Figure 1). Specifically, it combines three sources of qualitative data: 1) primary data on the factors relevant to 'manifest failing' determinations collected through an analysis of high-level political discourses on collective action in response to mass atrocities; 2) published secondary-source information on the factors relevant to the 'manifest failing' test (explored in Chapter I), and; 3) secondary data on 'unable or unwilling' factors obtained from publications on the 'unable or unwilling' threshold for self-defence against non-state actors.

Figure 1. Data Triangulation



[294] Hastings, 'Triangulation', p. 1539.; Miles and Huberman, *Qualitative Data Analysis*, p. 267.; Gretchen B. Rossman, and Bruce L. Wilson, 'Numbers and Words: Combining Quantitative and Qualitative Methods in a Single Large-Scale Evaluation Study', 9 (5) *Evaluation Review* (1985): 627-643, p. 637.

[295] Denzin, *The Research Act*, p. 295. Denzin distinguishes between four basic types of triangulation: *data triangulation, theory triangulation, investigator triangulation* and *methodological triangulation*.: Ibid. pp. 294-5.

When adopting a research strategy that involves integrating primary data and information from secondary sources, Davies advises researchers to treat the latter with caution, particularly because the published literature does not always rely on a robust methodology for the collection, interpretation and assessment of data.[296] Taking due precautions, the previous chapter engaged with the methodology of publications on the 'manifest failing' threshold in order to ascertain whether they were sufficiently reliable to inform the 'manifest failing' framework to be developed in this thesis. Where applicable, the next chapter also assesses whether there are any methodological issues that would preclude the consideration of proposals for the factors that should inform 'unable or unwilling' determinations in the context of the extraterritorial self-defence against non-state actors.

Integrating themes from the analysis of documents (scanned for 'manifest failing' factors) with the analysis of publications on the 'manifest failing' threshold and research on the factors of the 'unable or unwilling' test will facilitate the identification of overriding metathemes that cut across contexts and potentially raise confidence in the transfer (and interpretation) of key findings between analogous contexts. This will enable bringing together multiple perspective on the research questions posed in this thesis and tapping into different aspects of the issue at hand, thereby contributing to a more complete and multidimensional understanding of what the 'manifest failing' threshold should entail substantively.

Triangulation protocol

Generating a unified set of findings from multiple sources of data is often an arduous task, as researchers face the challenge of interpreting a multitude of possible outcomes ranging from convergence to dissonance. Most triangulation studies have attracted criticism for relying on the so-called 'intuitive approach, in which a researcher intuitively relates information obtained from different instruments to each other' without outlining how they went about the process of employing triangulation.[297] To ensure procedural transparency

[296] Phil H.J. Davies, 'Spies as Informants: Triangulation and the Interpretation of Elite Interview Data in the Study of the Intelligence and Security Services', 21 (1) *Politics* (2001): 73-80, p. 78.

[297] Tracy Farmer, Kerry Robinson, Susan J. Elliott, and John Eyles, 'Developing and Implementing a Triangulation Protocol for Qualitative Health Research', 16 (3) *Qualitative Health Research* (2006): 377-394, pp. 377, 379.

and validity, the steps taken in the triangulation process followed here will be detailed below.[298]

To arrive at an integrated set of determinants and indicators I will follow a four-step triangulation protocol, inspired by Farmer et al.'s oft-cited six-stage procedural triangulation approach.[299] It comprises four steps: *categorising information*, *convergence assessment*, *global assessment*, and *completeness assessment* (see Table 1) and draws on three sources of qualitative data (See Figure 1). The first step, *categorising information*, involves reviewing the findings from each data source in order to sort them into 'similarly categorised segments that address the research question(s)'.[300] The second step, *convergence assessment*, compares the three sets of findings, namely themes, determinants and indicators, to determine overlap (or lack thereof) in presence, and where possible meaning and prominence. The type and degree of convergence of each compared segment are categorised in accordance with three characterisations: overlap (a theme, indicator, or determinant is present in more than one data set), omission (presence in one set of findings, but not the others) and dissonance (where disagreement over the essence of meaning or prominence of a theme, indicator or determinant can be observed). These activities are followed by the *global assessment* stage – the evaluation of 'all compared segments to provide a global assessment on the level of convergence'.[301] Finally, the *completeness assessment* compares the scope and nature of the unique themes, indicators and determinants, for each method or data source 'to enhance the completeness of the united set of findings and identify key differences in scope and/or coverage'.[302] Determining the complementarity of different data sources in this way can help 'expose multiple dimensions of the same research issue and thereby increase our level of understanding'[303] by captur[ing] a more complete, holistic, and contextual portrayal of the units under study'.[304]

[298] This is known as the 'procedural approach' to triangulation: Ibid. p. 379.

[299] Ibid. pp. 381-389.

[300] Ibid. p. 383.

[301] Ibid.

[302] Ibid. p. 383.

[303] Ibid. pp. 378-379.

[304] Jick, 'Mixing Qualitative and Quantitative Methods: Triangulation in Action', p. 603.

Table 2.1. Triangulation Protocol

Step	Activity
1. Categorising information	Identify key themes, indicators and determinants of 'manifest failing' from each data source to compile a unified list of themes, indicators and determinants to be compared for presence and meaning in the next stage.
2. Convergence assessment	Identify areas of overlap and divergence of the segments mentioned in the first step (themes, indicators and determinants, relevant to making 'unable or unwilling'/'manifest failing' determinations) and characterising them in accordance with three categories – overlap, omission and dissonance.
3. Global assessment	Evaluate all compared segments to reflect on the overall convergence of themes, indicators and determinants.
4. Completeness assessment	Reflect on the scope and nature of unique themes, indicators and determinants drawn from each method/data source.

EXTENDING THE 'UNABLE OR UNWILLING' TEST

As indicated above, this thesis draws on the substantive dimension of the 'unable or unwilling' test in the context of the extraterritorial use of force against non-state actors to identify indicators and determinants that are applicable in the context of the R2P. The rationale behind this move is multifaced. First, the international political discourse over the threshold for international action in response to atrocities is marked by the continuous use of the 'unable or unwilling' language instead of, or in conjunction with, the 'manifest failing' terminology.[305] The latter formulation was embedded in the original 'responsibility to protect' concept, introduced by the International Commission on State Sovereignty and Intervention (ICISS) in its 2001 'The Responsibility to Protect' report to specify the threshold for international action to protect civilians,[306] and featured in earlier

[305] For examples, see Gallagher, 'What constitutes a "Manifest Failing"?', pp. 433-435.
[306] ICISS, 'The Responsibility to Protect', para. 4.1.

drafts of the 2005 Outcome Document.[307] In other words, the 'unable or unwilling' threshold is not only the immediate predecessor of the 'manifest failing' test, but also coexists with it, by virtue of the interchangeable use of the two concepts as though they mean the same thing.[308] Hence, given the substantive indeterminacy of the 'manifest failing' threshold, ICISS' 'unable or unwilling' conception offers an alternative gateway into learning more about the substance of the threshold for collective action in mass atrocity situations. Second, as the next chapter elaborates, even though the ICISS report provides little detail as to the substance of the 'unable or unwilling' threshold, my personal correspondence and conversations with ICISS Commissioners confirm that the architects of the R2P borrowed the phrase because of what it meant and how it was being used in the context of counter-terrorism. That is to say, they purposefully transferred the 'unable or unwilling' language to convey the idea that a threshold for external intervention has been crossed from its original area of application (self-defence against non-state actors) to the destination context of the R2P with the idea of *norm migration* in mind.[309]

The transfer of legal concepts from one area of international law or global governance to another is not uncommon. The ICISS report, as the next chapter details, is not the only instance in which the 'unable or unwilling' language alone has migrated across different contexts in which the question of the legality or legitimacy of the extraterritorial use of force arises. This and similar phenomena are well-documented in the literature characterising the various modes of norm travel.[310] Perhaps the most oft-discussed norm movements are those that fall under the rubric of *norm diffusion* – a term which implies the one-way vertical top-down journey of norms from the international to the domestic context, from the global to non-global.[311] Instead, I employ the term *norm migration* to characterise the horizontal movement of the 'unable or unwilling' test between different

[307] Gallagher, 'What constitutes a "Manifest Failing"?', pp. 430-432.

[308] For examples of the interchangeable and hybrid uses of the two concepts by scholars and policy-makers alike, see Gallagher, 'What constitutes a "Manifest Failing"?', pp. 429-430, 433-434, 443.

[309] This is a case of *intentional* norm migration, defined as 'involving the conscious efforts of norm entrepreneurs'.: Roderick A. Macdonald, 'Three Metaphors of Norm Migration in International Context', 34 (3) *Brooklyn Journal of International Law* (2009): 603-653, p. 603. The distinction between intentional an unintentional norm migration is discussed and illustrated in Finn Makela's work on viral norm migration: Finn Makela, 'The Drug Testing Virus', 43 (3) *Revue Juridique Themis* (2009): 651-703.

[310] In a nutshell, 'norms outline acceptable behaviour and create a circumscribed framework within which debate on appropriate action can take place'.: Aidan Hehir, *Hollow Norms and the Responsibility to Protect*, (Hampshire: Palgrave Macmillan, 2019), pp. 5-6.

[311] Suzanne Zwingel, 'How Do Norms Travel? Theorizing International Women's Rights in Transnational Perspective', 56 (1) *International Studies Quarterly* (2012): 115–129, p. 124.

contexts in which the extraterritorial use of force can be deemed permissible under the international laws on the use of force. In a similar fashion, Suzanne Zwingel uses the term 'translation' instead of 'diffusion' to emphasise that 'differently contextualized norms may be translated into another realm'.[312] She draws on the anthropological literature using the term 'translation' to describe how 'the researcher has to convert both cultural concepts and the transmitters of these concepts such as language and customs into his or her own system of meaning to enable cross-cultural understanding'.[313] This explanation is consistent with the manner in which this thesis proposes to translate the factors relevant to the 'unable or unwilling' requirement, as in the threshold for the extraterritorial use of force against non-state actors, into the context and language of the R2P. What is more, as mentioned above, this proposition to extend the 'unable or unwilling' test follows ICISS' own design for the inclusion of the 'unable or unwilling' terminology in the original R2P concept. A design, which in itself is perfectly consistent with the logic of norm migration (also typically explained as 'norm harmonisation' or 'transplant'), 'predicated upon the existence of explicit efforts to import or export norms, which can be described in terms of a unified process that develops according to a linear logic departing from an identifiable moment'.[314]

That the 'unable and unwilling' terminology migrated from the realm of the extraterritorial use of force against non-state actors into the arena of collective international humanitarian action to protect innocent populations from egregious violations of their human rights is neither accidental, nor surprising, because the two scenarios are analogous. With regards to the right to self-defence, the 'unable or unwilling' test is ordinarily employed in the following situations:

> 'Non-state actors, such as Al-Qaeda, operating within weak states such as Afghanistan and Somalia ("host states") are often suspected of launching attacks against states such as the United States ("victim states"). These victim states have at times directly attacked such actors within the territory of the

[312] Ibid.

[313] Ibid.

[314] Makela, 'The Drug Testing Virus', p. 660.

host state, without the host state's consent, alleging that the latter is *unwilling or unable* to prevent attacks' (emphasis added).[315]

In such circumstances, as Ahmed explains, states have a responsibility 'to prevent non-state actors from executing armed attacks against other states from within their territory', as reflected in ICJ rulings, and UN Security Council Resolutions (before and after 9/11).[316] The recognition of this responsibility reflects a recent trend observed by Theresa Reinold, where '[t]he language of sovereign responsibility is increasingly being exported from the human rights field to counterterrorism operations, where it provides the intervening states with a powerful vocabulary to press for forceful action against "irresponsible" sovereigns'.[317] This brief description of the way in which the 'unable or unwilling' test is employed in the context of self-defence against non-state actors offers sufficient insight to establish a few key parallels with the substantive threshold for collective action in the context of the R2P.

First, host states from which non-state actors are operating and host states in which mass atrocities are taking place find themselves in analogous positions. Both wish to safeguard their territorial integrity, avoid the external use of force within their sovereign jurisdiction and to be viewed as fulfilling their international legal obligations, be they towards other states or their own populations.[318] Similarly, just like the affronted victim state has an interest to prevent / put an end to the harmful attacks against them, the states of the international community have an interest, and an extra-legal responsibility, to prevent / put an end to atrocity crimes that innocent populations in third states are subjected to. Relatedly, the victim state and the international community alike have an interest to avoid using force in the host state (and expending resources in the process), especially if the host state, who has a responsibility to take the requisite action to address the threat, is able to do so.

[315] Dawood I. Ahmed, 'Defending Weak States against the Unwilling or Unable Doctrine of Self-Defense', 9 *Journal of International Law and International Relations* (2013): 1-37, p. 2.

[316] Ibid. p. 7.

[317] Theresa Reinold, 'State weakness, irregular warfare, and the right to self-defense post-9/11', 105 (2) *American Journal of International Law* (2011): 244-286, p. 285.

[318] Deeks makes a similar presentation of the parallels between the use of force under neutrality laws and self-defence against non-state actors, so as to justify extending the 'unable or unwilling' test from the former to the latter context: Deeks, '"Unwilling or Unable"', p. 503.

What is more, both situations in which the 'unable or unwilling' test may apply are consistent with the following formula: State X is invested with certain rights (those of a sovereign), contingent upon State X fulfilling its associated obligations (towards other states / towards their populations). Failure to fulfil these obligations, namely when State X is 'unable or unwilling' to suppress a threat within its territory (of a non-state actor / mass atrocity crimes), means that Actor Y (the affected victim state / the international community) may be justified in using force in State X's territory, without State X's consent, to eliminate the threat. The above discussion has evidenced the strong comparability of the two scenarios in which the 'unable or unwilling' test has been employed and meant to serve as a guide for those making decisions on the appropriateness of the use of force as well as coercive measures short of the use of force in the context of the R2P. The robust foundation of the proposed analogy suggests that some of the factors relied upon to make 'unable or unwilling' determinations in the context of self-defence against non-state actors may apply and be transferable to inquiries as to whether a host state is 'manifestly failing' in the context of R2P scenarios. This suggests that the detailed investigation into the substance of the 'unable or unwilling' test conducted in Chapter III is both reasonable and worthwhile.

For the purposes of the present chapter, suffice it to say that at the heart of the move to extend the 'unable or unwilling' factors that legal scholars have identified in the context of self-defence against non-state actors into the context of the R2P lies the logic of *reasoning by analogy* – a method commonly deployed in legal scholarship. In a nutshell, '[r]easoning by analogy involves saying that, if a number of different things are similar to each other in a number of different specific ways, they are, or should be, similar to each other in other ways as well'.[319] As seen in the previous chapter, Rosenberg and Strauss identify 'the most appropriate and fair set of standards for RtoP' by examining analogous 'source areas [of domestic and international law], highlighting the potential overlap in normative, political and legal concerns between each area and that of the RtoP'.[320] Similarly, by examining the analogous sources area of the extraterritorial self-defence against non-state actors, this thesis argues that there are reasonable grounds to believe that as the 'unable or unwilling' tests has moved from one area of international law to

[319] Ian McLeod, *Legal Method.* 2nd Edition. (Basingstoke: Macmillan, 1996), p.16.
[320] Rosenberg and Strauss, 'A Common Approach', p. 58.

another, its properties (or at least some of them) in the area of origin could apply in the area of destination (R2P).

What makes this analogy even more compelling is the fact that it rests on the abovementioned strong genealogical connection (established through the ICISS report) between the 'manifest failing' and the 'unable or unwilling' tests rooted in the context of self-defence against non-state actors. A genealogical approach to the study of the 'manifest failing' threshold is consistent with the metaphor of *norm migration*.[321] Similarly to Finn Makela's genealogical investigation of *viral norm migration* in the legal framework governing employment drug testing in Canada, instead of seeking the origins of 'manifest failing' in some foundational moment, I focus on demonstrating its 'relationship to other propositions at other times and in other places'.[322] This, Makela argues, is not much 'different from the standard notion of genealogy' commonly understood as a '"descent" into the past'.[323] Much like tracing a family tree 'one is not doing history *per se*; it is rather a question of following relationships backwards'.[324] This approach is particularly apposite for the investigation of the 'manifest failing' threshold for several reasons. Notably, as Gallagher underscores, '[t]he phrase 'manifest failing' was introduced in the final drafting stage of the WSOD in order to replace the terminology "unable or unwilling", but little guidance is offered to help decision-makers in determining when a state is "manifestly failing"'.[325] Furthermore, the phenomenon that the term 'manifest failing' describes is not new, given that, as noted above, it is used interchangeably with the term that preceded it, 'unable or unwilling', to convey the same meaning. As practitioners and scholars alike continue to oscillate between the two phrases, it is clear that the move to introduce the 'manifest failing' concept was not entirely successful in discarding its predecessor, and by extension, the concomitant historical or legal associations it carries.

Contextualising the 'manifest failing' terminology within a wider historical perspective offers the opportunity to explore the continuities with its predecessors, namely, the 'unable or unwilling' concept from the ICISS report and the latter's ancestor – the 'unable

[321] Similarly, Makela relies on the compatibility between a Foucauldian-inspired notion of genealogy and *viral norm migration*.: Makela, 'The Drug Testing Virus', p. 660.

[322] Ibid. p. 659.

[323] Ibid.

[324] Ibid.

[325] Gallagher, 'What constitutes a "manifest failing"?', p. 430. For more details, see further discussion on pp. 131-132 of this thesis.

or unwilling' test as applied in the context of the extraterritorial use of force in self-defence against non-state actors. The goal here is not to advocate that they are substantively identical (they are, after all, embedded in different political and legal contexts), but to better comprehend the role and substance of the newer 'manifest failing' threshold requirement in contemporary interpretive debates over international action that cohere with the spirit of R2P's reactive dimension. It is on this basis that this thesis maintains that tracing the 'historical lineage' of the 'unable or unwilling' test in order to gain a better understanding of the underpinnings of this norm in its context of origin may help to inform substantive normative guidelines for its 'transplant', the 'manifestly failing' threshold, in the analogous context of the R2P.

DERIVING 'MANIFEST FAILING' FACTORS FROM THE VIEWS OF STATES

This endeavour to develop a methodology for identifying the indicators and determinants that should inform inquiries into whether a host state is 'manifestly failing' to protect its populations from the threat of mass atrocities, through an analysis of the views of states involved in decisions over coercive action in response to such situations, will draw inspiration from Ashley Deeks' prominent corresponding attempt to identify the factors relevant to the 'unable or unwilling' test in the context of self-defence against non-state actors. Deeks derives most of the 'unable or unwilling' factors in her framework 'from identifiable trends in the information that [victim] states have offered in their own defense'.[326] Based on the assumption that 'victim states proffer those aspects of their situation that they perceived as most compelling to other states', Deeks uses these 'recitations of facts' to 'extrapolate general categories of information that […] states are most inclined to proffer'.[327] This method is inspired by Reisman's conception that 'incidents may serve as a type of 'meta-law,' providing normative guidelines for decision-makers in the international system in those vast deserts in which case law is sparse'.[328]

For the purposes of examining the 'manifest failing' threshold, Deeks' approach was adapted as follows. The factors that will inform the framework proposed here will be

[326] Deeks, '"Unwilling or Unable"', p. 516.

[327] Ibid. A list of such categories of information offered by Deeks: 'state's armed forces, information that suggests a relationship between the territorial state's leadership and the nonstate actors, the territorial state's real and claimed levels of control over particular parts of its territory, and the types of requests that the victim state has submitted to the territorial state'.: Ibid. pp. 516-17.

[328] W. Michael Reisman, 'The Incident as a Decisional Unit in International Law', 10 *Yale Journal of International Law* (1984): 1-19. Quoted in Deeks, '"Unwilling or Unable"', p. 517.

derived from identifiable trends in the information that states tend to offer when making a case for adopting coercive measures against a 'manifestly failing' host state. These statements of facts will then be used to extrapolate general categories of information that states are most inclined to proffer when justifying their decision to take timely and decisive action to put an end to atrocities in a third state. The logic here is also similar to that of Reissman in the sense that the analysis of an incident of the de-facto operationalisation of the collective responsibility to protect (the 2011 Libyan intervention) will aid the formulation of normative guidelines for decision-makers in an area otherwise barren of instances where this non-legal responsibility of the international community has been exercised. Despite adhering to the same method of identifying relevant factors, the resulting approach for developing 'manifest failing' determinants and indicators differs from Deeks' original proposition in important ways due to the differences between the contexts in which the two thresholds apply.

First, when a state is 'unable or unwilling' to prevent a non-state actor within its territory from launching attacks against a third state, this could impel the 'victim state' to act in self-defence. In R2P scenarios, however, it is the 'victim populations' that are negatively affected by a 'manifestly failing' host state that does not protect them. In such cases, the wider international community of states could act in other-defence to put an end to the violence against 'victim populations'. Relatedly, in the former case the burden of making the case that the host state is 'unable or unwilling' falls only on the affected 'victim state' determined to act in self-defence. However, as Paragraph 139 of the WSOD stipulates, decisions on the adoption of coercive measures under Chapter VII in the event of a 'manifestly failing' host state rest with the UN Security Council. This has two important implications. First, it means that the burden of providing justifications for the adoption of coercive measures lies with the 15 Security Council member states. Since this makes them the equivalent of Deeks' 'victim states', their statements are to be analysed for the presence of indicators and determinants relevant to the 'manifest failing' threshold. Second, the decision-making role of the Security Council on the threshold for coercive action also means that in order to determine the factors that member states tend to proffer when making the case that the host state is 'manifestly failing', their statements at UN Security Council debates on the adoption of coercive measures cannot be overlooked.

The consideration of Security Council debates on coercive measures to protect civilians makes the approach adopted here markedly different from the one proposed by Deeks,

who rejects them 'as an accurate indicator of views about the "unwilling or unable" test, either as a legal norm or as applied'.[329] Such debates, according to Deeks, 'are strongly coloured by the politics surrounding the uses of force — apartheid, colonialism, or the Israeli-Palestinian conflict, for example — and may not reveal the actual legal views of states'.[330] In other words, Deeks' pursuit of the legal arguments made by states as evidence of *opinion juris* (broadly understood as evidence of a state's belief that they are acting in accordance with international law) dissuades her from engaging with UN Security Council deliberations.[331] The scarcity of legal argumentation, however, is in no way a prohibiting factor against turning to the UN Security Council in the case of the R2P. Given that the international responsibility associated with R2P's coercive aspect is not a legal one, the present account does not look for evidence of *opinio juris* and thus does not depend on whether the statements of Security Council member states are indicative of their belief they are acting in accordance with international law. Besides, as Hurrel explains, the vital role of the UN Security Council should not be viewed in 'strict legal constitutionalist terms as the authoritative body that can rule on the legality or illegality of force'.[332] Instead, it is best understood as 'a deeply flawed and heavily politicised authoritative body in which arguments can be presented and policies defended, because other, better forums simply do not exist'.[333] What is more, as 'the sole organ with the capacity for collective judgement and mobilization of force in interpreting the UN Charter,'[334] the UN Security Council is an important inter-state forum where states

[329] Deeks, '"Unwilling or Unable"', p. 517.

There are also practical reasons for Deeks to avoid focusing on UN Security Council debates. One is the scarcity of records of the reactions of 'uninvolved states' (i.e. states other than the 'victim state'), especially in the pre-Charter era. The other stems from the fact that it is the 'victim state' that needs to make a case for self-defence against a non-state actor in order to justify its incursion upon the sovereignty of the 'host state'. Arguably, the reactions of 'uninvolved states' in the Security Council can tell us much about the 'unable or unwilling' test, as the receiving audience of such claims that can judge their validity.: Ibid. Either way, neither of these objections warrant further consideration, since they do not apply in the context under study. I have therefore focused on unpacking the far more relevant objection to considering Security Council conversations raised by Deeks, namely their heavily politicised nature.

[330] Ibid.

[331] Andrew T. Guzman and Timothy L. Meyer, 'Customary International Law in the 21st Century' in Russell A. Miller and Rebecca M. Bratspies (eds.) Progress in International Law (Leiden: Martinus Nijhoff Publishers, 2008): 197-218, p. 199.; See also Andrew T. Guzman and Timothy L. Meyer, 'International Common Law: The Soft Law of International Tribunals', 9 (2) *Chicago Journal of International Law* (2009): 515-532, p. 526.

[332] Andrew Hurrell, 'Legitimacy and the Use of Force: Can the Circle Be Squared?'. 31 *Review of International Studies*: *Force and Legitimacy in World Politics* (2005): 15-32, p. 24.

[333] Ibid.

[334] C. Cora True-Frost, 'The Security Council and Norm Consumption', 40 *International Law and Politics* (2007), p. 116.

routinely justify their policies through giving reasons.[335] When the situation under consideration is one that concerns the use of force and the UN Charter's articles on international peace and security, the principal site of the international discourse is precisely the Security Council.[336] Therefore, like others before me (e.g. Morris, Hehir and Dunne and Gelber),[337] I shall draw on the official high-level UN discourse in my analysis of the justifications provided by UN Security Council member states in support of coercive measures to put an end to atrocities.

Since states conduct their international relations largely through 'diplomatic conversation – explanations and justifications, persuasion and dissuasion, approval and condemnation,'[338] the justifications provided by states can tell us a great deal about the concerns they deem important when making difficult decision over the use of coercive measures against other states. However, as Deeks herself recognises, deriving 'unable or unwilling' factors from the explanations and statements of 'victim states' will unavoidably contain bias, since the identified factors will be shaped by states that are not only willing to use force in another state's territory but possess the capacity and power to do so.[339] The same is true of deriving 'manifest failing' indicators from the statements and justifications provided by UN Security Council member states. Since representatives to the UN Security Council often recite statements fashioned by their respective state departments, the door remains open for states to manipulate their position in a way that fits with their strategic interests.[340] Although short- and long-term interest-based calculations and power undoubtedly play role in this process, the influence of such factors is to some extent limited by discourse and norms.[341]

---

[335] Jennifer Mitzen, 'Reading Habermas in anarchy: multilateral diplomacy and global public spheres', 99 (3) *American Political Science Review* (2005): 401-417, p. 401.

[336] Ian Johnstone, 'UN Security Council Deliberations: The Power of the Better Argument', 14 (3) *European Journal of International Law* (2003): 437-480, p. 452.

[337] Morris, 'Libya and Syria'; Aidan Hehir, 'The Permanence of Inconsistency: Libya, the Security Council, and the Responsibility to Protect', 38 (1) *International Security* (2013): 137-159. Aidan Hehir, 'The Dog That Didn't Bark? A Response to Dunne and Gelber's Analysis of RtoP's Influence on the Intervention in Libya', 7 (2) *Global Responsibility to Protect* (2015): 211-224; Dunne and Gelber, 'Arguing Matters', pp. 326-349.

[338] Abraham Chayes and Antonia Handler Chayes, *The New Sovereignty: Compliance with International regulatory Agreements* (Cambridge, MA: Harvard University Press, 1995), p. 119.

[339] Deeks, '"Unwilling or Unable"', p. 518.

[340] Ibid.

[341] Johnstone, 'UN Security Council Deliberations', p. 477.

In line with a substantial tradition of legal writing, reason-giving or 'providing persuasive reasons as to why a course of action, a rule, or a political order, is right and appropriate' is understood to have a pivotal legitimating function in legal, political or moral debates.[342] This conception rests on the premise that states involved in such debates are inevitably compelled to justify their actions with reference to something other than self-interest, be it because they seek support domestically or internationally (from other states) or because they have 'internalized standards of appropriate behaviour [...] embedded in international legal norms'.[343] Providing convincing reasons for a state's political actions requires more than, as Watts cynically suggests, 'advanc[ing] a legal justification for their conduct which is not demonstrably rubbish'.[344] Legal arguments play an important role in justificatory discourses by limiting the range of reasonable arguments that can be invoked by governments in the process of substantiating a particular course of action, in this instance obtaining approval for uses of force and/or less intrusive coercive measures.[345] As Hurrell puts it, the vital role of law in legitimacy has less to do with the legal system's capacity to furnish an unequivocal 'answer as to what the law is' (e.g. the legality of the Iraq war).[346] Instead, it emanates from the presence of deep-rooted patterns of argumentation within law 'about the use of force, about the rules that have governed and might govern the use of force, about the ways in which political interests can be expressed in a common language of claim and counter-claim'.[347]

What is more, although legal norms play an important role in the justificatory discourse necessary for decision-making on the use of force, it is also guided by rules and norms (enjoying varying degrees of shared understanding, compliance and determinacy) that are not codified in international law. [348] All of the above, repeated over time, play a role in shaping the legal and moral debates over the use of force and other forms of coercive

[342] Hurrell, 'Legitimacy and the Use of Force', pp. 24-25. For a detailed explanation, see Johnstone, 'UN Security Council Deliberations', p. 441.

[343] Erik Voeten, 'The Political Origins of the UN Security Council's Ability to Legitimize the Use of Force', 59 *International Organization* (2005): 527-557, p. 537.

[344] Sir Arthur Watts, 'The Importance of International Law', in Michael Byers (ed.) *The Role of Law in International Politics: Essays in International Relations and International Law* (New York: Oxford University Press, 2001): 5-16, p. 7.

[345] Voeten, 'The Political Origins of the UN Security Council's Ability to Legitimize the Use of Force', p. 537.

[346] Hurrell, 'Legitimacy and the Use of Force', p. 25.

[347] Ibid.

[348] Voeten, 'The Political Origins of the UN Security Council's Ability to Legitimize the Use of Force', p. 537.

action against the wishes of sovereign states. This is reflected in Michael Walzer's explanation of how our understanding of what constitutes a just war was formed: 'Reiterated over time, our arguments and judgements shape what I want to call the moral reality of war – that is, all those experiences of which moral language is descriptive or within which it is necessarily employed'.[349] To elucidate how political principles limit the behaviour of states, I turn to Quentin Skinner's explanation that the principles professed by states have constraining effects on their behaviour and actions, 'obligat[ing] them to behave in such a way that [their] actions remain compatible with' the professed principles even if they were not genuinely motivated by any of them.[350]

Skinner concedes that political behaviour is rarely motivated by an agent's professed political principles or ideals.[351] Nevertheless, he maintains that even in the most extreme cases of mismatch between an agent's genuine and professed motives, be they 'ex post facto rationalisations' or not, 'it still does not follow [...] that we have no need to refer to this agent's professed principles if we wish to explain his behavior'.[352] Such an argument overlooks what was already established with the introduction of the notion of justificatory discourse, namely 'the fact that any agent possesses a standard motive for attempting to legitimate his untoward social or political actions'.[353] To demonstrate the potency of his argument, Skinner turns to consider the hardest case in which 'an agent's professed principles are never his real motives'.[354] He describes this situation as one in which an agent is involved in a form of 'untoward social or political action', but as is normally the case with such actions, it is vital to 'legitimate them to others who may have doubts about their legality or morality'[355] on some sort of 'accepted set of social or political principles'.[356] Significantly, this implies that the agent will be obligated to behave in a way that his actions remain consistent with their claim that it is the above principles that genuinely impelled them to act.[357] Accepting these implications is equivalent to

[349] Michael Walzer, *Just and Unjust Wars: A Moral Argument with Historical Illustrations.* Fourth Edition. (New York: Basic Books, 2006), p. 15.

[350] Quentin Skinner, 'Some Problems in the Analysis of Political Thought and Action', 2 (3) *Political Theory* (1974): 277-303, p. 299.

[351] Ibid. p. 292.

[352] Ibid. pp. 293, 299.

[353] Ibid. p. 299.

[354] Ibid. p. 292.

[355] Ibid. p. 291-292.

[356] Ibid. pp. 293, 299.

[357] Ibid. p. 299.

recognising that the avenues of action available to 'any rational agent in this type of situation must in part be determined by the range of principles which he can profess with plausibility'.[358] The availability, nature and range of evaluative concepts that any agent can apply in a bid to justify their behaviour 'is a question about the prevailing morality of the society in which the agent is acting, [whereas] their applicability is a question about the standard meaning and use of the terms involved, and about how far these can be plausibly stretched.'[359] And since actions may be hindered in the absence of acceptable legitimation, the above factors that constrain what states can reasonably profess also serve as 'constraints and directives' about the specific conduct that allows an agent to bring his 'actions in line with some accepted principle' in order to justify 'what he does while still gaining what he wants'.[360] Thereby, agents can neither set the principles that they can appeal to, nor stretch them indefinitely, but merely 'legitimate a restricted range of actions'.[361]As Skinner concludes, all of this means that 'to study the principles which the agent finally chooses to profess must be to study one of the key determinants of his decision to follow out any one particular line of actions'.[362]

R2P sceptics and humanitarian intervention proponents alike have drawn on Skinner's thesis to analyse international response to humanitarian crises. Representative of the former is Justin Morris' analysis of UN deliberations on the decision to intervene militarily in Libya, which questioned the role that the R2P played in the decision to intervene.[363] Nevertheless, Morris recognized that 'consideration of those factors declared by the UNSC to be decisive to the manner in which votes were cast over Resolution 1973 is itself revealing as it demonstrates the ways in which states sought to legitimize their position'.[364] My own analysis of Security Council meetings on resolutions sanctioning coercive measures in line with the R2P is premised on the same logic.

The 'central role played by the Security Council in legitimating the threat or use of force' on humanitarian grounds has also been recognized by humanitarian intervention advocate

[358] Ibid.
[359] Ibid. pp. 299-300.
[360] Ibid. p. 300.
[361] Ibid. pp. 299-300.
[362] Ibid. p. 300.
[363] Morris, 'Libya and Syria', p. 1272.
[364] Ibid. p. 1271.

Nicolas Wheeler.[365] In *Saving Strangers*, he incorporated Skinner's contribution to our understanding of language into a constructivist approach to analysing the humanitarian justifications offered by intervening states for their actions during Security Council deliberations on the UN-authorised interventions in Iraq, Somalia and Rwanda.[366] Following Skinner, Wheeler rejected the idea that states can legitimise any action in a way convenient to themselves with reference to indeterminate rules, because they are constrained and enabled by the language they use, which forms part of a bounded reality.[367] Drawing on Hedley Bull's argument that 'rules are not infinitely malleable and do circumscribe the range of choice of states which seek to give pretexts in terms of them',[368] Wheeler confirms the relevance of Skinners' argument that 'any course of action is inhibited from occurring if it cannot be legitimated' at the international level.[369] Following this line of argumentation, Wheeler claims that through 'justify[ing] intervention in humanitarian terms [states] establish a normative benchmark against which' their future actions can be judged, as they become 'entangled in their justifications'.[370] Consequently, 'actors who accept the "need to legitimate"' are constrained to actions that can be reasonably justified with reference to the humanitarian reasons they have professed to incite their response in any given situation.[371]

As the above examples illustrate, Skinner's thesis serves to provide a compelling answer to those critical of the heavily politicised nature of UN Security Council discussions and the concomitant lack of assurance that states will not misrepresent their true motivations, whilst giving credence to the arguments in favour of reason-giving as a form of legitimation. It may be that strategic incentives are the only motivation states have for pursuing a certain course of action, but their success in achieving the desired outcome depends on existing norms of acceptable behaviour. Therefore, much like Morris' assessment of the influence of R2P on UN Security Council member states' voting behaviour in the context of the Libyan intervention, my analysis of states' attempts to legitimise coercive action on humanitarian grounds is 'concerned with the normative development and the justifications for actions given by states, rather than with underlying

[365] Wheeler, *Saving Strangers*, p. 139.
[366] Ibid. pp. 7, 9, pp. 139-141.
[367] Ibid. p. 6, 9.
[368] Ibid. p. 25.
[369] Ibid. p. 25
[370] Ibid. p. 39.
[371] Ibid. p. 40.

motives'.[372] This focus on the professed justifications offered by states has the distinct advantage of circumventing some complex empirical questions surrounding the epistemic uncertainty as to the real (and most likely mixed and complicated) motives behind UN Security Council member states' decision to sanction coercive action, which in no way affect the validity of my general analysis and argument.

By explaining how an agent's professed motivations play a role in constraining their behaviour, I have substantiated the rationale behind turning to the justificatory rhetoric employed by UN Security Council member-states when they seek to legitimise the use of coercive measures consistent with the UN Charter. Paragraph 139 of the WSOD specifies the host state's 'manifest failing' to fulfil its protection responsibilities as the only acceptable justification for what are undoubtedly some of the most untoward political actions – the non-consensual intervention in another state's affairs through the adoption of coercive measures intended to induce individual states to observe their protection responsibilities. Hence, at least in theory, an analysis of the justifications offered by Security Council members in favour of coercive action on humanitarian grounds in an R2P scenario should unveil the factors that states deem most relevant to making a case that a host government is 'manifestly failing' in its responsibility to protect its populations. As Deeks notes, although relying on the statements of states wishing to use force in another state's affairs may not result in an ideal normative outcome, states are more likely to 'respect standards rationally related to concerns they find appropriate' and abide by constraints that they have themselves created and fortified through the justifications they have offered.[373] Hence, the high-level justificatory discourse on the adoption of coercive measures is the closest we can get to a 'reasonable' source for identifying the factors that state regard as germane to the 'manifest failing' threshold.

The case study method

This thesis focuses on one concrete manifestation of this discourse, namely UN Security Council deliberations over the threshold for coercive action in the context of the 2011 Libyan crisis, the analysis of which constitutes the third prong in the overall triangulation

[372] Morris, 'Libya and Syria', p. 1299.
[373] Deeks, '"Unwilling or Unable"', p. 518.

strategy adopted in this thesis. The rest of this section will explain why this method is deemed to be the most appropriate for the purposes of this piece of research.

To begin with, Gerring observes that the case study approach is 'generally more useful…when inferences are descriptive rather than causal'.[374] In other words, 'what' questions, such as the one posed in this thesis, 'are easier to answer without recourse to cross-unit analysis than *Why?* questions'.[375] The present study passes this test for the aptness of the case study method, premised on the form of the research question it poses and its research objectives, as it sets out to draw descriptive inferences about the threshold for coercive action from relevant state discourses in a bid to clarify what constitutes a 'manifest failing'. An equally important reason to opt for the case study approach is its 'natural advantage in research of an exploratory nature'.[376] Despite facing criticism on grounds of its 'subjectivity' in comparison with cross-case analysis, 'it is the very fuzziness of case studies that grants them an advantage in research at the exploratory stage'.[377] Given that the study of the 'manifest failing' threshold and its indicators is still at the exploratory stage, a corresponding strategy is optimal because it 'allows for the generation of a great number of … insights that might not be apparent to the cross-case researcher who works with a thinner set of empirical data across a large number of cases and with a more determinate (fixed) definition of cases, variables, and outcomes'.[378] Hence, a cross-case research design, would hardly be worthwhile in the context of the 'manifest failing' test at present, given the absence of sufficient understanding of this threshold for coercive action and what it entails.

Eckstein associates the advantage of the case study for exploratory research with so-called 'heuristic case studies (heuristic means "serving to find out")', defined by their use to 'deliberately…stimulate the imagination towards discerning important general problems and possible theoretical solutions'.[379] His reasoning for this approach is that: 'Theories do not come from a vacuum, or fully and directly from data. In the final analysis they

[374] John Gerring, 'What is a Case Study and What Is It Good for?', 98 (2) *American Political Science Review* (2004): 341-354, p. 352.

[375] Ibid. p. 345.

[376] John Gerring, *Case Study Research: Principles and Practices* (New York: Cambridge University Press, 2007), p. 39.

[377] Ibid. p. 41.

[378] Ibid.

[379] Harry Eckstein, 'Case Study and Theory in Political Science', in Roger Gomm, Martyn Hammersley, and Peter Foster (eds.) *Case Study Method* (London: SAGE, 2000): 119-164, p. 137.

come from the theorist's imagination, logical ability and ability to discern general problems and patterns in particular observations'.[380] In this sense, the case study (unlike the cross-case) approach allows for 'intensive analysis that does not commit the researcher to a highly limited set of variables, and thus increases the probability that critical variables and relations will be found'.[381]

The way in which the case study approach is used in this thesis is consistent with Eckstein's description of the 'heuristic case' because in addition to purposefully seeking 'potentially generalizable' insights, it also 'directly tie[s] into theory building'.[382] In the context of the present study, 'theory is simply regarded as any mental construct that orders phenomena or inquiry into them… including… "analytic" schemes that decompose complex phenomena into their component elements; frameworks and checklists for conducting inquiry…; any empirical patterns found in properly processed data, or anything considered to underlie such patterns'.[383] Eckstein's 'soft' definition of theory accurately reflects the specific objectives of this thesis, namely to identify the factors relevant to determining whether a state is 'manifestly failing' in its responsibility to protect in a bid to design a framework that could help future inquiries into this matter.

Although case studies 'are often aimed at drawing lessons [in the form of conclusions] that are generalizable to new contexts', their potential to do so is ordinarily regarded as a fundamental weakness of this method.[384] In the methodological literature, 'issues of generalizability are usually discussed under the name of external validity'.[385] In the context of case study research external validity is best understood as '[t]he extent to which the findings from a case study can be analytically generalized to other situations that were not part of the original study'.[386] According to Yin, '[t]he challenge of making analytic generalizations [from case studies] involves understanding that the generalization is not statistical (or numeric) and that you will be making an argumentative claim'.[387] He further

[380] Ibid. p. 138.

[381] Ibid.

[382] Ibid. p. 137.

[383] Ibid. p. 125.

[384] Attilia Ruzzene, 'Drawing Lessons from Case Studies by Enhancing Comparability', 42 (1) *Philosophy of the Social Sciences* (2012): 99-120, pp. 99-100.

[385] Ibid. p. 99.

[386] Robert K. Yin, *Case Study Research and Applications: Design and Methods*, 6th Edition (Thousand Oaks, CA: SAGE, 2018), p. 287.

[387] Ibid. p. 41, see also p. 267.

asserts that to think about extrapolating from case studies as a form of statistical generalisation is a 'fatal flaw' in case study research, because 'cases are not "sampling units" and also will be too few in number to serve as an adequately sized sample to represent any larger population'.[388] A case study should instead be regarded as 'the opportunity to shed empirical light on some theoretical concepts or principles'.[389] Analytic generalisations could either be premised on 'corroborating, modifying, rejecting, or otherwise advancing theoretical concepts' and propositions that originally informed the case study design or 'emerge from the case study's findings alone'.[390] In this thesis, the triangulation research design entails that the case study findings could empirically enhance the theoretical propositions emerging from secondary sources and reasoning by analogy in both ways described above.

One way to address the issue of external validity here is in terms of a trade-off between making generalisations and 'the opportunity to study a single unit in great depth [which] constitutes one of the primary virtues of the case study method'.[391] As Gerring puts it, the case study is particularly useful 'when propositional depth is prized over breadth and boundedness'.[392] Ultimately, when faced with a dilemma between a case study and a cross-case approach one has to choose 'between knowing more about less, or less about more'.[393] The case study method is adopted in this thesis because, given the scarcity of existing research on the 'manifest failing' test, there is more to be gained by being in a position to generate more findings and say much more about the 'manifest failing' phenomenon with accuracy, than by the pursuit of an increased inferential scope. It is important to keep 'in mind that case studies often tackle subjects about which little is previously known or about which existing knowledge is fundamentally flawed'.[394] In fact, the choice between depth and breadth is often guided by such prosaic considerations, as 'much of the debate over the utility of the case study method has little to do with the method itself and more to do with the state of current research in a particular field'.[395]

[388] Ibid. p. 38.
[389] Ibid.
[390] Ibid.
[391] Gerring, 'What is a Case Study', p. 345. Yin, *Case Study Research*, p. 4.
[392] Gerring, 'What is a Case Study', p. 352.
[393] Gerring, *Case Study Research*, p. 49.
[394] Gerring, 'What is a Case Study', p. 345.
[395] Gerring, *Case Study Research*, p. 62.

What is more, despite being seen as largely weaker for the purposes of generalisation across cases, even a single case can provide valuable clues that can play a vital role in informing a general framework. In his book *Case Study Research Design and Applications,* Yin provides a wealth of prominent and highly regarded examples of single-case study research that have made impactful contributions in political science and other branches of the social sciences.[396] One such example of a descriptive case study, which bears resemblance to the present research, is Neustadt and Fineberg's 'analysis of a mass immunization campaign… that took place under President Gerald Ford's administration, when the United States was faced with a threat of epidemic proportions from a new and potentially lethal influenza strain'.[397] In his analysis of this case, Yin points out that:

> 'Because the case study has become known as an exceptionally well-researched case study, contemporary policy makers have continued to consult it for any generalizable lessons for understanding the quandaries of health crises and public actions in light of new threats by flu epidemics, such as the H1N1 strain of 2008 2010 and by viruses such as the Ebola and Zika outbreaks of 2013 to the present'.[398]

This is just one example of how a single well-executed descriptive case can aid the understanding of future cases and serve as a guide for policy makers when a similar problem arises in the future.

Conversely, as Gerring observes, 'the sheer number of examples of a given phenomenon does not, by itself, produce insight. How many times did Newton observe apples falling before he recognized the nature of gravity?'.[399] This apocryphal illustration aptly illuminates the point that a case study could be 'more useful than cross-case studies when a subject is being encountered for the first time or is being considered in a fundamentally new way'.[400] Even though the 'manifest failing' threshold has been studied before in the

---

[396] Graham Allison's explanatory study of the 1962 Cuban missile crisis (contrasting different theoretical explanations of the crisis), which 'has been a political science best seller' for over 40 years since its publication 1971, is a noteworthy example that 'forcefully demonstrates how a single case study can be the basis for insightful generalizations'.: Yin, *Case Study Research*, p. 7.

[397] Yin, *Case Study Research*, p. 8. The case was first published as a 1978 government report titled '*The Swine Flu Affair: Decision-Making on a Slippery Disease*, and later published independently as *The Epidemic That Never Was* (1983)' (emphases added).: Ibid.

[398] Ibid.

[399] Gerring, *Case Study Research*, p. 40.

[400] Ibid.

context of an empirical example (viz. Gallagher's investigation of the Syrian crisis), it has never been done with a focus on the factors that states professed when making the case that this threshold has been crossed (or not) in debates over the need for collective action in response to atrocity crimes. The value of and rationale behind investigating the subject matter in this new way was already detailed in the previous section. In addition to the justifications for adopting the case study method discussed so far, there are other considerations that played a role in the decision to opt for a single case approach. These will be articulated below, as they are best understood within the context of the ensuing discussion on case selection.

Case selection: Libya

The analysis of decision-making states' views on the 'manifest failing' threshold is naturally premised on the existence of UN Security Council discourses over the threshold for collective action. In light of this thesis' objective to address the substantive indeterminacy of the 'manifest failing' threshold with a view to its important function as a substantive licence for, and a leash against, coercive international action, the selected case should lend insights into the threshold for the implementation of the most controversial (coercive) aspects of the R2P. Therefore, in order to identify relevant cases, this thesis has focused on locating instances of mass atrocity crimes in the period since the endorsement of the R2P at the 2005 World Summit, in which the UN Security Council has deliberated on the adoption of coercive measures against another state to induce them to respect their protection responsibilities towards their populations. In so doing, preference has been given to cases in which the 'manifest failing', 'unable or unwilling' or closely related concepts were explicitly invoked in debates and state justifications relating to the adoption of coercive measures, because they are definitive markers that the substantive threshold requirement for collective action was consciously emphasised and discussed by decision-makers.

In view of these inclusion criteria, the 2011 atrocity crisis in Libya has been singled out in this thesis for several reasons. First, it was widely looked upon as having met the 'manifest failing' threshold. That the threat to civilian populations had escalated beyond human rights violations alone was evident from the Security Council's unanimous recognition of the gravity of the humanitarian situation in Libya as the grounds for UN

action,[401] as well as from earlier statements by the UN Secretariat warning that the violence in Libya may amount to crimes against humanity.[402] Notably, the UN High Commissioner for Human Rights, Ms. Navi Pillay, and the UN Secretary-General explicitly invoked the 'manifest failing' threshold and reminded the international community of their responsibility to take collective action when it is crossed.[403] UN Security Council member states also referred to the threshold for collective action when explaining their decisions to support coercive measures in response to the Libyan crisis.[404] Such compelling evidence that the notion of a threshold being crossed was an integral part of the discourse surrounding the adoption of coercive measures in Libya is perhaps the most ostensible indicator of the suitability of this case for the present examination of the debate over the substantive threshold requirement embodied by the phrase 'manifest failing'.

Second, Libya is one of the most, if not the most, definitive cases in which the UN Security Council has drafted and passed resolutions under a clear human protection mandate explicitly stated as the primary objective for the adoption of coercive measures. This was evident in the wording of the first two paragraphs of Resolution 1970 'condemning the violence and use of force against civilians' and '[d]eploring the gross and systematic violation of human rights' in Libya, and reinforced with subsequent references to 'widespread and systematic attacks' against civilians that may constitute crimes against humanity, and Libya's 'responsibility to protect its population'.[405] Likewise, Resolution 1973 makes the objective of human protection clear from the very start by citing 'the escalation of violence, and heavy civilian casualties' as well as Libya's primary protection responsibilities.[406] The vocabulary of human rights and humanitarian

[401] UNSC, S/PV.6491, 26 February 2011; S/PV.6498.

[402] S/PV.6490 (2011), p. 3; United Nations press release, 'UN Secretary-General Special Adviser on the Prevention of Genocide, Francis Deng, and Special Adviser on the Responsibility to Protect, Edward Luck, on the Situation in Libya', New York, 22 February 2011. [Accessed 17 January 2021]. Available at: http://www.un.org/en/genocideprevention/documents/media/statements/2011/English/2011-02-22-OSAPG,%20Special%20Advisers%20Statement%20on%20Libya,%2022%20February%202011.pdf

[403] United Nations Human Rights Office of the High Commissioner (OHCHR), 'Situation of Human Rights in the Libyan Arab Jamahiriya: Statement by Navi Pillay, UN High Commissioner for Human Rights', Human Rights Council - 15th Special Session, Geneva, 25 February 2011 [Accessed 17 January 2021] Available at: https://newsarchive.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=10760&LangID=E; S/PV.6490 (2011), p. 3.

[404] See for instance, the statement by France in S/PV.6498, pp. 2-3. Further evidence provided in Chapter IV.

[405] S/RES/1970 (2011).

[406] UNSC, S/RES/1973, 17 March 2011.

law violations has been featured in other UN Security Council Resolutions authorising Chapter VII measures as one of the criteria employed by the Council to designate state and non-state actors for sanctions, as seen in Somalia 751, DRC 1533, Côte d'Ivoire 572, and Sudan 1591.[407] According to *Security Council Report*, '[t]he protection of civilians from human rights and humanitarian law violations can … be seen as an additional objective of the Council in these cases'.[408] Yet, the sanctions regime established in Resolution 1970 on Libya appears to be the first instance where 'at the outset the protection of civilians was explicitly stated as the principal objective'.[409] In this sense, the authorisation of international action in response to the Libyan crisis accorded with the spirit of R2P's reactive dimension in way that most cases linked to civilian protection did not (with the notable exception of the Central African Republic, discussed below).

That the international response to the Libyan crisis is the most definitive instance of the operationalisation of R2P's reactive component has been affirmed by numerous scholarly interpretation, portraying it as characteristic of R2P's coercive component. Gallagher refers to Libya (alone) as an example of R2P's reactive dimension and of a UN Security Council judgement that a state is 'manifestly failing' in its protection responsibilities.[410] Jess Gifkins reaffirms this interpretation of the Libyan intervention, suggesting it 'represents a decisive example of implementing the coercive aspects of the…R2P'.[411] Likewise, John-Mark Iyi labels the intervention as the 'the first test case of operationalisation of the "responsibility to react"'.[412]

More recently, Maggie Powers suggested that the UN Security Council mandate to use 'all necessary measures' in response to atrocity crimes in the Central African Republic (CAR) was 'an expression of R2P's most controversial aspect'.[413] As hinted above, the international community's response to the situation in CAR was comparable to Libya in the sense that UN Security Council employed strong civilian protection language in

[407] Security Council Report, 'Special Research Report: UN Sanctions', No 3.

[408] Ibid.

[409] Ibid.

[410] Adrian Gallagher, 'The promise of pillar II: analysing international assistance under the Responsibility to Protect', 91 (6) *International Affairs* (2015): 1259-1275, p. 1259.

[411] Jess Gifkins, 'R2P in the UN Security Council: Darfur, Libya and beyond', 51 (2) *Cooperation and Conflict* (2016): 148-165, p. 149.

[412] John-Mark Iyi, 'Emerging Powers and the Operationalisation of R2P in Africa: The Role of South Africa in the UNSC', 7 (1) *African Journal of Legal Studies* (2014): 149-176, p. 149.

[413] Maggie Powers, 'Responsibility to protect: dead, dying, or thriving?'. 19 (8) *The International Journal of Human Rights* (2015): 257-1278, p. 1258.

conjunction with references to CAR's primary responsibility to protect in resolutions authorising coercive measures.[414] Despite the similarity between the two cases, the debate surrounding the adoption of coercive measures in CAR was deemed less suitable for gaining insights into states' understanding of the 'manifest failing' threshold for the following reason. Since the coercive measures in question were deployed with the 'nominal consent of the transitional government',[415] taking CAR and similar instances of the use of force by the international community as exemplary of a 'manifestly failing' state and R2P's reactive dimension is problematic. As Gallagher reminds us, with regards to the UN-authorised use of force in Côte D'Ivoire, coercive measures can be employed as part of R2P's 'international assistance' component in the presence of host state consent.[416] What is more, Powers herself attests, such measures were deployed in CAR to 'assist the state in meeting its responsibility to protect',[417] rather than to sanction enforcement against it. Since the provision of international assistance is not premised on the 'manifest failing' of the host state (as explained in the Introduction), cases in which Chapter VII measures were employed with host state consent cannot offer a clear representation of the debate over the threshold for coercive action because the latter becomes redundant once host state consent in obtained. From this follows another important reason why Libya was considered the most suitable case for the present task of exploring the conversation surrounding the 'manifest failing' threshold.

Third, Libya stands out from the rest as the only case in which the UN Security Council sanctioned the use of 'all necessary means', 'the UN code words for armed intervention', under a civilian protection mandate against the will of the host state.[418] Paul D. Williams reflects on several cases in which the Security Council was close to authorising such a military intervention, but they were ultimately treated either as instances of the use of force 'in the absence of a central government rather than against one' (Somalia) or with the consent of the state's *de jure* authorities (Haiti, Democratic Republic of Congo,

[414] See UNSC, S/RES/2127, 5 December 2013; UNSC, S/RES/2134, 28 January 2014; and UNSC, S/RES/2399, 30 January 2018.

[415] Ibid.

[416] Gallagher, 'The promise of pillar II'.

[417] Maggie Powers, 'The Responsibility to Protect after Libya – dead, dying or thriving?', 24 June 2014 [Accessed 27 February 2020] Available at: https://www.opendemocracy.net/en/openglobalrights-openpage-blog/responsibility-to-protect-after-libya-dead-dying-or-thriving/

[418] Doyle, 'The Politics of Global Humanitarianism', p. 14.

Sudan, and Côte D'Ivoire) or interim government (Rwanda).[419] Likewise, Bellamy framed Libya as the only case of the 'the use of force against a UN Member State' amongst 'over a dozen' post-2005 atrocity cases, in which the R2P was invoked explicitly, including five instances '(Mali, Lord's Resistance Army, DRC, Somalia, and the CAR) [where force was] authorized or invited to support de jure governments against non-state armed actors'.[420] In a nutshell, as Bellamy and Williams put it, whilst it is not uncommon for the UN Security Council to authorise force for human protection purposes under the umbrella of peacekeeping missions, Libya is the only case in UN history when it did so 'against the wishes of a functioning state'.[421] In this sense, it sets a precedent not only post-2005, but also amongst atrocity cases predating the pledge of heads of states and government to uphold the principles of the R2P. Whereas other cases could shed some light on the threshold for coercive action as far as non-pacific measures short of the use of force are concerned, only Libya can offer insights into the professed role of the 'manifest failing' threshold in legitimising the authorisation of the full range of coercive measures available to the Council.

Finally, Libya is a textbook example of a scenario in which the international community was left with an uneasy choice between timely and decisive international response and a grave humanitarian tragedy. As Alex Bellamy observers:

> 'The March 2011 crisis in Libya was unexpected and escalated rapidly. Indeed, none of the world's existing genocide/atrocity risk assessment frameworks identified Libya as being at risk, despite some of those lists extending to 68 countries. To their credit, the UN's Special Advisers warned of the risk of crimes against humanity in Libya before most non-governmental agencies did so, but even this was very late in the day'.[422]

The singular circumstances of the Libyan case remonstrate that there is a very real possibility for an incident to slip through existing early warning mechanisms and develop

[419] Williams, 'Briefing: The Road to Humanitarian War in Libya', p. 249. See also Bellamy and Williams, 'The new politics of protection?', pp. 825-850.

[420] Alex J. Bellamy, *The Responsibility to Protect: A Defense* (New York: Oxford University Press, 2015), p. 130.

[421] Bellamy and Williams, 'The new politics of protection?', p. 825.; Paul D. Williams, 'Briefing: The Road to Humanitarian War in Libya', 3 (2) *Global Responsibility to Protect* (2011): 248-259, p. 249; Glanville, 'Intervention in Libya', p. 325.

[422] Bellamy and Williams, 'The new politics of protection?', p. 838.

so rapidly that by the time it comes to the attention of policy makers, it has already gone way beyond the scope of preventive efforts. As such, it highlights the need for a clearer understanding of what a 'manifest failing' entails and a framework that can aid the timely consideration of mass atrocity situations in order to determine whether the host state is fulfilling its responsibility to protect its populations from the four crimes or not. This makes it all the more suitable to explore in this thesis, the ultimate aim of which is to propose a common framework of reference to guide such determinations.

As the above discussion illustrates, there is a relatively small universe of cases that could provide insights into states' understanding of the 'manifest failing' threshold.[423] Gerring points out that in such instances more can be learned from the in-depth study of a single case or a small number of cases than the cross-case analysis of a large number of cases.[424] This is a common 'problem that affects the study of many social science questions where only one unit, or a small number of units, undergoes the event that is to be explained [including] the use of nuclear weapons (United States) [and] world war (WWI,WWII)' and 'explains why case studies often focus on rare ("historical") events', e.g. the French Revolution.[425] The same type of reasoning comes into play when one is faced with the choice between the analysis of a single case versus a small number of cases. Specifically, 'the issue is not whether natural experiments – or real experiments for that matter – are desirable. They are always desirable. The issue is how many experiments are available and how (truly) experimental are they?'.[426]

Aside from Libya, the process of 'manifest failing' can also be studied, to some extent, in the context of other mass atrocity cases where coercive measures were considered (e.g. Syria) and implemented with the consent of the host state (e.g. CAR). However, at the case selection stage, it became obvious that the debate over threshold in other high-profile instances of mass violence was not as clearly discernible as it was in Libya. As mentioned in Chapter I, Gallagher recognises that the Security Council did not explicitly discuss the issue over whether the Syrian regime was 'manifestly failing', which makes it a far less suitable and plausible source of states' views on the subject. Put simply, alternative cases

[423] For a list of atrocity cases detailing UNSC involvement between 2011 and 2015 see Alex J. Bellamy and Blagovesta Tacheva, 'R2P and the Emergence of Responsibilities Across Borders', in Richard Beardsworth, Garrett Wallace Brown, and Richard Shapcott (eds.) *The State and Cosmopolitan Responsibilities* (Oxford: Oxford University Press, 2019): 15-40, pp. 30-31.

[424] Gerring, 'What is a Case Study', p. 351.

[425] Ibid.

[426] Ibid.

could only provide a limited insight into the factors that states professed as critical to informing and guiding their position over the threshold for coercive action and even so presented serious limitations owing to the largely implicit nature of these debates. As Yin highlights, irrespective of 'the rationale for doing single-case studies [they] require careful investigation of the candidate case, to minimize the chances of misrepresentation and to maximize the access needed to collect the case study evidence', so as to avoid a situation in which 'a case may later turn out not to be the case it was thought to be at the outset'.[427] In the context of the present research, where the focus is on studying state discourses on the 'manifest failing' threshold, it is essential that such discourses are present and sufficiently prominent for an in-depth analysis to be possible. Libya stood out as the strongest case in this crucial regard due to the richness and detail of state discussions over the issue of threshold both within and outside the UN Security Council. Hence, Libya was selected because it is the case that best meets the case selection requirements set out at the start of this section and is, by far, the most fruitful case for investigating states' understanding of the 'manifest failing' threshold.

While the analysis of additional far-from-ideal cases could enhance the framework by corroborating the factors of the 'manifest failing' threshold derived from the Libyan case, it may or may not generate new insights or help to identify new factors. Besides, it is worth noting that supplementary heuristic cases would carry much less weight than Libya, because the insights into states' understanding of the process of 'manifest failing' that they can offer by comparison are far more limited and less reliable. In this regard, Gerring points out that the choice of research design is sometimes guided by 'the quality and quantity of information that is currently available, or could easily be gathered, on a given question'.[428] The inferior quantity and quality of relevant information that can be obtained from other high-profile cases was only one prohibiting factor against opting for an overarching case study approach based on two or three cases. More importantly, however, a case study approach alone cannot compare to or replace an approach that triangulates different perspectives for informing a general framework concerning a phenomenon about which little is known and is rarely discussed by academics and policy-makers. The most sensible first step towards addressing these challenges effectively is through an approach that actively draws on as wide a range of sources (i.e. secondary, legal and state

[427] Yin, *Case Study Research*, p. 51.
[428] Gerring, *Case Study Research*, p. 57.

practice) and methods as possible to maximise the chances for obtaining new and diverse insights and perspectives on the phenomenon in question and provide a theoretical basis for analytic generalisations. After all, '[t]he triangulation essential to social scientific advancement demands the employment of a variety of (viable) methods, including the case study'.[429]

As with any new or advanced framework, it could be refined and adjusted to reflect a wider range of empirical conditions through further research. The best way to do this would be through employing 'the so-called building-block technique', which involves examining supplementary heuristic cases in a bid to refine the theoretical construct arrived at on the basis of studying one heuristic case.[430] This, however, is a laborious and time-consuming task, which goes beyond the focus and scope of this thesis, constrained by the standard three-year time frame for completion and word count restrictions. It is important to stress that such additional work is not essential to coming up with a workable normative framework at a level of detail that would cover most cases and could be effectively applied to new contexts, as the last chapter of this thesis will demonstrate. This is, as already discussed, ensured by triangulation strategy adopted in this thesis, which relies on the integration of theoretical and empirical components to produce generalisable insights into the factors that should inform 'manifest failing' determinations.

In sum, the 'methodological affinity between descriptive inference and case study work' make the case study approach adopted in this thesis a perfect fit for achieving the objectives of this research.[431] The fact that many of the most prominent case studies in political science are descriptive is a testimony to the fact that '[d]escriptive inference remains an important, if undervalued, trope within the social sciences'.[432] Descriptive information alone can be revelatory, especially when there is little known about a phenomenon or a problem is approached from a new angle. Likewise, 'single-case studies are a common design for doing case study research'.[433] Case study approaches and especially single case studies are often criticised for providing poor grounds for drawing general lessons transferrable to other concrete situations. However, as stressed above, the

[429] Ibid. p.63.

[430] Eckstein, 'Case Study and Theory in Political Science', p. 137.

[431] Gerring, 'What is a Case Study', p. 347.

[432] Ibid. See examples provided by Gerring: Ibid.

[433] Yin, *Case Study Research*, p. 54.

record of impactful single case studies, be they descriptive, exploratory or explanatory, provides compelling evidence to the contrary. Yin makes a crucial point in this regard, namely that it is important to 'remember that you are generalizing from your case study, not from your case(s)'.[434] Going back to the work of Mary Kennedy (1979) who was perhaps the first to 'call attention to the analogous process in the field of law', Yin reminds us that:

> '[i]nterpretations made from a single legal case may be used as precedents (i.e., generalizations) for future cases. Indeed, the body of legal knowledge appears to grow in this manner. However, the interpretations (i.e. generalizations) are about the ideas or principles established by the case, not about the case and its potentially idiosyncratic demographic features itself'.[435]

Once again, this example illustrates that generalising from a single case is not only possible, but indeed very common, and that it is premised on well-substantiated arguments and inferences specific to the studied case, rather than the supplementary analysis of other cases. Lastly, even though the case study approach has its own merits for drawing generalisable insights, it is important to bear in mind that the case study analysis is not the sole stepping stone on which the 'manifest failing' framework to be devised in this thesis is premised. Nonetheless, by providing the facility to flesh out the details and nuances in states' understanding of the 'manifestly failing' test, the case study findings will help to empirically ground and provide greater texture to the resulting framework for its assessment.

Political Discourse Analysis

The term discourse, as Fairclough clarifies, can be used in a 'general abstract way', 'in what is widely called "discourse analysis"[, to] signa[l] the particular view of language in use…as an element of social life'.[436] It can also be used in a 'particular way [to] refer to particular 'discourses' such as the 'Third Way' political discourse of New Labour'.[437] With the exception of this section of the thesis, the term discourse will be used primarily

[434] Ibid. p. 41.

[435] Ibid. p. 65.

[436] Norman Fairclough, *Analysing Discourse: Textual Analysis for Social Research* (London and New York: Routledge, 2003), p. 3.

[437] Ibid. p. 4.

in the 'particular way' – to refer to specific discourses, such as the discourse over threshold or the justificatory discourse on the Libyan intervention.

The notion of 'language in use' is regularly underscored in a large body definitions of 'discourse'.[438] It follows that discourse analysis is commonly understood as the study of 'what language is used for,'[439] or as Gee puts it, it 'is the study of language at use in the world, not just to say things, but to do things'.[440] Language plays a central role in the process of reason-giving, because justification and persuasion depend on the existence of a shared language 'through which claims can be articulated, addressed and received'.[441] Reason-giving is vital for 'enacting any large-scale project of political violence' such as wars and external military interventions, because their execution necessitates sacrificing significant public resources and possibly the lives of many in conflict, and therefore, 'requires a significant degree of political and social consensus […] not possible without language'.[442] Whilst rhetorical argumentation, especially in the domains of politics, law and morality, is commonly understood in a limited sense as the 'art of aiming for effective persuasive argumentation', according to the original Aristotelean conception of rhetoric it is 'intrinsically connected with deliberation over what to do when several alternatives are possible, with choice and action related to matters of common concern'.[443] Security Council deliberations over threats to international peace and security that may require timely and decisive collective action are precisely such 'matters of common concern' of the highest order. Moreover, as Glanville explains, interrogating the language deployed by states can allow us to 'discern the "standard of appropriate behavior" against which they are seeking to either justify their actions or to admonish the actions of others and from there we can begin to interpret the impact of this standard'.[444] Political discourse exemplified in forms of institutional text and talk such as parliamentary debates, as well as media interviews and political speeches, can also be 'singled out as a prominent way

[438] Adam Jaworski and Nikolas Coupland, 'Introduction: Perspectives on discourse analysis', in Adam Jaworski and Nikolas Coupland (eds.) *The Discourse Reader*. 2nd Edition. (London: Routledge, 2006): 1-37, p. 3.

[439] Gillian Brown and George Yule, *Discourse Analysis* (Cambridge: Cambridge University Press, 1983), p. 1.

[440] James Paul Gee, *How to do Discourse Analysis: A Toolkit* (New York: Routledge, 2011), p. ix.

[441] Hurrell, 'Legitimacy and the Use of Force', p. 24.

[442] Richard Jackson, *Writing the war on terrorism: Language, politics and counter-terrorism* (Manchester: Manchester University Press, 2005), p. 1.

[443] Isabela Fairclough and Norman Fairclough, *Political Discourse Analysis: A Method for Advanced Students* (Abingdon: Routledge, 2012), p. 58.

[444] Glanville, 'Does R2P Matter?', p. 190.

of "doing politics"', since 'most political actions (such as passing laws, decision making, meeting, campaigning, etc.) are largely discursive'.[445]

For the purposes of this thesis, discourses are understood as forms of language used to interpret the social world or produce meaning. Despite a diversity of opinions as to 'what discourse is beyond language in use,'[446] Jackson identifies a 'set of theoretical assumptions or commitments about the nature and consequences of language, text and discourse and social practices' that are generally shared among discourse analytical studies across research paradigms. The present account subscribes to several of these broad commitments that bear mentioning explicitly: 'an understanding of language as constitutive or productive of meaning rather than simply descriptive of an external reality',[447] 'an understanding of discourse as structures of signification that construct social realities [and relations, and] an understanding of discourse as productive of subjects authorized to speak and act, legitimate forms of knowledge and political practices, and, importantly, common sense within particular social groups and historical settings'.[448]

Whilst discourses include symbols, images, sounds, body language and other communicative acts / features,[449] the focus here will be on written and spoken texts, which shape elite political discourses over the threshold for international action in response to mass atrocities. The discourses analysed here are characterised as political in accordance with van Dijk's definition of political discourse as 'identified by its actors or authors, viz., politicians'.[450] In other words, it is 'the text and talk of professional politicians or political institutions, such as presidents and prime ministers and other members of government, parliament or political parties, both at the local, national and international levels'.[451] This straightforward definition of political discourse warrants an additional stipulation, namely that the text and talk of politicians only counts as political discourse when produced in a political context.[452] Taking the entire context as the key determinant for classifying a

[445] Teun van Dijk, 'What is Political Discourse Analysis?', in Jan Blommaert and Chris Bulcaen (ed.) *Linguistics* (Amsterdam: Benjamins, 1987): 12-52, p. 18.

[446] Jaworski and Coupland, 'Introduction: Perspectives on discourse analysis', p. 3.

[447] Richard Jackson, 'Critical Discourse Analysis', in Priya Dixit and Jacob L. Stump (eds.) *Critical Methods in Terrorism Studies* (Abingdon: Routledge, 2016), p. 80.

[448] Richard Jackson, 'Constructing Enemies: "Islamic Terrorism" in Political and Academic Discourse', 42 (3) *Government and Opposition* (2007): 394-426, p. 396.

[449] See Fairclough, *Analysing Discourse: Textual Analysis for Social Research*, p. 3.

[450] van Dijk, 'What is Political Discourse Analysis?', p. 12.

[451] Ibid.

[452] Fairclough and Fairclough, *Political Discourse Analysis*, pp. 17-18.

discourse as '"political" or not', Fairclough and Fairclough characterise 'political contexts [as] institutional contexts, i.e. contexts which make it possible for actors to exert their agency and empower them to act on the world in a way that has impact on matters of common concern'.[453] Examples of such institutional contexts range from parliament and government to the far less obvious internet discussion forums.[454] Here this institutional dimension extends beyond the traditional political domains of parliament and government to also include a focus on the discourses that take place in the context of international organisations, specifically the United Nations.

Alternative methods

The suitability of the proposed method cannot be fully appreciated without a brief reflection on alternative methods for textual analysis. Examining how discourses comprising specific vocabularies are used to convey meaning is markedly different from the widely employed 'method of "content analysis", which identifies certain words, codes them on the basis of different categories and counts them'.[455] Whist the leads that emerge from such techniques for textual data analysis can provide valuable insights in some cases, their limitations make them incompatible with the objectives of this thesis. For instance, as some have suggested 'the fact that people use the word 'nation' three times does not necessarily tell us anything about people's racism'.[456] What makes discourse analysis a fitting method for gaining detailed insights into different aspects of the 'manifest failing' phenomenon, in comparison with content analysis, is its emphasis on the wider context. Specifically, whilst 'discourse analysis focuses on the relation between text and context, content analysis focuses on the text abstracted from its contexts'.[457]

R2P proponents and critics alike have stressed the importance of context when analysing political texts in a bid to ascertain the role of the R2P in decision-making. Despite disparities in opinion over R2P's potency as a norm, Hehir, Morris and Glanville share the view that the impact of the R2P norm cannot be adequately ascertained simply based

[453] Ibid. p. 18.

[454] Ibid.

[455] Marianne Jorgensen and Louise J. Phillips, *Discourse Analysis as Theory and Method* (London: SAGE, 2002), p. 122.

[456] Ibid.

[457] Cynthia Hardy, Bill Harley and Nelson Phillips, 'Discourse Analysis and Content Analysis: Two Solitudes?', 2 *Qualitative Methods* (2004): 19-21, p. 20.

on the frequency of the use of R2P language in high-level political discourse.[458] Hehir places emphasis on the way in which the term is employed and how invocations of the R2P relate to actual state practice.[459] For Glanville, it also comes down to how the language of the responsibility to protect innocent populations is being used and 'how it is received by relevant audiences'.[460] As illustrated in the previous chapter, the failure to take into account the context and manner in which the R2P concept and associated vernacular are invoked can lead to erroneous interpretations and findings.

Although qualitative approaches to content analysis endeavour to overcome context-related limitations, they often end up 'failing to develop a complete understanding of the context, and thus failing to identify key categories [of information]' or remain 'blind… to contextual aspects of the phenomenon,' much like their quantitative counterpart.[461] For these reasons content analytical approaches are not well-suited to this endeavour to better understand what the 'manifest failing' threshold entails and chart the factors associated with this notion. This can only be achieved through a discourse analytic approach, which focuses on examining texts in relation to the context in which they are produced, the goals pursued by their authors, and / or their reception by the intended audience. Having addressed the broader questions surrounding the definition and objectives of the method of choice and justified its appropriateness for addressing the research objectives of this thesis, I will turn to elaborate on how the method of discourse analysis will be employed in the upcoming examination of the Libyan case in Chapter IV.

Employing discourse analysis to identify the factors of the 'manifest failing' threshold

The analysis of the dominant political discourse that legitimised the adoption of coercive measures in Libya in 2011 as the accepted response to the unfolding atrocity crisis in the North African state serves a two-fold purpose, corresponding to the two lines of inquiry identified in the Introduction. The first objective is to chart the debate over threshold in the context of the adoption of coercive measures in Libya with a focus on the vocabulary

[458] Glanville, 'Does R2P matter?', p. 190; Aidan Hehir, 'Assessing the influence of the Responsibility to Protection the UN Security Council during the Arab Spring', 51 (2) *Cooperation and Conflict* (2016): 166-183, p. 173; Morris, 'Libya and Syria', p. 1272.

[459] Hehir, 'Assessing the influence', p. 173.

[460] Glanville, 'Does R2P matter?', p. 190.

[461] Hsiu-Fang Hsieh and Sarah E. Shannon, 'Three Approaches to Content Analysis', 15 (9) *Qualitative Health Research* (2005): 1277-1288, pp. 1280, 1283.

used by states to refer to the notion of threshold. The second – to identify the factors that states tend to proffer when invoking the notion of threshold in a bid to justify coercive action.

To accomplish the first task, each of the selected texts will be examined for specific themes, concepts, labels and arguments invoked by political actors that are relevant to the notion of threshold. This entails identifying specific references to the 'manifest failing', 'unable or unwilling' and alternative terminologies deployed to denote the threshold for coercive action, as well as the ways in which sates made the case that this threshold had been crossed in the absence of specific formulations attached to the notion of threshold.

Gaining an insight into the terminology states use to discuss the 'manifest failing' threshold is a vital prerequisite for identifying instances of justificatory discourse where states defended their position on Libya with reference to this threshold. Having mapped the broader debate over threshold in the selected texts, attention shifts to states' attempts to justify the use of coercive measures with reference to the notion of a threshold being crossed. This narrower analysis of instances of justificatory discourse will focus on identifying the specific factors that states profess when making the case that the threshold for international action has been met in a bid to legitimise a coercive response to the Libyan crisis. As clarified above, these factors are one of the three pillars upon which normative framework to be articulated here is premised.

As Richard Jackson suggests, the descriptive analysis of discourse involves 'descriptively mapping the contours of the discourse in the chosen texts' and is deemed completed when 'adding new texts generates no new categories or insights outside of those developed through the examination of the earlier texts'.[462] The first step in this process is the collection and in-depth examination of texts characteristic of the specific part of the discourse under study, 'particularly [including] actors presumed to be authoritative or authorized speakers of the dominant discourse'.[463] In line with the above objectives, the specific parts of the discourse on the Libyan intervention that I am interested in studying are the wider discourse over threshold and the justificatory discourse surrounding the authorisation of coercive measures in UN Security Council Resolutions 1970 and 1973. The second analytical step encompasses an immanent critique, which draws on the

[462] Jackson, 'Critical Discourse Analysis', p. 80.
[463] Ibid.

'internal contradictions, mistakes, misconceptions and instabilities within and between texts [to] critique the discourse on its own terms, and expose the events and perspectives that the discourse fails to acknowledge or address, or indeed actively subjugates'.[464] Jackson elucidates that the objective of this 'internal critique is not necessarily to establish the "correct" or "real truth" of the subject beyond doubt',[465] but to challenge dominant interpretations and reflect possible variation of and alternative understandings within the discourse. As evidenced by Gallagher's analysis of the broader 'manifest failing' discourse, discussed in the Introduction, it is a discourse riddled with inconsistencies, ambiguity, misuse of concepts, and misunderstandings. Therefore, it is critical to look for such misconceptions, mistakes and internal contradictions between and within texts to determine the extent to which the discourse on collective action in Libya was marked by contestation over the threshold at the heart of R2P's reactive dimension.

Through examining the discourse on threshold in the proposed way, this thesis makes several contributions. First, it provides an empirical analysis of a particular case study that has not been analysed with a focus on the 'manifest failing' threshold (with the exception of Bode's cursory and flawed attempt to do so discussed in Chapter I). Second, this analysis will show evidence of the debate over threshold for coercive action to protect civilian populations from atrocity crimes and give an insight into the vocabulary of this debate. In so doing, it will shed light on states' understanding (or lack thereof) of the 'manifest failing' notion. Third, illustrating the various terminologies employed by states in debates over threshold could aid the identification of such discussions in future atrocity cases and potentially result in the issue of the 'manifest failing' host states receiving more attention from commentators. Fourth, states' justificatory discourses on Libya will be analysed for evidence of patterns that signify the presence of specific determinants and associated indicators that policy actors tend to proffer when justifying a coercive approach to addressing perceived mass atrocity crises. The findings of the latter will inform the proposed triangulation strategy for designing a common reference framework that could aid future deliberations on the 'manifest failing' threshold.

The focus on the P3

Whilst the contributions of all 15 UN Security Council member states to the discourse over threshold are considered, the views of so-called Permanent Three (P3) members of

[464] Jackson, 'Constructing Enemies', p. 397.
[465] Ibid.

the Security Council – France, the UK and the US receive special attention, manifested in the analysis of their government statements on Libya, in addition to contributions to Council debates. The rationale behind this move rests on a combination of several factors, as a result of which the P3 actively engaged with the notion of threshold when justifying their position on the Libyan crisis. Not only were they the three permanent members of the UN Security Council that voted in favour of both Resolutions 1970 and 1973 authorising coercive action in Libya, but they also headed diplomatic efforts in shaping the international community's response to the Libyan crisis. They were instrumental in guiding and framing the international reaction to suspected crimes against humanity in Libya by virtue of their influential role in the Security Council, as well as due to their active role in sponsoring and drafting key Security Council resolution texts on Libya, either collectively or individually. Notably, the P3 participated in drafting Resolutions 1970 and 1973.[466] In addition to their diplomatic role in forging international consensus, the P3 made significant military contributions during the intervention, which placed a heavy burden of justifying their approach to the Libyan crisis domestically. As the most vocal proponents in the Libyan campaign, who also took an active role on the ground, it is no surprise that the P3 were the principal contributors to the justificatory discourse over coercive action in Libya in the lead up to and the immediate aftermath of the vote on the above resolutions. The alignment of these factors makes the public justifications for the coercive response to the Libyan crisis offered by these states the most nuanced and valuable authoritative contributions to the debate over threshold. Therefore, P3 discourses on Libya's 'manifest failing' are worth exploring in detail in their own right, as well as because they can provide a frame of reference against which the views of other states on the issue of threshold can be judged.

The analysis of the views of the P2 and the Elected Ten (E10) members of the Security Council at the time of the vote on the two resolutions (Bosnia and Herzegovina, Brazil, Colombia, Gabon, Germany, India, Lebanon, Nigeria, Portugal and South Africa) is equally important, albeit less exhaustive than that of the P3. Sampling the statements of the P2 and E10 found in their justifications for their vote on Resolutions 1970 and 1973 is invaluable to mapping the entire universe of factors and concerns that states highlighted when making arguments as to whether the threshold for coercive action had been crossed. The assessment of how these states sought to publicly justify their position on the

[466] S/PV.6491, p. 2; S/PV.6498, p. 2.

adoption of coercive measures in Libya will not only allow to gauge their reaction to P3 claims that this threshold had been met, but more importantly reveal whether the determinants and indicators of 'manifest failing' professed by the P3 resonated with the rest of the Council.

Finally, the contributions of UN member states invited to take part in relevant regional and thematic meetings on Libya are also taken into account. Again, this is done to gain a better understanding of the views of a broach church of states on the threshold for coercive actions in a bid to draw a more comprehensive kaleidoscopic picture of the wider debate over threshold in the context of the Libyan crisis. In addition, considering the views of UN member states expressed in relevant thematic meetings provides an opportunity to gain an insight into the normative stance of states over how a 'manifest failing' should be determined in principle. The suggestions of member states at these meetings as to what should be done in theory are worth considering with a view to being instructive for the present endeavour to devise a normative framework of reference to guide such assessments, because they supply the wider normative context within which argumentation on host states' 'manifest failings' can take place.

Selection of Sources

For the purposes of the two-tier analytical strategy outlined above the Libyan crisis is provisionally divided into two time periods. Specifically, the period before the intervention when deliberation over coercive action took place and states actively sought to justify the adoption of coercive measures, and the successive period of the active intervention phase when sates invoked the notion of threshold in the context of normative discussions in hindsight of the Libyan crisis. The first period from 15 February to 31 March 2011 captures the events from onset of demonstrations in Benghazi through to the end of the initial phase of the military intervention in Libya launched to enforce the mandate of Resolution 1973. It encompassed the adoption of the two UN Security Council resolutions sanctioning the adoption of coercive measures against the Libyan state, Resolution 1970 on 26 February and Resolution 1973 on 17 March and the duration of military operations led by the P3 states from 19 March to 31 March 2011. The cut-off point at the end of March extends slightly beyond the adoption of Resolution 1973 authorising a no-fly zone and 'all necessary measures… to protect civilians', to allow an

examination of P3 statements in immediate aftermath of the Security Council's vote.[467] It is during this period from mid-February to the end of March 2011 that the broader debate and justificatory discourse over threshold were most intensified. The P3 states issued statements promoting, advocating and justifying the adoption of coercive measures. UN Security Council meetings on the adoption of Resolutions 1970 and 1973 saw Council member states justifying their vote on the two resolutions authorising collective against Libya. Hence, both stages of analysis – the mapping of the discourse over threshold and the identification of relevant determinants and indicators, draw extensively on the selected P3 statements and UN Security Council meetings records from this period.

The second period from 1 April to 22 December 2011 is included to supplement the first stage of analysis by expanding the sample of relevant statements in a bid ensure the views of the broader international community are better represented, especially those of Security Council members other than the P3 and the wider UN membership. The above period encompasses the seven-month intervention in Libya under the NATO-led Operation Unified Protector from 1 April to 31 October 2011, which saw the unanimous adoption of another four resolutions.[468] It extends beyond the end date of international military action until 22 December 2011 (when the last UN Security Council meeting on the situation in Libya in 2011 was held) to allow for the examination of state discourses on the threshold for coercive action through to immediate aftermath of the intervention.

In both periods, the analysis encompasses the elite-level discourse on Libya depicted in official documents and formal public statements of states that shaped the international response to the Libyan crisis. It focuses on key contributors to the discourse on the threshold for international action over the course of the Libyan crisis, specifically the P3 states who fronted the Libyan intervention, the P2, E10 who deliberated on and subsequently decided to mandate coercive measures in Libya, and other UN member states that contributed to the selected Security Council meetings.

The contributions of P2, E10 and UN member states are drawn from Security Council debates only. The choice of UN documents is narrowed down to the official record of UN

[467] S/RES/1973 (2011).

[468] UNSC, S/PV.6620, 16 September 2011; UNSC, S/RES/2009, 16 September 2011; UNSC, S/RES/2016, 27 October 2011; UNSC, S/RES/2017, 31 October 2011; UNSC, S/RES/2022, 2 December 2011.

Security Council deliberations, because they capture the discourses within the Security Council in a way that other documents, such as the products of these debates, UN Security Council Resolutions and Presidential Statements, do not. The relevant Security Council meeting records are identified through a search of the index to proceedings of the Security Council for the year 2011.[469] They comprise specific meetings on Libya under the agenda item 'The Situation in Libya', regional meetings under the rubric 'Peace and Security in Africa' where the Libyan crisis was discussed within the context of the broader geographic region, and thematic meetings held on topics relevant to the threshold for international action, namely 'Protection of Civilians in Armed Conflict'. The latter, in the words of Emergency Relief Coordinator Ms. Valerie Amos, 'are an important opportunity for focused discussion of the protection of civilians in a number of situations and serve to convey some sense of what we see in conflicts, which is marked by the consistent failure of the parties concerned to comply with their legal obligations to respect and protect civilians'.[470] This is what makes debates on the protection of civilians a perfect opportunity to explore states' views on what it means for a state to be 'manifestly failing' in its primary responsibility to protect its populations. Whilst some of these meetings include contributions from a range of actors (e.g. diplomats, and senior officials within the UN and regional organisations), only the statements made by state representatives to the UN will be taken into account in line with the goal to draw on formal expressions of state views on the threshold for coercive action.

Respectively, the P3 states' approach to the threshold for international action is examined in their statements to the Security Council of the meeting records specified above, as well as in public statements and remarks made by the three states outside the Council, including government statements, press releases/statements, and statements in parliamentary debates. The analysis of P3 justificatory discourse outside the Council focuses on public statements produced by the authoritative speakers of the dominant discourse that legitimised coercive action in Libya, namely the heads of state and government of the three states – US President Barack Obama, UK Prime Minister David Cameron, the French President Nicolas Sarkozy and Prime Minister François Fillon, and where relevant other senior government officials.

[469] Index to proceedings of the Security Council: 66th year (2011). [Accessed 05 November 2019] Available at: http://research.un.org/en/docs/sc/quick/meetings/2011

[470] UNSC, S/PV.6650, 9 November 2011, p.7.

The statements made by the P3 outside the Council will be identified through a search of official government websites containing archives of government statements. For the US, the necessary documents will be obtained from 'The White House: President Barack Obama' [471] which provides a comprehensive archive of Barack Obama's statements during his presidency. For the UK – the official site of the UK government for government statements[472] and the website of the UK parliament for publications of parliamentary debates.[473] For France, the websites of the French Embassy in the US and the UK served as repositories of government statements,[474] whereas the website of the French Assembly was searched for parliamentary debates on the Libyan crisis.[475] In some cases where the full texts of statements were not available from the above repositories, they were obtained from other online sources, such as the websites of the Guardian and the Washington Post.

Initially, all state contributions in all selected texts are to be considered to map the debate over threshold in the Libyan context. The examination of P3 statements will be succeeded by an examination of statements by the P2, E10, and eventually, the wider UN memberships in the selected thematic, regional and specific meeting records on Libya. Thematic and regional debates will be particularly valuable in examining the debate over threshold because they include contributions of states from the wider UN membership, invited to participate, without the right to vote, in Security Council discussions as per rule 37 of the Council's provisional rules of procedure.[476] The selected documents will be read for references to the international community's responsibility to protect, the 'unable or unwilling,' 'manifest failing' and other terminologies used to refer to the notion of a threshold. Thematic meetings will also be read for references to Libya in order to distinguish statements made specifically about the Libyan crisis from general remarks on the situation in the region or other states within the region (e.g. Syria).

[471] See Obama White House website [Accessed 29 January 2020] Available at: https://obamawhitehouse.archives.gov/

[472] See Government UK website [Accessed 29 January 2020] Available at: https://www.gov.uk/

[473] See UK Parliament website [Accessed 29 January 202] Available at: https://www.parliament.uk/

[474] See French Embassy in the US website [Accessed 29 January 2020] Available at: https://franceintheus.org/ and French Embassy in the UK website [Accessed 29 January 2020] Available at: https://uk.ambafrance.org/

[475] See National Assembly website [Accessed 29 January 2020] Available at: http://www.assemblee-nationale.fr/13/cri/2010-2011/

[476] UNSC, Provisional Rules of Procedure of the Security Council, S/96/Rev.7, 1 January 1983. See, for instance, UNSC, S/PV.6486, 22 February 2011.

Following this first phase of documentary study, the second stage of analysis focuses on the detailed examination of selected documents in the period between 15 February and 31 March 2011 to identify the determinants and indicators that states invoked when making the argument that the threshold for collective international action had been crossed. The material from this period includes contributions from all Security Council members in the official record of UNSC deliberations over Resolutions 1970 and 1973, as well as statements on the situation in Libya made by the P3 states outside the Council in which they sought to justify the adoption of coercive measures. Again, the initial focus will be on the justificatory discourse of the P3 states, followed by an examination of the justifications offered by the P2 and E10 states.

CONCLUSION

This chapter detailed how a set of determinants and corresponding indicators will be compiled through a triangulation approach which involves integrating primary and secondary qualitative data on the factors relevant to the 'manifest failing' and 'unable or unwilling' tests. The advantage of this approach is that it draws on sources of data and areas of international law that have not been tapped into within previous investigations of the 'manifest failing' test, whilst taking into account the research findings of previous contributions to the study of what the 'manifest failing' threshold entails (reviewed in Chapter I). Combining existing knowledge with original research in this way will contribute to developing a comprehensive standalone framework providing guidelines for assessments as to whether a host state is 'manifestly failing' or not in its responsibility to protect its own populations.

Once the indicators and determinants that will shape the skeleton of the resulting framework are identified, this thesis will proceed to draw on the scholarship on mass violence and the 'unable or unwilling' threshold in order to provide guidance for the interpretation of its constitutive elements. Finally, in order to illustrate how the resulting framework could be utilised, it will be applied to a contemporary case in which mass atrocities are allegedly being committed – the Philippines' 'war on drugs'.

CHAPTER III: THE ORIGINS OF 'MANIFEST FAILING' THRESHOLD – THE 'UNABLE OR UNWILLING' TEST IN INTERNATIONAL LAW

This chapter takes a step back to relate how the 'manifest failing' threshold came into existence and trace its historical origins in search of clues as to what it may entail. This avenue of research leads to excavating the lineage of its predecessor – the 'unable or unwilling' threshold for international action in response to gross and systematic violations of human rights embedded in the R2P concept by ICISS. The ICISS report provides an entry point for foraying into an analogous area of international law and international relations – the study of the 'unable or unwilling' test*[477] in the context of self-defence against non-state actors. By drawing insights from the latter, this chapter endeavours to fill the gaps in our understanding of the threshold for collective action in mass atrocity scenarios. In so doing, it will advance three main goals:

1) unravelling the logic behind the threshold for international action and shedding more light on what the phrase 'manifest failing' should mean normatively by reference to its long-standing predecessors in international law;
2) providing important insights for establishing substantive normative guidelines for making 'manifest failing' determinations, and;
3) locating and addressing caveats in thinking about the 'unable or unwilling' test that will help resolve analogous problems with specifying the substance of the 'manifest failing' threshold.

The chapter proceeds in four sections. The first section sheds light on the emergence of the 'manifest failing' formulation and justifies tracing its historical ancestry via the 'unable or unwilling' requirement set out by ICISS. The second section introduces the 'unable or unwilling' threshold for the extraterritorial use of force against non-state actors, briefly touching upon its origins, scholarly debates on its status as a legal norm and the lack of clarity surrounding its substance. The third and last section of this chapter

[477] *The phrases 'unable or unwilling' and 'unwilling or unable' are used interchangeably in legal scholarship and state practice.

is dedicated to examining proposals for clarifying the substance of the 'unable or unwilling' test and the lessons that can be learned from debates amongst legal scholars over the factors and considerations that should inform 'unable or unwilling' determinations.

I. THE NOTION OF A 'MANIFESTLY FAILING' STATE AND ITS PREDECESSOR THE 'UNABLE OR UNWILLING' STATE

The concept of 'manifest failing' was first introduced in Paragraph 139 of the WSOD to replace the 'unable or unwilling' concept, which appeared in all previous drafts of the Outcome Document.[478] GCR2P's justification for this terminological shift, obtained by Gallagher, reads as follows:

> 'There will be no documents on this point. At the final days of negotiation, all was done very very informally with no official drafts but through discussions of a few of the key drafters. Manifest failure was a Canadian suggestion, trying to remove the subjectivity of 'unable or unwilling' that had appeared in previous drafts, and insert what they believed to be a more evidence-based standard. It was accepted without difficulty'.[479]

Despite his extensive efforts, Gallagher was unable to procure additional evidence that either confirms or denies these suggestions. GCR2P's statement seems to find some support in Pollentine's detailed account of the 2005 negotiations. Drawing on private interviews with individuals closely involved in the negotiations, Pollentine explains that 'though the unwilling/unable language commanded support from a number of states', the more sceptical majority 'were concerned to guard against unwarranted interpretation and abuses of R2P', and therefore reluctant 'to accept language which spoke more to the motivation of individual states than to an evidence-based assessment of a specific situation'.[480] Hence, the intention behind introducing the 'manifest failing' formulation 'was to define a higher threshold based upon available evidence (so a greater burden of

[478] Gallagher, 'What constitutes a "Manifest Failing"?', pp. 429, 430, 432. Marc Pollentine, *Constructing the Responsibility to Protect*, Ph.D. thesis (Cardiff University, 2012), pp. 334-337.

[479] Gallagher, 'What constitutes a "Manifest Failing"?', p. 432.

[480] Pollentine, *Constructing the Responsibility to Protect*, p. 334.

proof) rather than subjective judgements relating to the political motives of a government'.[481]

But, as Gallagher suggests, 'even if one accepts the view set out by GCR2P, it is not clear (a) how the phrase "manifest failing" is more objective than "unable or unwilling", (b) what evidence is required to prove a "manifest failing" and (c) what the unintended implications of this terminology change may be'.[482] Consequently, as he discovers, this line of enquiry does not yield any definitive answers to the question of 'What constitutes a "manifest failing"?', as those who conceived of the 'manifest failing' formulation during the final stages of negotiation preceding the adoption of the 2005 Outcome Document offered little guidance on how to determine whether this threshold has been crossed.[483] What we do know, however, is that the threshold for collective action was represented by the phrase 'unable or unwilling' until the late stages of the 2005 drafting process, when it was introduced to replace the language of 'unable or unwilling' implanted in the notion of R2P as originally articulated by ICISS in 2001.

The notion of an 'unable or unwilling' state played an important part in ICISS' conception of the R2P and featured prominently in the 2001 report. The Commission placed it at the heart of R2P at the first mention of this new concept:

> '"The Responsibility to Protect", the idea that sovereign states have a responsibility to protect their own citizens from avoidable catastrophe – from mass murder and rape, from starvation but that when they are *unwilling or unable* to do so, that responsibility must be borne by the broader community of states' (emphasis added).[484]

Later, Paragraph 2.25 of the report specifies:

> 'The emerging principle in question is that intervention for human protection purposes, including military intervention in extreme cases, is supportable when major harm to civilians is occurring or imminently apprehended, and

[481] Ibid., pp. 336-337.
[482] Gallagher, 'What constitutes a "manifest failing"?', p. 432.
[483] Ibid., p. 430.
[484] ICISS, 'The Responsibility to Protect', p. VIII.

> the state in question is *unable or unwilling* to end the harm, or is itself the perpetrator' (emphasis added).[485]

Likewise, in Paragraph 2.29 the Commission clarifies:

> '[T]he responsibility to protect acknowledges that the primary responsibility in this regard rests with the state concerned, and that it is *only if the state is unable or unwilling to fulfill this responsibility, or is itself the perpetrator, that it becomes the responsibility of the international community to act in its place'* (emphasis added).[486]

From the above selection of quotes, it is clear that the 'unable or unwilling' concept plays a central role in determining which avoidable humanitarian catastrophes warrant intervention by the international community and which do not. Commenting on the second of the abovecited excerpts of the ICISS report, Gallagher points out that 'The statement highlights that the phrase "unable or unwilling" goes right to the very heart of the R2P'.[487] According to him, it suggests that '[i]f, for whatever reason, the state in question was judged to be "unable or unwilling" to protect its population, then, and only then, could the UN Security Council act without the state's consent'.[488] Likewise, the third of the passages cited above clearly specifies that the 'unable or unwilling' requirement functions as a critical marker of the transfer of protection responsibilities from the host state to the international community. The 'unable or unwilling' concept is integral to the shift in terminology and perspective that the Commission sought to prompt in the debate about humanitarian intervention to 'focus not on the "right to intervene" but on "the responsibility to protect"'.[489] Unlike the 'intrinsically more confrontational' language of the 'right or duty to intervene', ICISS conceived the '"responsibility to protect" [as] more of a linking concept that bridges the divide between intervention and sovereignty'.[490] The notion of 'unable or unwilling' does precisely that by making external intervention by the international community conditional upon the way in which the host state discharges its primary responsibility to protect its populations. In this

[485] Ibid. para. 2.25.
[486] Ibid. para. 2.29.
[487] Gallagher, 'What Constitutes a "Manifest Failing"?', p. 431.
[488] Ibid.
[489] ICISS, 'The Responsibility to Protect', para. 2.29.
[490] Ibid.

context, the report recognises that '[i]n many cases, the state will seek to acquit its responsibility in full and active partnership with representatives of the international community'.[491]

The precise function of the 'unable or unwilling' threshold crystallises later in the report, in the first paragraph of Section Four, detailing 'the responsibility to react', the text of which is will be replicated in full:

> 'The "responsibility to protect" implies above all else a responsibility to react to situations of compelling need for human protection. When preventive measures fail to resolve or contain the situation and *when a state is unable or unwilling to redress the situation, then interventionary measures by other members of the broader community of states may be required.* These *coercive measures* may include political, economic or judicial measures, and in extreme cases – but only extreme cases – they may also include military action. As a matter of first principles, in the case of reaction just as with prevention, less intrusive and coercive measures should always be considered before more coercive and intrusive ones are applied' (emphasis added).

From this paragraph, it becomes clear that the 'unable or unwilling' threshold performs a function similar to that of the 'manifest failing' threshold found in the WSOD. It is not just a threshold that signifies a transfer of protection responsibilities, but a threshold for the use of coercive 'interventionary' measures by the international community. By contrast with the WSOD's threshold for international action (the inadequacy of peaceful measures plus the 'manifest failing' requirement), ICISS' threshold for R2P's reactive dimension comprises the 'unable or unwilling' requirement coupled with the failing of preventative measures (associated with the 'responsibility to prevent' specified in the report).[492] Although the composite conception of the threshold for collective action in the ICISS report and the WSOD differ on account of the above disparity, the 'manifest failing' and 'unable or unwilling' requirements remain substantively alike.

With regards to coercive action, the measures envisioned in the ICISS report are consistent with the measures under Chapter VII of the UN Charter, specified in Paragraph

[491] Ibid.

[492] For the 'responsibility to prevent' see: Ibid. pp. 19-23.

139 of the WSOD, although the requirement for Security Council authorisation is absent from the ICISS report. The latter also makes an additional stipulation that the bar for coercive measures short of the use of force is lower than the 'threshold or "trigger" conditions' that need to be met before contemplating military intervention.[493] This idea of progressively more demanding trigger conditions to match increasingly intrusive measures is hardly controversial. Notably, it finds support in Ban Ki-moon's 2009 implementation report, which suggests that the threshold for Pillar III action should be higher than the threshold for Pillar II action, as well as that within Pillar III the threshold for sanctioning Chapter VII measures should be higher than that for Chapter VI measures.[494]

Judging from the text of the 2005 Outcome Document and the 2001 ICISS report, there is considerable continuity between the 'unable or unwilling' and 'manifest failing' substantive threshold requirements, even though these two conceptions of the R2P notion differ meaningfully in other ways. This conclusion was reaffirmed by my conversation with one ICISS Commissioner, who noted that the idea of a threshold expressed as 'manifest failing' in the WSOD is consistent with the notion of 'unable or unwilling' introduced by ICISS.[495] This continuity extends to the substantive clarity of the two thresholds. Like the WSOD, the ICISS report does not offer much guidance as to what the 'unable or unwilling' concept entails, nor guidance as to the factors that could indicate that the threshold it represents has been crossed. According to the aforementioned ICISS Commissioner, the Commission decided not to specify what the threshold may entail and instead 'used general words, [so as] to increase the chances of the Security Council being interested without putting parameters and limits defining where the Security Council gets interested and where it does not'.[496] This decision was informed by three considerations: 1) 'the Security Council is first and foremost a political body' and its decision would ultimately be contingent upon a balance of political calculations, reflective of the national interests of its constitutive members; 2) such political decision would always 'depend on the nature of the crisis, the location of the crisis, [and] the context in which it was occurring [to determine] what constitutes mass atrocities'; 3) and because one cannot 'anticipate every possible contingency' and its nuances, 'to specify something today,

[493] Ibid., paras. 4.1 and 4.2.

[494] A/63/677 (2009), p. 22.

[495] Private conversation with the author, 6 July 2018.

[496] Ibid.

carried the risk of missing out some of the details of the crisis as it actually occurred'.[497] The Commissioners had good reasons to avoid articulating any rigid standards as to what the 'unable or unwilling' test entails, which might have been perceived as intrusive or limiting on the former's decision-making prerogative. Thus, 'telling the Security Council exactly what to do, what the threshold might be'[498] would have jeopardised the positive reception of the report's proposals.

Even though the ICISS report did not seek to answer the questions surrounding the substance of the 'unwilling or unable' threshold requirement, it provided a hint as to its origin and consequently a trail to excavate its historical lineage further in search of clues for the substantive factors that should inform the 'manifest failing' test. Specifically, paragraph 2.10 of the ICISS report states:

> 'The established and universally acknowledged right to self-defence, embodied in Article 51 of the UN Charter, was sometimes extended to include the right to launch punitive raids into neighbouring countries that had shown themselves *unwilling or unable* to stop their territory from being used as a launching pad for cross-border armed raids or terrorist attacks' [emphasis added].[499]

Unlike the rest of the report, the 'unable or unwilling' phrase in the above paragraph was used to represent the threshold for the extraterritorial use of force in self-defence against non-state actors. In response to a question as to whether the text of the above passage is an indication of the origins of the 'unable or unwilling' phrase, the abovementioned ICISS Commissioner confirmed that their proposal to include the 'unable or unwilling' formulation in the original conception of the R2P was inspired precisely by the use of the phrase in the context of self-defence against non-state actors. To put this into context, they cited the infamous 9/11 attacks organised by al-Qaeda, as a case in which 'the government in power in Afghanistan did not take adequate measures to prevent groups based on its soil from launching an attack on the United States'.[500] As they explained, 'if

[497] Ibid.

[498] Ibid.

[499] ICISS, 'The Responsibility to Protect', para. 2.10.

[500] Ibid. The Commissioner clarified that even though 'there was no suggestion at the time… that the Taliban government of Afghanistan was itself directly involved in planning or executing the attacks on the United States.... the overwhelming sentiment, not just in the West,… was that the attack did constitute an armed attack on the United States, had been carried out by al-Qaeda, al-Qaeda had an Osama bin Laden affiliation, they were based in Afghanistan, the government of Afghanistan had tolerated their

there is a large-scale attack from one state and the circumstances are such that it is simply not conceivable that the host state would have been unaware of what was being done, which is the case in Afghanistan – they knew the nature of the regime, they knew exactly whom they were hosting and the rhetoric of belligerents coming from these groups was clear –' then the host state is either 'unwilling' or 'unable'.[501] It is worth noting that the war against Afghanistan initiated in 2001 is 'the only precedent having given rise to an implicit general endorsement of a broad interpretation of Art. 51' of the Charter.[502] According to Corten, the precedent could be interpreted in line with two different arguments commonly used to justify the extraterritorial use of force against a non-state actor: 1) that the host state has sent the non-state actor or 'has been "substantially involved" in its activities', i.e. 'the Taliban regime was "substantially involved" in the activities of Al Qaeda'; or 2) there is proof of 'another form of complicity between the host state and the non-state actor', 'if we prefer to characterise [this case] as a mere situation of complicity'.[503]

Likewise, the argument that Israel has a right to act in self-defence in response to armed attacks carried out by non-state groups from neighbouring territories, where the state concerned is 'unable or unwilling' to confront them, was also brought up by the aforementioned ICISS Commissioner, in a private conversation with the author, as an illustration of the context from which the 'unable or unwilling' formulation was borrowed. Israeli counter-terrorist operations in Lebanon in response to attacks perpetrated by Palestinians and other anti-Israeli groups have also been discussed by O'Connell as representative of a group of cases in which it is argued that 'where a territorial state is unable to prevent on-going attacks, some limited force may be used to prevent future attacks' (emphasis added), despite the fact that 'the attacks are not attributable to the territorial state'.[504] However, as Corten underscores that situation did

---

presence knowing the nature of the regime without being directly involved in the attack and therefore the right to self-defence included the right to go into Afghanistan'.: Ibid.

[501] Ibid.

[502] Olivier Corten, 'Has Practice Led to an "Agreement Between the Parties" Regarding the Interpretation of Article 51 of the UN Charter?' in Anne Peters and Christian Marxsen (eds) *Self-Defence Against Non-State Actors: Impulses from the Max Planck Trialogues on the Law of Peace and War*, Max Planck Institute for Comparative Public Law & International Law (MPIL) Research Paper No. 2017-07 (MPIL, 2017): 14-16, p. 15. [Accessed 16 November 2020] Available at SSRN: https://ssrn.com/abstract=2941640 or http://dx.doi.org/10.2139/ssrn.2941640.

[503] Ibid.

[504] Mary Ellen O'Connell, 'Evidence of terror', 7 (1) *Journal of Conflict and Security Law* (2002): 19-36, p. 34.

not set a precedent, as '[t]he wars launched by Israel against Lebanon (2006) and in Gaza (in 2009 and 2014) were widely condemned, and not only because of the disproportional character of the ripostes'.[505]

The statements and examples provided by the abovementioned ICISS Commissioner with regards to the origins of the 'unable or unwilling' formulation in the ICISS report affirm that the 'unable or unwilling' threshold has migrated from the context of self-defence against non-state actors to that of the R2P in the way described in Chapter II, via intentional norm migration (see section 'Extending the "unable or unwilling" test'). Hence, in addition to the substantive parallels Chapter II already detailed between these two contexts in which the 'unable or unwilling' test has been employed, the above discussion provided further evidence of the compelling genealogical basis for drawing on the substance of the 'unable or unwilling' test in the present investigation of the 'manifest failing' threshold. In so doing, it attested to the strength of this analogy.

II. THE 'UNABLE OR UNWILLING' TEST IN THE CONTEXT OF SELF-DEFENCE AGAINST NON-STATE ACTORS

The 'unable or unwilling' requirement has recently gained prominence in the polemics on the rules of international law regulating the use of force in international relations, whose limits are increasingly contested in light of the contemporary security threats posed by non-state actors.[506] The 'unable or unwilling' lexicon has become an invariable part of the conversations surrounding the permissibility and legality of extraterritorial use of force in self-defence against a threatening non-state actor located in a host state that is ineffectual in supressing this threat. This spike in interest can be attributed to the fact that the right of self-defence against non-state actors is not provided for in the UN Charter and thereby has largely been left out of mainstream academic debate until perhaps the most

[505] Corten, 'Has Practice Led to an "Agreement Between the Parties" Regarding the Interpretation of Article 51 of the UN Charter?', p. 15. For instance, in a statement of September 2006 the Non-aligned Movement 'denounc[ed] a "relentless Israeli aggression", without any reference to the criterion of proportionality, and similar statements were issued to denounce the military interventions in Palestine'.: Ibid. For further analysis of this case in relation to the 'unable or unwilling' test see: Ahmed, 'Defending Weak States', p. 22.

[506] For a discussion of the so-called new security threats and international law see Allen S. Weiner, 'The Use of Force and Contemporary Security Threats: Old Medicine for New Ills?', 59 (2) *Stanford Law Review* (2006): 415-504. See also: Danish Institute for International Studies (DIIS), *New Threats and the Use of Force* (Copenhagen: Danish Institute for International Studies, 2005). [Accessed 16 July 2020]. Available at: https://www.files.ethz.ch/isn/16807/new_threats_whole_web.pdf. For an explicit focus on the 'unable or unwilling' concept under existing international law see Elizabeth Wilmshurst, 'The Chatham House Principles of International Law on the Use of Force by States in Self-Defence', 55 (4) *The International and Comparative Law Quarterly* (2006): 963-972.

harrowing act of violence by a non-state actor – the abovementioned 9/11 attacks conducted by al-Qaeda.[507] Notwithstanding the rising prominence of the 'unable or unwilling' terminology in legal discourses since then, the phrase has been in use in legal debates surrounding the use of force since before the adoption of the UN Charter.

The origins of the 'unable or unwilling' test

The most thorough investigation of the roots of the 'unable or unwilling' test has been conducted by Deeks, whose notable 'normative framework for extraterritorial self-defence' was introduced in Chapter II and is to be discussed in the next section. In addition to examining state practice in a bid to identify a core set of normative factors that ought to inform 'unable or unwilling' inquiries, Deeks also 'identif[ies] the law of neutrality,… developed in situations of international armed conflict between states, as the original source of the ['unable or unwilling'] test'.[508] For Deeks, '[n]eutrality law offers a useful starting point from which to understand a rule that allows one state to use force on another state's territory against a third entity in certain circumstances [and] anchors the [unable or unwilling] test's legitimacy'.[509] O'Connell sees the effort to draw on this analogy, linking the unable or unwilling test to existing law, as superfluous, 'given that international law has directly applicable rules and princip1es to govern State uses of force against non-State actors',[510] a position which is consistent with her principled scepticism of the 'unable or unwilling' test. Despite the objections as to the way she uses this analogy, Deeks' exploration of the origins of this threshold requirement is a well-supported and interesting theory that is worth elaborating on for the purpose of shedding light on the movements of the 'unable or unwilling' test between various contexts, how the substance of this threshold has evolved through time and acquire knowledge that may help us to gain a more complete insight into its use in the context of the extraterritorial self-defence against non-state actors. Without pursuing a comprehensive account of the

[507] Britta Sjöstedt, 'Applying the Unable/Unwilling State Doctrine – Can a State Be Unable to Take Action?', in Anne Peters and Christian Marxsen (eds) *Self-Defence Against Non-State Actors: Impulses from the Max Planck Trialogues on the Law of Peace and War*, Max Planck Institute for Comparative Public Law & International Law (MPIL) Research Paper No. 2017-07 (MPIL, 2017): 35-38, p. 36. [Accessed 16 July 2020] Available at SSRN: https://ssrn.com/abstract=2941640 or http://dx.doi.org/10.2139/ssrn.2941640.

[508] Deeks, '"Unwilling or Unable"', p. 496.

[509] Ibid. p. 497

[510] Mary Ellen O'Connell, 'Self-Defence, Pernicious Doctrines, Peremptory Norms' in Mary Ellen O'Connell, Christian J. Tams, and Dire Tladi. *Self-Defence against Non-State Actors*, Max Planck Trialogues (Cambridge: Cambridge University Press, 2019): 174–257, p. 226.

wide-ranging laws of neutrality, it is worth laying out the basics for the sake of the ensuing discussion.

In their most basic form, as found in customary international law and the 1907 Hague Conventions, the laws of neutrality establish rules concerning states that do not participate in an armed conflict.[511] These rules can provisionally be divided in two categories: 1) rules protecting the rights of neutral states by ensuring that they incur minimal damage as a result of conflict, and 2) rules protecting the rights of belligerent states by guaranteeing that the neutrals respect the principles of non-participation in the conflict and non-discrimination between the warring parties (including the prohibition of activities, such as the passage of war materials and belligerents through the neutral state's territory, enrolment in belligerent armed forces, and providing military supplies to belligerent parties).[512] It is important to note that '[n]eutrality is not contingent upon any declaration of neutrality, any state that is not a Belligerent Party is considered neutral and the laws of neutrality then apply (subject to being overridden by a binding decision of the Security Council…)'.[513] Naturally, the occurrence of situations in which a state fails to uphold its neutral obligations effectively, whether they are ignoring their duties or simply lack the capacity to prevent a belligerent from using their territory to engage in war-related activities, is to be anticipated.[514]Although such instances of non-compliance are not provided for in treaty law, legal commentators, and subsequently, state military manuals acknowledged that in line with neutrality law it is permissible for the affected belligerent state to use force in self-defence on a neutral state's territory, if the latter proves *'unable or unwilling'* to preclude an opposing belligerent from violating its neutrality.[515] Whether the 'unable or unwilling' test has emerged as a customary law norm in this context, as some claim (see examples provided by Deeks), or not, Deeks maintains that it is at the very least a 'well-entrenched' norm as evidenced by state practice, exemplified by references 'to the "unwilling or unable" test in the law of neutrality' in official UK, US and Canadian military manuals.[516]

---

[511] See Yoram Dinstein, *War, Aggression and Self-Defense*. 3rd Edition. (Cambridge: Cambridge University Press, Virtual Publishing, 2003), pp. 169-81, esp. p. 125 (discussing the basic characteristics of the law of neutrality).

[512] Ibid.

[513] Ibid.

[514] Deeks, '"Unwilling or Unable"', p. 499.

[515] Ibid.

[516] Ibid. pp. 499-500.

Soon after establishing its presence 'in the context of international armed conflict [the 'unable or unwilling' test] migrated into the rules governing a state's use of force extraterritorially against nonstate actors'.[517] Initially, some states (concerned with maintaining their neutral status) brought the laws of neutrality into their domestic sphere in order 'to prohibit [their] citizens' to take actions against a state they are not in conflict with, in breach of the 'laws of neutrality'.[518] This is how, according to Deeks, the test was extended 'to govern acts by nonstate actors during peacetime'.[519]

Deeks's investigation shows that the scholarship on the laws of neutrality recognises that when a neutral state is either unwilling or lacks the capacity 'to prevent violations of its neutrality by another belligerent', it is permissible for the offended belligerent to use force on the territory of the 'unable or unwilling' neutral state 'to prevent violations of its neutrality by another belligerent'.[520] The host of direct quotations amassed in Deeks' footnotes illustrate that early commentators employed a range of alternative formulations to acknowledge the application of what later became widely known as the 'unable or unwilling' test under neutrality law. Some speak of neutrals who 'cannot or will not enforce [their] rights', who 'ha[ve] neither the desire nor the power' / 'possess neither the power nor disposition' to resist violent belligerents using their territory as a launching pad for attacks against other states they are at war with.[521] Others, refer to such ineffective states as 'neutrals [who] acquiesce in or are unable to prevent the violation of' their neutrality, those who cannot resist such violations because of 'complaisance… or… inability, through weakness or otherwise', due to 'the helplessness of the country or by means of intrigues with a party within it'.[522] Though not always using the exact phrase (i.e. 'unable or unwilling'), commentators generally recognise that failings to uphold neutrality law can be equally attributed to either a mere lack of capacity or the intentional (in)actions of a neutral state, who conspires with / provides support to the offending belligerent or deliberately chooses not to oppose them.

Whilst Deeks' rationale for drawing on the analogous context of the 'laws of neutrality' appears to be sound and has not been met with opposition by critics, her understanding

[517] Ibid. p. 501.
[518] Ibid. p. 502.
[519] Ibid.
[520] Ibid. pp. 499-500.
[521] Ibid.
[522] Ibid.

of what can be learned or indeed claimed on the basis of this analogy is problematic. Specifically, Deeks argues that [a]lthough the law of neutrality offers a clear pedigree in international law for the "unwilling or unable" test, however, it tells us little about what standards should or do attach to that test'.[523] This rejection of seeking normative support for the 'unable or unwilling' test and the factors that should inform it by exploring the way in which it is interpreted in the analogous context of the laws of neutrality is rather odd. After all, as discussed in the previous chapter, this is precisely the way in which the method of reasoning by analogy has been employed in international law, resulting in its expansion. This critique finds support in Heller's discussion of Deeks' methodology, where, contra Deeks, he maintains that the law of neutrality 'may provide *normative* support for the "unwilling or unable" test in the context of attacks against NSAs [non-state actors], but it does not provide *legal* support for it'.[524] To dispel any ambiguity as to what the first part of this claim entails, the normative contribution Heller is referring to in his article is Deeks' 'propos[al for] a core set of substantive and procedural factors that should inform the "unwilling or unable" inquiry'.[525] That said, Heller is less concerned with Deeks' failing to recognise how she could draw on the laws of neutrality to support her normative framework than with her problematic claim that the legal status of the norm can be inferred by relating it to the venerable law of neutrality. Heller is quick to point out that even by her own admission, 'neutrality law does not directly govern uses of force between states and non-state actors', although the two contexts are analogous.[526] He also sheds light on 'another problem with using the law of neutrality to support the "unwilling or unable" test [(recognised but unaddressed by Deeks)]: that [neutrality] law predates the adoption of the UN Charter, which strictly regulates the use of interstate force'.[527] As the ensuing discussion elucidates, there are a number of other reasons why Deeks' claim that the 'unable or unwilling' test migrated into the context of the use of force against non-state actors carrying over its customary law status, established in the context of neutrality law, does not find support amongst scholars of international law.

[523] Ibid. p. 497.

[524] Kevin Jon Heller, 'Ashley Deeks' Problematic Defense of the "Unwilling or Unable" Test', 15 December 2011 [Accessed 21 July 2020] Available at: http://opiniojuris.org/2011/12/15/ashley-deeks-failure-to-defend-the-unwilling-or-unable-test/

[525] Ibid.

[526] Deeks, '"Unwilling or Unable"', p. 497.

[527] Ibid.

Is the 'unable or unwilling' test part of customary international law?

The legal debates surrounding the 'unwilling or unable' threshold are underpinned by the following question: 'How can an actor A (a State) invoke a use of force attributable to an actor B (an NSA) to justify using force against another actor C (a State whose territory is used by the NSA)?'.[528] In this regard, the proposition made by proponents of the 'unable or unwilling' test (see ensuing discussions of Bethlehem and Deeks' views) that 'when a state is "unable or unwilling" to control terrorism, the legal requirements of attribution [(i.e. showing that the attacks are attributable to the host state)] or consent no longer apply' has attracted criticism from prominent legal scholars, such Mary Ellen O'Connell.[529] Accordingly, as Corten frames the issue,

> '[t]he problem…lies not in the invocation of self-defence in the case of an armed attack launched by an NSA, but rather in the possibility to invoke self-defence against a State which is not itself the direct author of the armed attack, and is therefore entitled to the protection of its sovereignty and its political independence under Art. 2.4 of the UN Charter'.[530]

The 'unable or unwilling' test has become the topic of many a legal debate concerning the extraterritorial self-defence against non-state actors. Pivotal among them is the debate as to whether it is part of customary international law, which features prominently in discussions of its substance. It is widely acknowledged that customary international law has two elements: 1) *state practice* ('an objective requirement that concentrates on state behaviour'); and 2) *opinio juris* ('a requirement that focuses on a state's subjective belief that they are bound by a particular rule of international law [i.e.] the customary rule in question').[531] Opinions as to whether there is sufficient evidence that these prerequisites for the acceptance of the 'unable or unwilling' test diverge greatly among legal scholars.

---

[528] Corten, 'Has Practice Led to an "Agreement Between the Parties" Regarding the Interpretation of Article 51 of the UN Charter?', p. 15. Corten highlights four arguments used to justify the use of force against non-state actors, namely 1) '[e]stablishing that the territorial State has "sent" the NSA, or has been "substantially involved" in its activities', 2) 'proving another form of complicity between' the host state and the non-state actor, 3) deeming the host state to be 'objectively responsible' for the actions of the non-state actor, which 'correspond[s] to a situation of "inability" or of a "Failed State" [and 4)] invoking necessity as the only criterion justifying the intervention against the NSA, without even searching to establish a form of responsibility of the territorial State'.: Ibid.

[529] O'Connell, 'Self-Defence, Pernicious Doctrines, Peremptory Norms', pp. 212-213.

[530] Ibid. pp.14-15.

[531] Guzman and Meyer, 'Customary International Law in the 21st Century', p. 199.; See also Guzman and Meyer, 'International Common Law', p. 526.

Without aiming to resolve this dispute, addressing it here is key to making sense of the propositions regarding the test's content and will inevitably crop up in discussions of the substance of the test later in this chapter.

As a proponent of the 'unable or unwilling' test, Deeks claims that '[m]ore than a century of state practice suggests that it is *lawful* for State X, which has suffered an armed attack by an insurgent or terrorist group, to use force in State Y against that group if State Y is *unwilling or unable* to suppress the threat' (emphases added).[532] She makes this claim in the context of her analysis of the widely discussed US military operation to capture and kill Osama bin Laden, which took place on 2 March 2011 on Pakistan's territory, purportedly without the latter's consent. Specifically, she finds a compelling illustration of the use of the 'unable or unwilling' test in President Obama's justification for taking actions against high-value al-Qaida officials in a statement made in August 2007 when he was still a presidential candidate: 'if we have actionable intelligence against bin Laden or other key al-Qaida officials … and Pakistan is *unwilling or unable* to strike against them, we should' (emphasis added).[533] However, subsequent claims concerning the test in justifications offered in the aftermath of the military operation in question were markedly different. Having initially justified the legality of this action as 'consistent with [the US'] inherent right to self-defense under international law'[534], the Obama Administration later reasoned that killing US citizens linked to Al-Qaeda is congruent with international law if they are found to '"pose an imminent threat of attack to the United States" in host states without host state consent if that state is "unable or unwilling" to suppress the "threat"'.[535] As Ahmed points out, '[n]o such limitation has been specified for the killing of non-state actors who may not be United States citizens – that is, the overwhelming majority of non-state actors targeted', as is the case with the killing of Osama bin Laden.[536] Nonetheless, commentators have continued to interpret

[532] Deeks, '"Unwilling or Unable"', p. 486.

[533] Andy Merten, Presidential Candidates Debate Pakistan, MSNBC. 28 February 2008. Available at: http://tinyurl.com/78paob3 , quoted in Deeks, '"Unwilling or Unable"', p. 486.; Obama made a similar statement on the same day, namely 'If the United States has al-Qaida, bin Laden, top level lieutenants in our sights and Pakistan is unable or unwilling to act, then we should take them out': AFP News Agency, 'Obama Vows to "Take Out" Terror Targets in Pakistan', 28 September 2008. [Accessed 11 January 2021] Available at: https://www.youtube.com/watch?v=LqOncFQUZsw

[534] Harold Hongju Koh, 'The Lawfulness of the U.S. Operation Against Osama bin Laden', 19 May 2011. [Accessed 26 July 2020] Available at: http://opiniojuris.org/2011/05/19/the-lawfulness-of-the-us-operation-against-osama-bin-laden/

[535] Ahmed, 'Defending Weak States', p. 2.

[536] Ibid.

this case as providing support for the use of the 'unable or unwilling' test to justify the legality of the use of force, despite serious doubts being raised over the matter in light of Pakistan's objection to the intervention on the grounds that it was a 'violation of sovereignty'.[537] For instance, Goldsmith insists that the 'unable or unwilling' test 'was almost certainly the one the United States relied on' in support of the claim that even though 'it is not settled in international law, it is sufficiently grounded in law and practice that no American president…could refuse to exercise international self-defense rights when presented with a concrete security threat in this situation'.[538]

Given the strong opposition to claims that the 'unable or unwilling' test has been established in international law in the above two examples, it is not surprising that Deeks has come under fire for her portrayal of the 'unable or unwilling' test as established in customary international law. In a book published six years after Deeks' article, O'Connell notes that 'only five States have attempted to justify a use of force by invoking "unable/unwilling" and those States have invoked other justifications in addition to "unable/unwilling"'.[539] In other words, even when the 'unable or unwilling' test is invoked, it is difficult to claim with certainty that it legitimated the use of force under international law or whether this function was rather served by the other legal justifications it was intermingled with. Besides, even though Deeks initially suggests that the test appears far more frequently in state practice, it later transpires that she has only identified five states that have explicitly mentioned it.[540] When one consults 'Appendix I' at the end of her article, it is clear that out of the thirty-nine identified cases of non-consensual uses of force against a non-state actor on host state territory, there are only ten in which one of five victim state (Israel, Russia, US, UK or Turkey) has 'specifically invoked the "unwilling or unable" test or a closely related concept' to justify its use of force.[541] Besides, while O'Connell acknowledges that 'the UK Attorney General has used [the terms 'unable' and 'unwilling'] and they appear in letters to the Security Council with respect to the use of force in Syria by Australia, Canada, Turkey and the United

[537] Owen Bowcott, 'Osama bin Laden death: Pakistan says US may have breached sovereignty', 5 May 2011 [Accessed 24 July 2020] Available at: https://www.theguardian.com/world/2011/may/05/osama-bin-laden-pakistan-us-sovereignty

[538] Jack Goldsmith, 'Fire When Ready', 20 March 2020 [Accessed 24 July 2020] Available at: https://foreignpolicy.com/2012/03/20/fire-when-ready/

[539] O'Connell, 'Self-Defence, Pernicious Doctrines, Peremptory Norms', p. 226.

[540] Deeks, '"Unwilling or Unable"', p. 503.

[541] Ibid. p. 549.

States', she underscores that they 'are not used in the UN Charter, the drafts of the Charter, nor decisions of the ICJ on the use of force'.[542] On a related note, Ahmed takes issue with Deeks' lack of satisfactory engagement with two oft-cited ICJ decisions that do not support the doctrine, the *Armed Activities of Congo* and *Wall Opinion* cases, despite being delivered post-9/11 (which means that the Court's awareness of state practice and existing scholarship in the area can be assumed).[543] Notably, in both cases the ICJ opined that 'where attacks by a non-state actor cannot be attributed to a host state, the use of such force without obtaining the host state's consent would be illegal and therefore cannot be justified on grounds of ineffectiveness'.[544] As for the other element in the formation of custom O'Connell adds that '[t]he case for *opinio iuris* is almost non-existent'.[545] Deeks' admission that she has 'found no cases in which states clearly assert that they follow the test out of a sense of legal obligation (i.e., the *opinio juris* aspect of custom)' paradoxically lends support to O'Connell's assertion.[546]

With neither of the two aspects of custom adequately satisfied and a lack of engagement with cases that reject the test, Deeks ultimately fails to convince that it is lawful for a state to act in self-defence against non-state actors in another state's territory on the basis of the 'unable or unwilling' test being satisfied. Even though Deeks' approach to finding legal support for the 'unable or unwilling' test is 'methodologically unsound'[547], the debate surrounding its status is far from resolved. Ahmed's analysis illustrates that both legal scholars and the ICJ remain divided over the matter.[548] What is more, he finds that even those sceptical of the test (such as Mary Ellen O'Connell and Antonio Cassese) concede that the use of force (or limited force) in justifiable in some cases of self-defence against non-state actors launching attacks from an infective host state.[549] In other words, whilst there is considerable disagreement as to whether it is legal to act in self-defence on

[542] O'Connell, 'Self-Defence, Pernicious Doctrines, Peremptory Norms', p. 223.

[543] Ahmed, 'Defending Weak States', p. 11.

[544] Ibid. p. 9.

[545] O'Connell, 'Self-Defence, Pernicious Doctrines, Peremptory Norms', p. 226.

[546] Deeks, '"Unwilling or Unable"', p. 503. For a detailed critique of Deeks' contradictory admissions that dilute her claims about the existence of relevant practice see: Heller, 'Ashley Deeks' Problematic Defense of the "Unwilling or Unable" Test'.

[547] Heller, 'Ashley Deeks' Problematic Defense of the "Unwilling or Unable" Test'.

[548] Ahmed, 'Defending Weak States', p. 10.

[549] Ibid.

the grounds of the 'unable or unwilling' test, many are willing to concede that in some instances a case for legitimate self-defence could be made.

More recently, the debate surrounding the 'unable or unwilling' test has found prominent expression in the context of the fight against Islamic State of Iraq and the Levant (ISIL) in Syria. Some states involved in this fight (the US, Canada, Australia and Turkey) have referred to the 'unable or unwilling' test in an attempt to justify their collective and/or individual right to self-defence to prevent armed attacks by ISIL launched from Syria's territory.[550] Notably, in a letter to the UN Security Council, the US claimed that:

> 'States must be able to defend themselves, in accordance with the inherent right of individual and collective self-defence, as reflected in Article 51 of the Charter of the United Nations, when, as is the case here, the government of the State where the threat is located is unwilling or unable to prevent the use of its territory for such attacks'. [551]

The questions of whether 'the "unable or unwilling" test as… invoked by the US' has been accepted and whether the Syrian case constitutes a precedent supporting the test as employed in this context[552] has divided legal scholars. On the one hand, Scharf regards UNSC Resolution 2249, recognising ISIL as a threat to international peace and security and calling upon states 'to take all necessary measures…to prevent and suppress terrorist acts committed' by the non-state group, as a confirmation of a 'newly accepted change in the international law of self-defense […] that any State can now lawfully use force against non-state actors (terrorists, rebels, pirates, drug cartels, etc.) that are present in the territory of another State if the territorial State is unable or unwilling to suppress the threat posed by those non-state actors'.[553] In support of this argument, Scharf invokes the view of the 13th Commission of the Institute of International Law that 'where a rule of customary law is (merely) emerging or there is still some doubt as to its status, a unanimous resolution [of the UN General Assembly] can consolidate the custom and

[550] For a detailed analysis of the arguments made by these four states and others see Olivier Corten, 'The "Unwilling or Unable" Test: Has it Been, and Could it be, Accepted?', 29 *Leiden Journal of International Law* (2016): 777-799, pp. 780-785.

[551] UNSC, 'Letter dated 23 September 2014 from the Permanent Representative of the United States of America to the United Nations addressed to the Secretary-General', S/2014/695, 23 September 2014.

[552] Corten, 'The "Unwilling or Unable" Test', p. 779.

[553] UNSC, S/RES/2249, 20 November 2015.; Michael P. Scharf, 'How the War Against ISIS Changed International Law', 48 *Case Western Reserve Journal of International Law* (2016): 1-54, p. 53.

remove doubts which might have existed'.[554] In other words, although Scharf concedes that 'Resolution 2249 did not provide a new stand-alone legal basis or authorization for use of force against ISIS in Syria', it provided the last push for the 'rapid formation of a new rule of customary international law… where a context of fundamental change served as an accelerating agent, enabling customary international law to form much more rapidly, and with less State practice, than is normally the case'.[555] Nonetheless, other in-depth investigations into the situation in Syria suggest that Scharf's interpretation of Resolution 2249 is a far too generous reading of the facts.

For instance, Corten finds that none of the numerous UN Security Council and UN General Assembly statements and resolutions 'dedicated to the Syrian crisis mention Article 51 of the UN Charter, the right to self-defence, or *a fortiori* any "unwilling or unable" argument'.[556] Quite the opposite, most of these texts, including Resolution 2249, 'seem to be incompatible with any unilateral action'.[557] Nothing in Resolution 2249, in which the Council 'reaffirm[ed] the principles and purposes of' the UN Charter, its respect for state sovereignty and '[c]all[ed] upon Member States… to take all necessary measures in compliance with international law, in particular with the United Nations Charter', can be understood as a general acceptance of the claim that it is lawful to use of force in self-defence against non-state actors in a third state 'unable or unwilling' to control them.[558] Examining the 'unable or unwilling' doctrine post-Syria, Corten concludes that the Syrian case has not led to the acceptance of the 'unable or unwilling' test by the international community. The US' attempt to invoke the 'unable or unwilling' formulation in a bid to justify launching air strikes on Syrian territory against ISIL was undermined by both 'the absence of a common *opinio juris* shared by the intervening states themselves with respect to the "unwilling or unable" test', as well as 'the absence of a general acceptance of the 'unwilling and unable' test'.[559] Theodore Christakis also argues that the test 'was "unable" to make its entry in positive international law' with

[554] International Law Association, Committee on the Formation of Customary (General) International Law, Report of the Sixty-Ninth Conference, 25-29th July 2000, London, p. 64. [Accessed 20 October 2020] Available at: https://www.law.umich.edu/facultyhome/drwcasebook/Documents/Documents/ILA%20Report%20on%20Formation%20of%20Customary%20International%20Law.pdf

[555] Scharf, 'How the War Against ISIS', pp. 51-53.

[556] Corten, 'The "Unwilling or Unable" Test', p. 789.

[557] Ibid. pp. 789, 791.

[558] S/RES/2249 (2015).

[559] Corten, 'The "Unwilling or Unable" Test', pp. 780, 785.

reference to the Syrian case for very similar reasons: 1) Resolution 2249 did not provide any support for the 'unable or unwilling' test; 2) while the US-led military coalition against ISIL justified action against the non-state group 'on the basis of individual or collective self-defence, the grounds for this claim…varied' from state to state and often did not involve the 'unable or unwilling' test (e.g. in the case of Germany, France, Belgium, Norway or the Arab states), despite its invocation by the US, and; 3) 'the debates within the [UN] demonstrate that the…test was not [endorsed] by the vast majority' of UN members.[560]

That the debate remains unsettled is unsurprising, given the sustained, yet mostly unsuccessful, attempts by political actors, legal advisers and scholars of international law to justify military interventions without the consent of the host state with reference to this threshold. As Reinold argues, despite the challenge of ascertaining, based on existing state practice, whether the unwillingness or mere inability of a host state automatically provide legal grounds for the use of force in a host state, because 'the intervening states…proffered various legal rationales or, in some cases, none at all…one cannot help but note the emerging trend that states are making indiscriminate use of the unwillingness and inability scenarios to justify military action in states harboring irregular forces'.[561] The above discussion of prominent 'unable or unwilling' cases suggests that there has been a notable variance of opinion as to whether a particular action or behaviour are covered by the norm and whether or not the norm has emerged as customary law. This evidences a weak shared understanding of the 'unable or unwilling' test at best, which suggests that it is a long way from reaching the taken for granted status that a norm of international law should possess. The unsettled status of the test can also be attributed to its 'substantive indeterminacy' that Deeks and others after her have sought to resolve.

III. THE FACTORS OF THE 'UNABLE OR UNWILLING' TEST

It is not difficult to see that the lack of clarity surrounding the 'unable or unwilling' test is problematic for a number of reasons. Most importantly, Ahmed argues that '[a] lack of

[560] Theodore Christakis , 'Challenging the "Unwilling or Unable" Test', in Anne Peters and Christian Marxsen (eds.) *Self-Defence Against Non-State Actors: Impulses from the Max Planck Trialogues on the Law of Peace and War,* Max Planck Institute for Comparative Public Law & International Law (MPIL) Research Paper No. 2017-07 (MPIL, 2017): 17-20, pp. 17-18. [Accessed 16 July 2020] Available at SSRN: https://ssrn.com/abstract=2941640 or http://dx.doi.org/10.2139/ssrn.2941640

[561] Reinold, 'State weakness, irregular warfare, and the right to self-defense post-9/11', p. 285.

any substantive clarity robs the "unwilling or unable" doctrine of much of its efficacy in guiding state behaviour'.[562] On a further note, Brunnée and Toope suggest that:

> 'The more unpredictable and uncertain a supposed rule becomes, the more difficult it will be to meet the congruence requirement [that state practice is congruent with the rule]. If we do not know what the rule is, or we find that the rule is actually without constraining content (because it is self-assessed, as was the case with the "Bush doctrine"), then congruence becomes a meaningless concept; there is nothing to be congruent with'.[563]

In light of these issues, it is not surprising that a number of legal scholars (notably Deeks and Ahmed) have argued that there needs to be more clarity as to what the 'unable or unwilling' requirement entails in order to regulate the use of force in self-defence against non-state actors 'within a clear international legal framework'.[564] The rest of the chapter focuses on concrete proposals for a more determinate 'unable or unwilling' threshold requirement with a view to gaining valuable insights into how to devise a robust normative framework that could effectively guide 'manifestly failing' inquiries.

The Principles of International Law on the Use of Force in Self-Defence

The earliest attempt to make a statement about the 'proper' interpretation of the 'unable or unwilling' test under contemporary international law can be traced back to a 2006 publication by the British foreign affairs think tank Chatham House titled 'The Chatham House Principles of International Law on the Use of Force in Self-Defence'. The publication, and the principles outlined in it, were the outcome of a 2004-2005 Chatham House study, sampling the views of 'a small group of international law academics and practitioners and international relations scholars [in the UK] on the criteria for the use of force in self-defence', in a bid to provide 'a clear statement of the rules of international law "properly understood" governing the use of force by states in self-defence'.[565]

Significantly, the 'unable or unwilling' requirement is incorporated in Subsection F of the report, namely 'Article 51 is not confined to self-defence in response to attacks by

[562] Ahmed, 'Defending Weak States', p. 14.

[563] Jutta Brunnée and Stephen J. Toope, *Legitimacy and Legality in International Law* (Cambridge: Cambridge University Press, 2010), p. 307.

[564] Ahmed, 'Defending Weak States', p. 13.

[565] Wilmshurst, 'The Chatham House Principles', p. 963.

States', which specifies the conditions under which 'the right of self-defence may apply also to attacks by non-state actors'.[566] The use of force in self-defence within the territory where the hostile non-state actor is found is premised upon three preconditions: 1) the attack being large-scale; 2) that 'it must be evident that that State is unable or unwilling to deal with the non-state actors itself, and that it is necessary to use force from outside to deal with the threat in circumstances where the consent of the territorial state cannot be obtained'; and 3) that force is used 'only in so far as it is necessary to avert or end the attack'.[567] The first requirement suggests that not every attack by a non-state actor would trigger an 'unable or unwilling' determination, but a certain threshold of gravity ought to be met. The second precondition ties the 'unable or unwilling' test to the requirement to obtain consent, suggesting that the latter can be overridden if the former is 'evident'. The third requirement stipulates that the use of force against a territorial state is only justifiable for the purposes of supressing the threat posed by the violent non-state actor and should not exceed what is necessary to achieve this goal. The publication makes only one further reference to the test, reiterating that a state's inability or unwillingness to assert its authority to restrain a terrorist actor found in its territory triggers the victim state's right of self-defence against the attacker, as a last resort.[568]

Consequently, although it is clear that the 'unable or unwilling' requirement plays a pivotal role in justifying the extraterritorial use of force in self-defence from the threat of non-state actors, Chatham House does not elaborate on what it means for a state to be 'unable or unwilling'. In other words, it remains unclear as to what actions or the lack thereof would render a state 'unable' or 'unwilling'. Perhaps the reason for this lack of explanation is that Chatham House did not deem it necessary, as implied by their statement that 'it must be it must be evident that that State is unable or unwilling to deal with the non-state actors itself'.[569] Similarly to the 'manifest failing' of host state to protect its populations from atrocities, whether a host state is 'unable or unwilling' to stop an aggressive non-state actor will seldom be self-evident. In a similar vein, Deeks argues that the latter inquiry will rarely be straightforward, because 'in any of a number of cases, it will not be clear to a victim state, at least initially, whether the territorial state is

[566] Ibid. p. 969.
[567] Ibid.
[568] Ibid. p. 970.
[569] Ibid. p. 969.

unwilling or unable to suppress the threat'.[570] Despite being deficient with regards to unpacking the substance of the 'unable or unwilling' test, Chatham House's publication was pivotal in putting it on the map in discussions over the use of force in self-defence against non-state actors.

Principles Relevant to the Scope of a State's Right of Self-Defense Against an Imminent or Actual Armed Attack by Nonstate Actors

One of the contributors to the Chatham House principles – former legal advisor to the British Foreign and Commonwealth Office, Sir Daniel Bethlehem – has been singled out for playing a key role in promoting the notion that victim states may attack a non-state actor on the territory of an 'unwilling or unable' host state.[571] Indeed, in October 2012, Bethlehem made an independent proposal to expand the scope of the right to self-defence in view of the threat posed by non-state actors, which involved formulating sixteen 'principles relevant to the scope of a state's right of self-defence against an imminent or actual armed attack by nonstate actors'.[572] The intent behind his proposition was 'to address a strategic and operational reality with which states are faced, and to formulate principles that reflect, as well as shape, the conduct of states in the particular circumstances in question'.[573]

Similarly to Chatham House's proposal, the first three of his principles specify that 'states have a right of self-defense against an imminent or actual armed attack by nonstate actors', on the basis of which they can use force 'as a last resort in circumstances in which no other effective means are reasonably available to address' the threat, which should be proportionate and 'limited to what is necessary' to counter the threat.[574] The latter two prerequisites are to be expected, given that proportionality and necessity, along with immediacy, are the customary international law requirement that self-defence is subject to.[575] However, Bethlehem goes much further in specifying the concrete circumstances

[570] Deeks, '"Unwilling or Unable"', pp. 505-506.

[571] O'Connell, 'Self-Defence, Pernicious Doctrines, Peremptory Norms', pp. 212-213.

[572] Daniel Bethlehem, 'Self-Defense Against an Imminent or Actual Armed Attack by Nonstate Actors', 106 (4) *The American Journal of International Law* (2012): 769-777, pp. 775-776. Specifically, Bethlemen's endeavour was directed at 'shaping the operational thinking of those within governments and the military who are required to make decisions in the face of significant terrorist threats emanating from abroad'.: Ibid. p. 773.

[573] Ibid. p. 774.

[574] Ibid. p. 775.

[575] Ahmed, 'Defending Weak States', p. 6. For details on what these requirements involve see: Ibid.

under which a state can use force to defend itself from the threat posed by a non-state perpetrator.

Central to his understanding of when states can exercise their right of self-defence against non-state actors is his conceptualisation of what constitutes an 'imminent armed attack'. For the purposes of Bethlehem's account, the notion '"armed attack" includes both discrete attacks and a series of attacks that indicate a concerted pattern of continuing armed activity', where the 'distinction between discrete attacks and a series of attacks' may inform the critical assessments of the necessity and proportionality of the use of force in self-defence.[576] As for the 'imminence' of the attack, it ought to be:

> 'assessed by reference to all relevant circumstances, including (a) the nature and immediacy of the threat, (b) the probability of an attack, (c) whether the anticipated attack is part of a concerted pattern of continuing armed activity, (d) the likely scale of the attack and the injury, loss, or damage likely to result therefrom in the absence of mitigating action, and (e) the likelihood that there will be other opportunities to undertake effective action in self-defense that may be expected to cause less serious collateral injury, loss, or damage'. [577]

The notion of an 'imminent threat', articulated by Bethlehem, is unpacked further by Akande and Liefländer in their discussion of anticipatory self-defence. Specifically, they suggest that '[i]mminence describes a certain pressing quality that a threat must have for anticipatory self-defense to be lawful'.[578] As Chapter IV will illustrate, this understanding of imminence was reflected in states' discussions of the threat that Libyan populations would be faced with, if the international community did not take strong measures to prevent this threat from materialising.

To clarify what the concept of imminence entails, Akande and Liefländer reflect on the notion of 'threat' that it is attached to, and, its constitutive elements:

> 'A threat is a situation where a causal chain can lead from the status quo (no attack) to an undesired future (attack). The essential components of a threat are fourfold: (1) type-what kind of attack is threatened? (2) likelihood-how

[576] Bethlehem, 'Self-Defense Against an Imminent or Actual Armed Attack by Nonstate Actors', p. 775.
[577] Ibid. pp. 775-776.
[578] Dapo Akande and Thomas Liefländer, 'Clarifying Necessity, Imminence, and Proportionality in the Law of Self-Defense', 107 (3) *American Journal of International Law* (2013): 563-570, p.564.

probable is it that the attack will occur? (3) gravity-how severe will the attack be? and (4) timing-when will the attack occur?'.[579]

They find that the above four aspects are embodied in Bethlehem's definition of imminence, more specifically in three of the factors he identifies as relevant to this determination: 'the nature and immediacy of the threat', 'the probability of an attack' and 'the likely scale of the attack and the injury, loss, or damage likely to result therefrom in the absence of mitigating action'.[580] They further suggest that considering 'whether the anticipated attack is part of a concerted pattern of continuing armed activity', as Bethlehem suggests, 'mak[es] it easier to establish the threat [by] allow[ing] a more reliable prediction that a threat within this pattern will materialize with sufficient gravity', but does not add anything that is not already captured by likelihood and gravity.[581]

Upon critical reflection, Akande and Liefländer argue that 'only likelihood and gravity of the attack are genuine elements of the concept of imminence', but not the fourth temporal factor.[582] While 'imminence seems, on first sight, to require that the threat will materialize within a short time frame', the authors advocate that it 'entails no independent temporal requirement'.[583] The reason to repudiate the existence of an autonomous time-based limitation is that it 'would mean that where a highly probable and severe threat exists, whose realization is temporally remote, no action could be taken even where no future opportunity will arise to eliminate the threat'.[584] Instead, Akande and Liefländer advocate that 'where a threat is sufficiently probable and severe, the mere fact that it is still temporally remote should provide no independent injunction against action where that action is necessary and proportionate'.[585]

While rejecting an '*independent* temporal limitation' attached to the requirement of 'imminence', they underscore the significant bearing of temporal factors on the prospect of making 'accurate predictions' about the gravity and likelihood of the threat.[586] The

[579] Ibid.

[580] Ibid., p. 565. Bethlehem, 'Self-Defense Against an Imminent or Actual Armed Attack by Nonstate Actors', p. 775.

[581] Akande and Liefländer, 'Clarifying Necessity, Imminence and Proportionality', p. 565.

[582] Ibid.

[583] Ibid., p.564-565.

[584] Ibid. p. 565.

[585] Ibid.

[586] Ibid.

further removed a threat is in a temporal sense, the more difficult it would be to determine whether it is 'sufficiently probable and severe', while 'the shorter a causal chain is, the easier it becomes to predict what will occur'.[587] What is more, the temporal factors have an impact on necessity, which 'allows using force only where no peaceful alternative is available'.[588] In line with Chatham House's interpretation of imminence, the authors argue that the focus should be 'on the last point in time at which an effective responsive action is possible, rather than temporal proximity *per se*'.[589] In other words, '[w]hat is really at stake is whether some sort of self-defense action is demonstrably necessary – without any alternative, including later in time – rather than how temporally remote the threat is'.[590] Thus, the authors make the case that the temporal dimension, frequently considered under imminence, 'properly belongs to necessity',[591] which in the context of the WSOD is represented by the threshold requirement that 'peaceful means be inadequate'. This leaves the gravity and the likelihood of the threat as the only 'genuine elements of the concept of imminence [whose] actual content is mostly left vague'.[592] With a view to devising a framework for 'manifest failing' inquiries, this thesis adopts Akande and Liefländer's understanding of imminence and its key constitutive elements. The content of this determinant and the relevant indicators that attach to it will be explored further in the context of the Libyan intervention in the next chapter of this thesis.

The factors of the 'unwilling or unable' test

Despite the above attempts to clarify the 'unable or unwilling' test, Deeks 'provid[ed] the first sustained descriptive and normative analysis of the test'.[593] Regardless of coming under severe criticism for her treatment of the 'unable or unwilling' test as part of customary international law, her pioneering attempt to construct a normative framework to guide 'unable or unwilling' determinations has become a starting point for subsequent

[587] Ibid.

[588] Ibid. A similar idea is embodied in one of Bethlehem's imminence conditions, namely the need to consider 'the likelihood that there will be other opportunities to undertake effective action in self-defense that may be expected to cause less serious collateral injury, loss, or damage'. Bethlehem, 'Self-Defense Against an Imminent or Actual Armed Attack by Nonstate Actors', p. 775.

[589] Akande and Liefländer, 'Clarifying Necessity, Imminence and Proportionality', p. 565.

[590] Ibid.

[591] Ibid. p. 566.

[592] Ibid.

[593] Heller, 'Ashley Deeks' Problematic Defense of the "Unwilling or Unable" Test'.

endeavours to pinpoint the substantive factors that should inform 'unable or unwilling' inquiries.

The rationale behind her endeavour to address the 'substantive indeterminacy' of the 'unwilling or unable' test is analogous to the reasoning for clarifying the substance of the 'manifest failing' threshold outlined in the Introduction. In particular, Deeks envisions that a more precise 'unable or unwilling' test will provide three principal advantages, namely '(1) serving as a substantive constraint on action by the victim state, (2) providing a basis on which the victim state can (and must) justify its actions, and (3) as a procedural matter, structuring decision-making by the victim and territorial states and by international bodies in a way that improves the quality of those decisions'.[594] To advance the goal of creating a more detailed 'unable or unwilling' test, Deeks puts forward six 'factors', describing procedural steps and substantive assessments that victim states must undertake, in order to determine whether a territorial state is 'unable or unwilling' to curtail the actions of a menacing non-state actor operating from its territory: (1) 'attempt to act with the consent of or in cooperation with the territorial state'; (2) assesses the 'nature of the threat posed by the non-state actor'; ([3]) 'ask the territorial state to address the threat itself and provide adequate time for the latter to respond; ([4]) assess the territorial state's control and capacity in the relevant region as accurately as possible; ([5]) reasonably assess the means by which the territorial state proposes to suppress the threat, and; ([6]) evaluate its prior (positive and negative) interactions with the territorial state on related issues'.[595]

Prioritising host state consent and cooperation

The requirement to obtain host state consent is an invariable part of discussions of the 'unable or unwilling' test. This is not only evidenced by its presence in the principles of extraterritorial self-defence, detailed above, but also by the debate surrounding the justification for launching air strikes against Syria on account of it being 'unable or unwilling' to supress the threat posed by ISIL. For instance, Corten, Pak and Son's applications of Deeks' 'unwilling or unable' framework in the above context all affirm that the use of force on Syrian territory to prevent armed attacks by ISIL cannot be justified on the basis of the 'unable or unwilling' test because they do not meet this critical

[594] Deeks, '"Unwilling or Unable"', p. 506.
[595] Ibid. pp. 490, 519-532.

requirement.[596] According to Corten, this factor was 'simply ignored by the US in the Syrian case', as 'due to the strong [political] opposition between [the two governments] consent was never requested, if it was ever envisaged in the first place'.[597] In this context it is hardly surprising that Deeks attributes considerable weight to the first factor of her framework, 'the prioritization of consent and co-operation', which entails that:

> 'where a victim state obtains a territorial state's consent to use force within the latter's borders, the victim state need not conduct an "unwilling or unable" inquiry… but if the territorial state denies the victim state's request for consent, the denial may prove relevant in the subsequent "unwilling or unable" analysis'.[598]

That said, Deeks further postulates that '[e]ven if the territorial state declines to give the victim state consent to use force unilaterally, the victim state should, as a rule, explore whether there is an opportunity to work cooperatively with the territorial state to suppress the threat'.[599] This requirement 'seeks to reduce the overall number of cases in which the victim state uses force unilaterally in the territorial state'.[600] In this sense, it acts as an active constraint on unwarranted uses of force by the victim state.

While Deeks does not specify how the host state's refusal to cooperate should be taken into account in 'unable or unwilling' determinations, others have expanded on this notion. For instance, Couzigou asserts that if the host state proves unable to independently neutralise the non-state actor, then 'it must' either ask for external assistance or be willing to accept such assistance from the victim and/or other states.[601] The failure of an 'unable' state to request or accept assistance would automatically render that state 'unwilling' to suppress the threat posed by the non-state actor.[602] Consequently, whether the state is 'an

---

[596] Corten, 'The "Unwilling or Unable" Test', p. 779. Hui-Chol Pak & Hye-Ryon Son, *Military Intervention in Syria and the "Unwilling or Unable" Test: Lawful or Unlawful?*, 7 (4) *Russian Law Journal* (2019): 73-98, esp. pp. 76-77.

[597] Corten, 'The "Unwilling or Unable Test" Test', p. 779.

[598] Deeks, '"Unwilling or Unable"', p. 519.

[599] Ibid. p. 520.

[600] Ibid. p. 533.

[601] Irène Couzigou, 'The Right to Self-Defence Against Non-State Actors – Criteria of the "Unwilling or Unable" Test', in Anne Peters and Christian Marxsen (eds.) *Self-Defence Against Non-State Actors: Impulses from the Max Planck Trialogues on the Law of Peace and War, Max Planck Institute for Comparative Public Law & International Law* (MPIL), Research Paper No. 2017-07 (MPIL, 2017): 47-49, p. 48. [Accessed 16 July 2020] Available at SSRN: https://ssrn.com/abstract=2941640 or http://dx.doi.org/10.2139/ssrn.2941640

[602] Ibid.

"unable" failed State', or proves to be 'unwilling' on account of failing to consent to external interference, Couzigou argues that 'the victim State should have a right to self-defence against the non-State actor', subject to the observing the requirements of necessity and proportionality.[603] This view seems deeply problematic because it overprotects the victim state by bestowing upon it significant leverage to justify a military intervention in another state's territory. In addition, as Sjöstedt suggests, the 'unwilling or unable' test 'does not appear to dictate that a state must accept assistance from other states', because '[s]uch a conclusion would challenge the territorial sovereignty of a state'.[604] On this understanding, '[t]he fact that Syria has not consented to the military operations conducted by the US, UK and France in its territory would probably not make Syria an unwilling state'.[605]

In sum, as far as proving a state's 'unwillingness' is concerned, refusing assistance or cooperation, or failing to request assistance, cannot automatically render the host state 'unwilling' and justify action against the host state. However, when considered in conjunction with other factors, it can contribute to 'unwilling or unable' inquiries, especially when a pattern of unwillingness to cooperate can be established in the context of a serious threat posed to the victim state's security. Consequently, this understanding of the relevance of host state cooperation will be carried over to the framework for making 'manifest failing' inquiries introduced in Chapter V. Nonetheless, it is important to note that a few alternations will be required.

First, since the framework proposed in this thesis does not dictate what decision-makers should do, instead focusing on the substantive determinants of the 'manifest failing' test, the consideration of host state consent and cooperation will not be a procedural requirement. What is more, given that decisions to take action against an 'unwilling or unable' host state are not made by a 'victim state', but by the UN Security Council, a procedural requirement to safeguard host states against unwarranted intervention is already in place. Second, since considering the 'manifest failing' of the host state is only applicable to cases in which the host state has not consented to external interference, the analogous determinant of 'host state consent and cooperation' specified in the framework for 'manifest failing' inquiries in Chapter V will concentrate on ascertaining the host

[603] Ibid. pp. 48-49.

[604] Sjöstedt, 'Applying the Unable/Unwilling State Doctrine', p. 37.

[605] Ibid.

state's 'willingness to cooperate' with external actors (such as regional and international organisations) to supress the threat of atrocities. Relevant indicators will be derived from those listed under Risk Factor 6 of the UN 'Framework of Assessment for Atrocity Crimes' (see Table 3.1 below) and the investigation of the Libyan case (see Table 4.2. in Chapter IV).

The nature of the threat posed by the non-state actor

The second factor offered by Deeks, 'the nature of the threat posed by the non-state actor', is understandably pivotal to 'unable or unwilling' inquiries, because the affected state's 'understanding of the nature and seriousness of the threat from the nonstate actor that attacked it will permeate its consideration of the territorial state's willingness and ability to suppress that threat (as well as the territorial state's view of its own ability to suppress the threat)'.[606] In this regard, Deeks offers a range of associated factors, i.e. specific aspects of the situation, that the affected state ought to consider, namely:

> 'the geographic scope and intensity of the nonstate actor's activities, the sophistication of the attacks the nonstate actors have undertaken and are expected to undertake in the future, the number of actors in a particular area, the seniority (or juniority) of those actors within the organization, and the imminence of the threat of further armed attacks'. [607]

There is considerable overlap between the above factors for assessing the gravity of the threat posed by non-state actors for the purposes of making 'unable or unwilling' determinations and the determinations of relevant human rights violations and their gravity, derived in Chapter I. For instance, Deeks also relies on a combination of qualitative and quantitative aspects of the situation. Most importantly, the fact that she sees the nature and gravity of the threat as a key determinant of the 'unwilling or unable' test, similarly to Chatham House and Bethlehem, reinforces the need to include analogous determinations of the nature and gravity of the threat in the normative framework for 'manifest failing' inquiries being developed in this thesis. The concrete indicators that will inform this determination have already been specified in Table 1.1 at the end of Chapter I.

[606] Deeks, '"Unwilling or Unable"', p. 521.
[607] Ibid.

In addition, Deeks' assessment of the nature of the threat involves not only evidence of the extant observable threat, but also an element of prospective assessment of the threat, based on present facts and circumstances, i.e. the imminence of the threat of further attacks by the non-state actor (similarly to Chatham House, Bethlehem, Akande and Liefländer). This is illustrated in Deeks' depiction of what would make the victim state doubt whether the host state is able or willing to address the threat, namely:

> '[i]f the nonstate actor undertook an armed attack against the victim state that killed hundreds of people, runs multiple training camps in the territorial state, and, according to the victim state's intelligence, is planning several additional attacks in the next week'.[608]

In addition to an appreciation of indicators such as 'death toll' and preparations for further attacks indicative of a prospective threat, Deeks underscores that '[t]he higher the density of actors, the more senior the actors present on the territory, and the more sophisticated the group's organization, the harder it will be for the territorial state — indeed, for any state — truly to counter the threat to the extent and with the speed required'.[609] With these concrete examples, she offers an idea of the combination of qualitative and quantitative aspects of a given situation that would cast serious doubts on whether the territorial state is 'unwilling or unable' to address the threat and would warrant a thorough investigation into the host state's capacity to address the threat.

Hence, the work of 'unable or unwilling' scholars on specifying the elements of the threat that should inform determinations as to whether this test has been passed reveal that the 'imminence' of the threat is a key element of determining whether the host state is 'manifestly failing' in its responsibility to protect that ought to be included in the normative framework assembled in Chapter V. While the importance of considering forward-looking assessments of the threat of atrocities has been recognised by those studying mass atrocities, such as prospective death toll (Gallagher and Pape) and prospective gravity (Rosenberg and Strauss), this has not translated into formulating a concrete determination and its manifestations, relevant to assessing the threat of atrocities. In this sense, the detailed work done by 'unable or unwilling' scholars on specifying the notion of imminence, is vital for concretising this key element of 'manifest

[608] Ibid.
[609] Ibid.

failing' assessments. An important part of this is drawing on the criteria relevant to determine the imminence of the extraterritorial threat posed by non-state actors, in order to derive analogous indicators that could aid the determination of 'imminence' in the context of the threat of mass atrocities. These indicators are detailed in Table 3.1 at the end of this chapter.

Ask the territorial state to address the threat itself and provide adequate time for the latter to respond

Deeks' third factor lays out the first procedural steps that a victim state should take in the event that it has not secured host state consent to use force against the non-state actor operating from its territory. According to her, 'the most obvious way to determine whether a territorial state is willing or able to address the threat is for the victim state to request that it do so and evaluate its response'.[610] Asking the host state to take measures to supress the activities of violent non-state groups is an important procedural requirement, because it ensures that host state inaction does not stem from its ignorance of the situation and that it is informed about the threat posed by the non-state actor, including its nature and location.[611] As Deeks points out, an 'unable or unwilling' test without this requirement would 'crea[te] a regime of strict liability for' the host state, granting 'the victim state too much leeway to violate' its sovereignty, which will 'systematically over-protect the national security of the victim state' and under-protect the sovereignty of the host state.[612] The requirement to give notice is also 'an important first step in' determining whether the host state is 'unwilling or unable', as the latter's response to this request will afford the victim state vital information as to how it ought to proceed.[613] Should the host state prove both willing and able to address the threat, the victim state need not take further action, thus sparing resources and a violation of the host state's sovereignty.

A requirement that involves making a request presupposes allowing time for the host state to respond or act accordingly. As Deeks phrases it, '[u]nless the territorial state unequivocally rejects a victim state's request to suppress the threat, the victim state should allow the territorial state a reasonable amount of time in which to respond to that

[610] Ibid. p. 522.
[611] Ibid.
[612] Ibid. pp. 522-523.
[613] Ibid. p. 523.

threat'.[614] According to the author, 'What constitutes a "reasonable" amount of time must be judged in relation to the imminence of the threat'.[615] Deeks envisions that in exceptional scenarios where the affected state 'truly faces "no moment for deliberation", as where the nonstate actor already has initiated another armed attack, the victim state may need to respond immediately, possibly without soliciting the territorial state's assistance'.[616] In the instance described by Deeks, there can be no doubt as to the imminence of the threat, because it has already materialised. Provided that the attack launched by the offending non-state actor is of a gravity that poses a significant threat to the national security of the victim state, it may be permissible for the victim state to act in self-defence without seeking the aid of the host state. Excepting this extreme hypothetical scenario, Deeks does not provide further guidance as to what may constitute 'reasonable amount of time' for the host state to respond when a request to supress the activities of the violent non-state actor operating from its territory has been made.

Michael Schmitt has made a more focused endeavour to answer this question. According to him, 'any formula for resolving a conflict between one State's right to self-defense and another's right of territorial integrity must [recognise] that the need for conducting the defensive operations arises only when the latter fails to meet its policing duties'.[617] Equally, it must also take into account the issue of 'territorial integrity', which entails that prior to taking defensive action in another state's territory, the state under attack 'must first demand that the State from which the attacks have been mounted act to put an end to any future misuse of its territory'.[618] Should 'the sanctuary State either prov[e] unable to act or choos[e] not to do so, the State under attack may' enter the former's territory without its consent 'following a reasonable period for compliance (measured by the threat posed to the defender)'.[619] While what Schmidt suggests is in essence almost identical to the third requirement proposed by Deeks, he makes two important qualifications. First, the victim state is allowed to cross into the host state's territory 'for the sole purpose of conducting defensive operations', which means that it 'may not conduct operations directly against sanctuary State forces and must withdraw as soon as its defensive

[614] Ibid.

[615] Ibid.

[616] Ibid.

[617] Michael N. Schmitt, '"Change Direction" 2006: Israeli Operations in Lebanon and the International Law of Self-Defense', 29 (2) *Michigan Journal of International Law* (2009): 127-164, p. 161.

[618] Ibid.

[619] Ibid.

requirements have been met'.[620] The requirement to only take measures that are necessary to thwart the threat posed by the offending non-state actor is important in ensuring the protection of the interests of the (much weaker) host state against the victim state's pursuit of self-interested gain under the pretext of self-defence. Second, unlike Deeks, Schmidt provides effective guidance as to how to judge what 'reasonable response time is' by clarifying that the amount of time the host state should be afforded to fulfil the victim state's request will depend on and be measured by 'the threat posed to the defender'. This means that while the period for compliance is not something that can be determined ahead of time, it depends on the nature of the threat, its gravity and imminence. Since these contingencies are bound to vary greatly between cases, anything more specific than general guidance as to the factors relevant to determining reasonable time, would be counter-productive for making 'unable or unwilling' determinations on a case-by-case basis. The same applies to determinations of 'manifest failing', where the period for compliance is context-dependent, premised on the nature, gravity and the imminence of the threat, and will be determined by the international actors making demands that the host state takes concrete actions to address the threat.

The key takeaway from this factor proposed by Deeks is that it is important to consider the host state's response to any existing requests made by the international community, which will be reaffirmed in the context of the discussion of the Libyan intervention in the next chapter. In this sense, Deeks' procedural factor suggesting that the victim state 'asks the territorial state to address the threat' posed by the non-state actor translates better into to a substantive assessment involving the determination of 'the host state's response to extant requests to address the threat'. The latter assessment will tie better into the framework being developed in this thesis, which does not prescribe concrete actions that states or decision-makers should take in a given case, but provides a set of substantive determinants that could aid assessments as to whether a national government is 'manifestly failing' to protect its populations from atrocity crimes.

Assessment of the host state's control and capacity

In the event that the host state responds favourably to the request made by the victim state, the question that inevitably arises is whether the state is able to fulfil its pledge to supress the threat. As noted earlier, in the first instance, the nature and gravity of the threat is one

[620] Ibid. pp. 161-162.

of the factors that can raise doubts as to whether the state is able to keep violent non-state actors operating from its territory at bay. Historically, states have underscored 'the lack of territorial control in deciding whether a territorial state is unable to respond to a threat'.[621] Given that '[u]ngoverned and under-governed spaces are a frequent problem in practice', assessing the level of control the host state has over the area in which the non-state group is operating, is indispensable to determining whether the host state is able to address the threat.[622] Deeks suggests that these assessments should be done based on 'well-established public information, that a territorial state is unable to act in a particular area'.[623]

The state's 'inability' to supress the threat could also be due to the capacity of its armed forces or police. For instance, Deeks points out that often 'a state cannot control all parts of its territory because it lacks a robust set of forces to keep order'.[624] There could be a case where the host state's military force or law-enforcement officials are capable but are sympathetic to the nonstate actors and thus are 'unwilling' to act.[625] Alternatively, the host state can be rendered 'unable' if it lacks effective control over its otherwise capable security forces. The assessment of 'host state control and capacity' will not always be straightforward, such as in cases where the host state's military/police are judged to be 'not fully adequate to achieve the task, but are improving' or when the host state questions an assessment that it lacks territorial control.[626] In such difficult cases, the victim state would have to rely on other factors to reach a final determination. Either way, it is clear that the lack of control over the territory from which the threat emanates or concerns surrounding the capacity or control over the sanctuary state's military or law-enforcement forces are among the key factors that a state under threat ought to consider in order to make an accurate assessment as to whether the host state is 'unable or unwilling' to counter the threat posed by the non-state actor launching attacks from its territory.

While the development of the R2P 'was motivated primarily by the scenario of state-led atrocity crimes', the 2009 Secretary-General report on the implementation of the R2P 'moved beyond this initial starting point, by recognizing that a predatory state is only one

[621] Deeks, '"Unwilling or Unable"', pp. 525-526.
[622] Ibid. p. 525.
[623] Ibid. p. 527.
[624] Ibid.
[625] Ibid. pp. 527-528.
[626] Ibid. p. 528.

source of threat to populations and by considering non-state armed groups in the framework for operationalizing R2P'.[627] The need to do so is well-illustrated in the current situation in the Central African Republic. A serious human rights and humanitarian crisis gripped CAR in September 2013, when following 'months of brutality by the predominantly Muslim Seleka ("alliance") forces, which had overthrown the government of President François Bozizé in March, Christian militias known as the anti-balaka ("anti-machete") began to organize counterattacks'.[628]

According to HCHR reports, in the period of armed conflict in the country between 2013 and 2015 'antibalaka and ex-Séléka forces may have committed war crimes and crimes against humanity'.[629] Since then, predatory armed groups, including factions of the ex-Séléka rebel alliance and anti-balaka militias, have continued to incite ongoing violence, 'target[ing] civilians, humanitarian workers and UN peacekeepers, committ[ing] sexual violence, recruit[ing] children, and perpetrat[ing] attacks on IDP camps, schools, medical facilities and places of worship'.[630] Even though a peace deal, under the auspices of the AU, was signed on 6 February 2019 by the government and 14 armed groups, 'some signatories continue to violate the agreement and/or have exploited the deal to consolidate their control over territory'. The latest issue of GCR2P's bimonthly bulleting 'R2P Monitor' determines that '[o]ngoing violence by armed groups leaves populations in the Central African Republic at imminent risk of atrocity crimes', where 'imminent risk' describes 'a situation is reaching a critical threshold and the risk of mass atrocity crimes occurring in the immediate future is very high if effective preventive action is not taken'.[631] More specifically, the GCR2P analysis of the situation suggests that:

> '[o]ngoing violence and growing allegations of serious human rights violations by armed groups, including parties to the peace agreement,

[627] Jennifer Welsh, 'R2P's Next Ten Years: Deepening and Extending the Consensus', in Alex J. Bellamy and Tim Dunne (eds), *The Oxford Handbook of the Responsibility to Protect* (Oxford: Oxford University Press, 2016): 984-1000, p. 991.

[628] Human Rights Watch news release, '"They Came To Kill" Escalating Atrocities in the Central African Republic', 18 December 2013. [Accessed 31 January 2021]. Available at: https://www.hrw.org/report/2013/12/18/they-came-kill/escalating-atrocities-central-african-republic#

[629] GCR2P, 'R2P Monitor', Issue 55. Human Rights Watch documented war crimes and crimes against humanity committed by both forces in the period between late 2014 and 2017.: Human Rights Watch, 'Killing Without Consequence: War Crimes, Crimes Against Humanity and the Special Criminal Court in the Central African Republic', 5 July 2017. [Accessed 31 January 2021]. Available at: https://www.hrw.org/report/2017/07/05/killing-without-consequence/war-crimes-crimes-against-humanity-and-special

[630] GCR2P, 'R2P Monitor', Issue 55, p. 14.

[631] Ibid.

> highlight the risks resulting from limited governmental capacity outside the capital. Armed groups continue to control the majority of territory in CAR and profit from illegal taxation and arms trafficking'.[632]

On 22 January, CAR 'declared a 15-day state of emergency throughout the country as a coalition of armed groups seeks to overthrow the newly re-elected President Faustin-Archange Touadera'.[633] In light of '[m]ilitais claiming to represent ethnic or other groups control two-thirds of the country's territory… questions about government's control' over the vast territory of CAR remain pertinent.[634] In the context of these events, it is hardly surprising that the Central African Republic has been ranked among the bottom 9 of the 'most fragile states' in the world, in Fragile State Index annual reports (9th in 2013, 3rd in the period 2014-2017, 5th in 2018 and 6th in 2019 and 2020). In light of recent developments, it appears that the problems of lack of territorial control and the state's inability to stop gross human rights violations by predatory non-state groups are here to stay.

In light of this context, Welsh's claims that one of the key challenges for the R2P in the next ten years is the one posed by non-state perpetrators of atrocity crimes is hard to contest.[635] Nonetheless, the capacity of the host state of dealing with the threat is not explored in the literature on the 'manifestly failing' threshold. This could be attributed to the fact that there is an implicit expectation that a state which does not possess the capacity to protect its populations from the four crimes, would be receptive of support offered by the international community and the situation would ultimately fall under R2P's second pillar. This is the way in which the Secretary-General framed such situation in his 2009 report, which specifies that:

> 'Non-state actors, as well as States, can commit egregious crimes relating to the responsibility to protect. When they do, collective international military assistance may be the surest way to support the State in meeting its

[632] Ibid.

[633] Jean Fernand Koena, 'Central African Republic declares state of emergency', 21 January 2021 [Accessed 21 January 2021] Available at: https://www.washingtonpost.com/world/africa/central-african-republic-declares-state-of-emergency/2021/01/22/852fc84e-5c90-11eb-a849-6f9423a75ffd_story.html

[634] Ibid.

[635] Welsh, 'R2P's Next Ten Years', p. 991.

obligations relating to the responsibility to protect and, in extreme cases, to restore its effective sovereignty'.[636]

In light of the above discussion, it is clear that non-state actors operating on the territory of a host state could not only pose an extraterritorial threat to third states, but also an internal threat of atrocity crimes to innocent populations occupying the territory of the host state. Both scenarios where non-state actors are involved require answer the exact same question: What if [the host] state asserts that it is willing to take steps against the nonstate actor, but the victim state [/international community] has real doubts about the territorial state's level of control over the area in which the nonstate actor is operating, or serious concerns about the capacity of the territorial state's armed forces or police?'. Consequently, the present discussion of the 'host state control and capacity as a factor of the 'unwilling or unable' test can provide valuable insights for informing an analogous determinant of a 'manifest failing', which will inform the framework proposed in Chapter V. Specifically, based on Deeks understanding of the 'unwilling or unable' test the host state's lack of control over its security forces and the weakness of the host state's armed or law-enforcement force are negative indicators of the host state's ability to address an atrocity crisis within its borders. Five of the indicators of Risk Factor 3 of the UN 'Framework of Analysis for Atrocity Crimes' unpack concrete dimensions of these indicators and hence could be particularly instructive in 'manifest failing' inquiries: 'law enforcement… institutions that lack sufficient resources, adequate representation or training', '[l]ack of effective civilian control of security forces', '[l]ack of awareness of and training on international human rights and humanitarian law to military forces, irregular forces and non-State armed groups, or other relevant actors', '[l]ack of capacity to ensure that means and methods of warfare comply with international humanitarian law standards' and '[i]nsufficient resources to implement overall measures aimed at protecting populations'.[637] These indicators and Deeks' insights will inform the determination of 'the host state's territorial control and capacity to supress the threat of atrocity crimes' in the framework outlined in Chapter V.

[636] A/63/677 (2009), para. 40.

[637] UN, 'Framework of Analysis for Atrocity Crimes', p. 12.

Proposed means to suppress the threat

The fifth factor in Deeks' framework is relevant to situations in which the host state has suggested a plan of action. According to her, in such scenarios, it is 'imperative' that the victim state assesses the steps the host state has offered to take in order to counter the threat posed by the non-state groups.[638] This review of proposed means to thwart the non-state actor 'would advance the victim state's efforts to determine the territorial state's ability and willingness to address the threat by bearing down on the details of how the territorial state proposes to apply its capabilities to the situation at hand'.[639] It further 'forces' the victim state to answer the related questions of 'What steps are required to suppress the threat effectively?' and 'Is the territorial state actually able to bear the burden of using force effectively in a lower cost way than the victim state, or is the victim state persuaded that it must employ force itself within the territorial state's borders?'.[640]

In answering these questions, Deeks suggests that 'the approach must be what a "reasonable state" believes would accomplish the core goal of the victim state: avoiding further armed attacks by the nonstate actors operating from within the territorial state'.[641] Having adopted this lens for determining the kind of response the case requires, Deeks stipulates that in a situation where the victim state is in doubt about the adequacy of the host state's proposition, 'it should err on the side of acquiescing to the territorial state's plan'.[642] This means that the affected state would have flexibility to 'reconsider [the host] state's ability to supress the threat', should successive events prove that the measures the latter proposed were not sufficient.[643]

Following Deeks, Couzigou argues that the victim state should be in charge of determining whether the host state has resorted to 'all possible means' to meet the threat posed by the non-state perpetrator or it has proved 'unwilling or unable' to do so.[644] In

[638] Deeks, '"Unwilling or Unable"', p. 529.

[639] Ibid.

[640] Ibid.

[641] Ibid. Here, Deeks relies on Michael Schmitt's conception of a reasonable state, suggesting that 'Reasonable States do not act precipitously, nor do they remain idle as indications that an attack is forthcoming become deafening': Ibid. p. 530.

[642] Ibid. p. 530.

[643] Ibid.

[644] Couzigou, 'The Right to Self-Defence Against Non-State Actors', p. 47.

order to make this determination, Couzigou proposes that the victim state considers the following criteria:

> 'whether there has been a continuous pattern of armed attacks; whether the State criminalises the commission of armed attacks; whether the State conducts detailed investigations into those attacks; whether the State arrests, prosecutes, or extradites the authors of those attacks; whether the State complies with United Nations (UN) Security Council resolutions, if any, that sanction the authors of those attacks'.[645]

With the exception of the first criterion of the above list, which has been singled out as one of the factors pertaining to determining the imminence of the threat in the previous section, Couzigou's criteria illustrate concrete measures that the host state could take to supress the threat posed by the non-state actor, whose implementation or lack thereof should be considered in determinations of the host state's 'inability' or 'unwillingness'. Hence, she provides a useful list of mitigating actions potentially available to the host state, which supplement Deeks' proposal to consider the 'proposed means to supress the threat' as factor in 'unwilling or unable' inquiries.

While Deeks acknowledges that it would be difficult to articulate 'comprehensive objective test for how a "reasonable victim state" should behave', she envisions that as a minimum it 'would evaluate in good faith and with an objective eye the territorial state's proposal, keeping in mind both those actions that it has determined are necessary to suppress the threat and the practical limitations that any state likely would face in addressing that particular threat'.[646] For instance, 'a reasonable victim' state should keep in mind that 'even a state with a robust military capacity is unlikely to be able to suppress the threat fully'.[647] Hence, if the measures proposed by the host state do not 'anticipate complete success in rooting out every last member of the nonstate group [, this] does not necessarily indicate that the territorial state is unwilling or unable'.[648]

Deeks envisions that even though the victim state determines the fitness of the plan or actions of the host state, 'an expectation ex ante that the victim state must act reasonably

[645] Ibid., p. 48
[646] Deeks, '"Unwilling or Unable"', p. 530.
[647] Ibid.
[648] Ibid.

imposes some de facto constraints on that state'.[649] Should it become 'known publicly' that the victim state dismissed a host state proposal 'that seems reasonable to most states in the international community', the affected state's consequent use of force on host state territory could be seen as illegitimate.[650] Although her suggestions about assessing whether the approach proposed by the host state promises to be effective is generally sound, relying on the victim state's 'reasonableness' when making this assessment is not enough. An expectation that the affected state will act 'reasonably' does not pose effective constraints on its behaviour, if the international community does not have access to the information about the grounds for this assessment. Relying on a fortuitous turn of events, where that the host state chooses to volunteer this information or that it will somehow become public knowledge is also not a reliable way to protect the equities of the host state. Hence, in addition to the guidance provided by Deeks, a more robust 'unable or unwilling' test should envision some way of keeping the victim state accountable. A more consistent way to somewhat regulate the use of force by the affected state and prevent gratuitous interventions into the host state's territory is the requirement that the victim state gives reasons as to why it has deemed the host state's proposal to meet the threat deficient. Following a similar logic, Ahmed proposes concrete accountability mechanisms that should be put in place to avoid abuse of the 'unable or unwilling' test.

Ahmed underscores that the 'substantive indeterminacy' of the test leaves it open to abuse at the hands of powerful states, given the 'operational reality that victim states often tend to be relatively powerful states and host states, conversely, often tend to be relatively weak and therefore unable to deter victim state misbehaviour undertaken as "self-defence"'.[651] As previously noted, this element of scrutinising the actions or claims of the victim state is markedly missing from Deeks' framework, as it exclusively counts on the victim state to act 'reasonably', leaving important determinations such as whether the host state has cooperated, whether they have been given 'reasonable time' or whether they have the capacity to address the threat at the hands of a party with a vested interest in intervening in the host state's territory. As Ahmed remarks, the ambiguity of the 'unable or unwilling' test cannot be rectified by 'hoping that the victim state acts reasonably in self-defence – no state would claim that its own behaviour was

[649] Ibid. pp. 530-531

[650] Ibid. p. 531.

[651] Ahmed, 'Defending Weak States', p. 4.

unreasonable'.[652] Clearly, the high risk of abuse of power and undermining the sovereign rights of the much weaker host state necessitate a fundamental modification of the 'unable or unwilling' framework advanced by Deeks and other advocates of the doctrine, which allows the international community to hold victim states to account to deter them from making erroneous assertions and ultimately restrain unwarranted uses of force. To this end, Ahmed's argues that ('subject to existing restrictions on self-defence contained in the UN Charter and customary international law') it may be permissible for a victim state to use force under the premise that the host state is 'unwilling or unable' to supress the threat posed by a non-state actor, provided that it is 'willing to bear the burden of disclosing to the Security Council why it deems the host state to be ineffective - or "unwilling or unable"'.[653] Ideally, he envisions that 'the role of the Security Council would be passive' and involve 'collecting information that the victim and host states disclose voluntarily' and 'draw on substantive expertise [and information] from its Counter-Terrorism Committee (CTC) to verify whether or not a particular host state is unwilling or unable'.[654] This proposal demonstrates how the suggestion made in the previous section that victim states give reasons as to why they believe the host state to be 'unwilling or unable' can be operationalised.

The proposed means to assess the threat is very similar to the determination of 'the measures for mitigating the risk of mass atrocities', identified as key to determining whether a national government is 'manifestly failing' by Rosenberg and Strauss (see p. 57). The fact that it overlaps with the criteria of the 'unable or unwilling' test, identified by legal scholars, and includes a similar assessment as to whether the host state has proposed 'an action plan with timelines for mitigating the most urgent risk factors' gives further support to the understanding that it is central to the determination of a 'manifest failing' and supports its inclusion as a key determinant in the framework to be compiled in Chapter V. In so doing, this thesis will not rely on the vague notion of what a 'reasonable state' would do for reasons discussed above, but on the concrete measures to mitigate the risk of atrocities derived by Rosenberg and Strauss from research into past cases of atrocities, as well as the abovementioned propositions for mitigating actions that the host state could take to supress the threat or demonstrate its willingness to do so made by Couzigou. The latter will serve as a blueprint for the analogous indicators specified in

[652] Ibid. p. 15.
[653] Ibid. p. 21.
[654] Ibid. pp. 22, 26.

Table 3.1, below. As Chapter IV will illustrate, these analogous indicators also find support in the criteria states used to determine whether the 'manifest failing' threshold was met in Libya.

Prior interactions with the territorial state

A further step that the victim state could undertake in its inquiry into whether a host state is 'unwilling or unable' to supress the threat is 'to evaluate its prior interactions with the territorial state on issues related to the attacking nonstate actor'.[655] This could involve an appraisal of what the host state 'has done in response to any previous requests to take steps against the nonstate actors or other groups conducting armed attacks against the victim state'.[656] According to Deeks, proof of inaction on the part of the host state in the face of such attacks, provided that it was aware of them, would give the affected state 'a strong reason to infer that the territorial state will be unable or unwilling to act in response to the latest attack'.[657] However, 'even if a territorial state has not responded favorably to prior requests…the victim state nevertheless should make a new request'.[658] Failure to respond to this new request or declining to take measures to supress the threat, would constitute further proof of the host state's 'unwillingness and increase[e] the legitimacy of the victim state's decision' to take matters into its own hands and pursue the offending non-state actor.[659]

Relatedly, Deeks observes that victim states often consider prior armed attacks originating from the host state in question, when trying to determining whether they are 'unwilling or unable' to supress a current threat by a non-state actor residing in that state.[660] In this regard, she find that it is appropriate for the affected state to see previous 'attacks that it [or other states have] suffered from actors within the territorial state as substantive indications of' the latter being 'unwilling or unable' to act.[661] Reciprocally, 'if a state's territory never before has served as a launching pad for an attack by a nonstate

[655] Ibid.
[656] Ibid.
[657] Ibid.
[658] Ibid. p. 532
[659] Ibid.
[660] Ibid.
[661] Ibid.

actor', the affected state should be particularly careful when inferring the host state is 'unable or unwilling' to meet the present threat.[662]

Deeks' detailed discussion of the host state's 'prior interactions with the affected territorial state' showcase how the historical conduct of host state can contribute to determining whether a host state is 'unwilling or unable'. It also suggests that this is a useful determination that could help to judge whether a host state is 'manifestly failing' to protect its populations when considered in conjunction with other relevant aspects of the situation at hand. As seen in Table 3.1. at the end of this chapter, this retrospective determination and its concrete manifestations are analogous to the last factor of Deeks' framework. Namely, the determination of 'the host states' prior interactions with the international community on matters relevant to atrocity crimes' of table 3.1 mirrors Deeks' 'prior interactions with the victim state'.

Given her detailed consideration of the factors relevant to 'unwilling or unable' inquiries, it is not surprising that Deeks' normative framework has become 'the reference work with respect to the 'unwilling or unable' test'.[663] While by no means perfect, it created the momentum for other contributions towards clarifying the substance of this threshold for the use of force. Crucially, Deeks' recognises that '[a]s with most situations in international relations… there often will be no one fact that clearly establishes that a territorial state is unwilling or unable to suppress the threat that the victim state faces'.[664] In response to this operational reality, the six factors she proposes are intended 'to guide the victim state through this ambiguity and improve its decision making without dictating a one-size-fits-all answer' and at the same time 'contribute to a more coherent consideration by the affected states and the international community of particular uses of force, both ex post and ex ante'.[665] In this regard, what Deeks' has achieved embodies the objective of this thesis to develop a normative framework that aims to provide guidance for case-by-case inquiries into whether a host state is 'manifestly failing' in its responsibility to protect. Accordingly, her framework for facilitating assessments of suspected 'unable or unwilling' host states has served as a valuable source of inspiration for the normative framework for 'manifest failing' inquiries introduced in Chapter V.

---

[662] Ibid.

[663] Corten, 'The "Unwilling or Unable" Test', p. 779.

[664] Deeks, '"Unwilling or Unable"', p. 533.

[665] Ibid.

CONCLUSION

This chapter started with a discussion of the 'unable or unwilling' test found in ICISS' definition of the R2P concept. In so doing, it elaborated on and fortified the basis for relying on the analogy between the 'unable or unwilling' test in the context of self-defence against non-state actors and the 'manifest failing' threshold, established in Chapter II, in the present endeavour to devise a normative framework that can guide relevant stakeholders in making case-by-case determinations as to whether a host state is 'manifestly failing' in its responsibility to protect. The chapter proceeded to provide some background information on the 'unable or unwilling' test used in the context of contemporary debates on counter-terrorism, which included a brief account of the origins of this threshold and a more detailed insight into the debate arising from claims that the test is part of customary international law. Most importantly, this chapter detailed the ways in which legal scholars have endeavoured to address the substantive indeterminacy of the 'unwilling or unable' state in the context of the extraterritorial self-defence against non-state actors, with a view to identifying some stepping stones that could help address the same problem with regards to the 'manifestly failing' threshold.

Chatham House and Bethlehem's principles were important first steps in promoting and exploring the 'unable or unwilling' test. Nonetheless, it is Deeks' normative framework of 'factors that define what it means for a…state to be "unwilling or unable" to supress attacks by non-state actors' that has become 'the reference work' with regards to the 'unable or unwilling' threshold[666] and the starting point of analysis for subsequent endeavours to substantively unpack this threshold for the extraterritorial self-defence. As some of Deeks' critics point out, the main problem with her attempt to construct an 'unable or unwilling' framework is her failure to convincingly establish her central premise – the argument that the 'unable or unwilling' test has become a part of international law.[667] Nonetheless, the failure to substantiate the latter argument does not undermine the relevance of the substantive factors comprising her framework for determining whether a host state is 'unable or unwilling' to supress the threat posed by a non-state actor. In fact, even her critics have pointed out that 'Deeks should be

[666] Corten, 'The "Unwilling or Unable" Test', p. 779.
[667] Ibid.

commended for trying to think systematically about what the "unwilling or unable" test would require in practice'.[668]

In her article, Deeks set out to 'demonstrate why it is useful to consider the historical development and applications of the test in ascertaining what its meaning should be'.[669] Likewise, tracing the evolution of the 'unable or unwilling' test has helped to situate the discussions surrounding the manifest failing test within the broader tradition of the debate on the use of force. It has also proved invaluable in understanding existing 'unable or unwilling' frameworks, endeavouring to clarify a threshold requirement for the use of force on the territory of a third state that is failing to fulfil its responsibilities. Due to the fact that 'unable or unwilling' and 'manifest failing' inquiries are analogous, some of the factors of the 'unable or unwilling' test identified by legal scholars are clearly applicable to determining whether a host state is 'manifestly failing' in its responsibility to protect. This does not mean that the factors of the 'unable or unwilling' test would automatically apply to the destination context of R2P, but that they can provide a rough blueprint for the key factors that 'manifest failing' assessments should encompass. The analogous indicators, which emerged from the study of the 'unwilling or unable' test in this chapter, are specified in the table below.

Table 3.1. 'Manifest Failing' Determinants Derived from Analogous 'Unable or Unwilling' Factors

DETERMINANTS	**INDICATORS**
The nature and gravity of the threat	See the indicators of the determination or relevant human rights/acts of violence and determination of the gravity of human rights violation on pp. 79-80.
The imminence of the threat	*The assessment should involve the following considerations*: whether the anticipated atrocities are part of a concerted pattern of continuing grave human rights violations; the probability of grave human rights violations, the likely scale of the human rights violations and the loss likely to result therefrom in the absence of mitigating action.

[668] Heller, 'Ashley Deeks' Problematic Defense of the "Unwilling or Unable" Test'.
[669] Deeks, '"Unwilling or Unable"', p. 488.

Host state response to requests to address the threat	The host state's response to requests to address the threat can be evaluated against specific demands made by the international community, which could be expressed in the resolutions of regional (EU, ASEN, AU, LAS etc.) and international organisations (UNGA, UNSC, UNHRC).
The consent and cooperation of the host state	*The host state's willingness to cooperate with the international community to address the threat of atrocities can be determined on the basis of the following:* inadequate cooperation of the host state with regional and international human rights mechanisms; demonstrable unwillingness on the part of the host state to engage in dialogue, make concessions and receive support from the international community; the host state does not allow/impedes the provision of humanitarian assistance to affected populations; the host state does not allow/impedes the access of international human rights monitors and investigations into the human rights situation in the country; the host state impedes the safe passage of humanitarian and medical supplies, and humanitarian agencies and workers, into the country; the absence or restricted presence of regional or international organisations / actors on the territory of the host state with access to populations; failure to meet / comply with concrete demands to enable or cooperate with regional or international actors to address the threat of atrocity crimes.
Host state control and capacity	*The assessment as to whether the host state is 'unable' to address the threat should be made based on the following indicators:* the host state's lack of control over relevant parts of its territory; lack of effective civilian control of security forces; security or military forces who possess the requisite capacity but are unwilling to act against a non-state actor; failure to hold perpetrators to account; weak, untrained or underresourced law enforcement; lack of capacity to ensure that means and methods of warfare comply with international humanitarian law standards; insufficient resources to implement overall measures aimed at protecting populations; lack of awareness of and training on international human rights and humanitarian law to military forces, irregular forces and non-State armed groups, or other relevant actors; law enforcement institutions that lack sufficient resources, adequate representation or training; lack of awareness of and training on international human rights and humanitarian law to military forces, irregular forces and non-State armed groups, or other relevant actors

Measures proposed/taken by the host state to supress the threat	*The implementation of the following measures should be considered*: whether the host criminalises the commission of atrocities; whether the State conducts detailed investigations into grave human rights violations/alleged atrocities; whether the host state arrests, prosecutes, or extradites the authors of these crimes.
The host state's past interactions with the international community on matters relevant to mass atrocities	The following could help determine whether the host state is 'manifestly failing' to protect its populations when considered in conjunction with other determinations: - historical lack of compliance with or failure to meet the demands of the UN and regional organisations to protect populations from the threat of atrocities - past inaction, unwillingness or refusal of the host state to take all possible measures to put an end to imminent or ongoing atrocities

CHAPTER IV: 'MANIFEST FAILING' IN LIBYA

This chapter turns to state practice to explore the debate over the threshold for collective action in mass atrocity situations and identify the factors that states tend to proffer when making the case that this threshold has been crossed. Specifically, it examines the international community's response to the 2011 humanitarian crisis in Libya. As justified in the methodology chapter (Chapter II), this case presents a unique opportunity to explore the debate over threshold not only because of the unprecedented UN mandate for a military intervention in a functioning state without the express consent of the latter,[670] but also because the state in question was widely regarded as one that was 'manifestly failing' to protect its populations (as will be evidenced in this chapter).

The Libyan intervention was far from uncontroversial with some Security Council members – Germany and the BRICS – expressing their concern over the use of force and the violation of Libya's sovereignty in Council discussions at the vote on Resolution 1973.[671] Dissatisfaction grew quickly as the Libyan campaign progressed and the conduct of the intervention came to be widely regarded as having exceeded the civilian protection mandate authorised by the UN Security Council.[672] This led some to question the role that the R2P and the moral imperative to act on humanitarian grounds played in the decision to intervene in Libya.[673] As Hehir notes '[i]n response to the operational evolution of the "no-fly zone" in Libya into "regime change" certain states – Russia, China and South Africa in particular – determined to be more forthright in their opposition to external

[670] Morris, 'Libya and Syria', p. 1271.

[671] See S/RES/1973 (2011).

[672] Patrick Goodenough, 'Russia, China Accuse West of Exceeding UN Resolution, Making Libyan Crisis Worse', 29 March 2011 [Accessed 25 March 2020] Available at: https://cnsnews.com/news/article/russia-china-accuse-west-exceeding-un-resolution-making-libyan-crisis-worse; France 24 Press Release, 'Russia says NATO strikes on Libya exceed mandate', 15 April 2011 [Accessed 25 March 2020] Available at: https://www.france24.com/en/20110415-russia-says-nato-libya-strikes-exceed-un-mandate; Geir Ulfstein and Hege Føsund Christiansen, 'The legality of the NATO bombing in Libya'. 62 (1) *International & Comparative Law Quarterly* (2013): 159-171, esp. pp. 166-7. Ramesh Thakur, 'R2P after Libya and Syria: Engaging emerging powers', 36 (2) *The Washington Quarterly* (2013): 61-76, p. 61; Tom Keating, 'The UN Security Council on Libya: Legitimation or Dissimulation', in Aidan Hehir and Robert Murray (eds.) *Libya: The Responsibility to Protect and the Future of Humanitarian Intervention* (Basignstoke: Palgrave, 2013): 162-190, p. 175.

[673] Morris, 'Libya and Syria'; Hehir, 'The Permanence of Inconsistency', p. 146; Hehir, 'The Dog That Didn't Bark?', esp. pp. 219, 223-224; Robert W. Murray, 'Humanitarianism, Responsibility or Rationality? Evaluating Intervention as State Strategy', in Aidan Hehir and Robert Murray (eds.) *Libya: The Responsibility to Protect and the Future of Humanitarian Intervention* (Basignstoke: Palgrave, 2013): 15-33, p. 24.; Alan J. Kuperman, 'Obama's Libya debacle: how a well-meaning intervention ended in failure', 94 (2) *Foreign Affairs* (2015): 66-77.

involvement in intra-state affairs'.[674] This was seen in the scepticism of the BRICS countries towards international involvement in the Syrian crisis at the vote on the first draft resolution on Syria, because of the implementation of the civilian protection mandate in Libya.[675] Notably, the Russian representative underscored that 'it is very important to know how [Resolution 1973] was implemented and how a Security Council resolution turned into its opposite'.[676] Comparative examination of the underlying motivations for the Libyan intervention, and detailed analysis of the events following the decision to intervene and its alleged negative impact on the R2P doctrine fall outside the remit of this thesis. Instead, this thesis considers the factors that states on the Security Council profess to have shaped their position on the use of coercive measures to protect civilians in Libya and that informed their vote on Resolutions 1970 and 1973. In doing so, this discourse analysis will exhibit the ways in which states sought to legitimise their approach with reference to the notion that a threshold for collective action had been crossed owing to the Libyan state's 'manifest failing' to protect its populations. Whilst states professed additional justifications for their actions related to 'national security,' 'regional stability' and 'regime change', the investigation of the 'manifest failing' threshold demands that the focus remains narrowly on the dominant justification and grounds for the decision to respond to the events unfolding in Libya cited in UN Security Council Resolutions 1970 and 1973 – 'to protect civilians'. By maintaining this tight focus, the findings from this chapter will constitute the determinants and indicators that states see as key to determining whether a host state is 'manifestly failing' in its primary protection responsibilities. The rationale for deriving indicators from the views of state and the importance of triangulating this with 'manifest failing' criteria draw from secondary sources and via reasoning by analogy were articulated in detail in Chapter II.

The chapter is structured as follows. After a breakdown of the sources that formed the evidentiary basis for the analysis and a brief chronology of events during the Libyan crisis with particular regards to the threshold for international action, the analytical portion of

[674] Hehir, 'Assessing the influence', pp. 171-173.

[675] S/PV.6627 (2011).

[676] Ibid. p. 4. The Russian representative elaborated on this claim as follows: 'The demand for a quick ceasefire turned into a full-fledged civil war, the humanitarian, social, economic and military consequences of which transcend Libyan borders. The situation in connection with the no-fly zone has morphed into the bombing of oil refineries, television stations and other civilian sites. The arms embargo has morphed into a naval blockade in western Libya, including a blockade of humanitarian goods. Today the tragedy of Benghazi has spread to other western Libyan towns — Sirte and Bani Walid. These types of models should be excluded from global practices once and for all'.: Ibid.

this chapter proceeds in two parts. The first part is dedicated to mapping the debate over threshold in the broad discourse on the use of coercive measures in response to the Libyan crisis from its onset to the immediate aftermath of the intervention, in order to illustrate the pronounced presence of this debate in both Security Council meetings and P3 statements. In so doing, it will shed light on the terminology employed by states to refer to the notion of threshold, ultimately arguing that despite the lack of uniformity in language, the threshold for collective action to prevent mass atrocity crimes was not only active in decision-makers' minds, but also underscored by the wider UN membership.

The second part of the chapter follows up with a narrow investigation of the justifications offered by states for the adoption of collective action measures with respect to declared civilian protection goals in Libya. The aim of this examination is to identify the trends in the type of information that states professed when making the case that a substantive threshold had been crossed or not. On the basis of this investigation, it will be argued that in addition to invoking the 'manifest failing' and 'unwilling or unable' threshold requirements in discussions over the adoptions of coercive measures, states also expressed some concrete ideas as to what this threshold entails substantively. By analysing these discussions, this chapter will shed light on the determinants and indicators that states have associated with the threshold for collective action, which will feed into the 'manifest failing' framework to be proposed in Chapter V.

The sources analysed

The analysis in this chapter draws on the evidence of P3 government statements and the official record of UN Security Council meetings in which states discussed the threshold for collective action. A search of P3 government statements offering justifications for their position on Libya in the period between 15 February and 31 March 2011 generated a core sample of 38 textual records – 9 for the US, 9 for the UK, 17 for France and 3 joint statements. The goal was to include all relevant statements of P3 statesmen identified in the selected archives (for details, see Chapter II). This strategy ensured that sufficient data was gathered so that the resulting sample captured the most significant P3 statements in the period under study and was comprehensive enough to not only reach but also go beyond the point of saturation, when adding more statements did not lend new relevant insights.

In the same period, the Security Council held six meetings to discuss the situation in Libya, starting with a closed meeting on 22 February with the participation of 75 invited countries, of which there was no public record. Out of the five publicly available meetings on the situation in Libya, two were held in February under the agenda item 'Peace and Security in Africa' and three in March under the item 'The Situation in Libya'. Security Council member states did not make statements in three of these meetings, which took the shape of briefings on the situation in Libya. Therefore, the remaining two meeting records on the voting on Resolutions 1970 and 1973 formed the evidentiary basis for identifying the factors that Security Council member states professed to justify their decision to take collective action in Libya. All of the P3 statements and UN Security Council meetings in this initial period of the Libyan crisis were read with a view to both mapping the debate over threshold and identifying the determinants and indicators of the threshold for collective action that states proffered in the Libyan context.

In addition to the above documents, the analysis of the debate over threshold was also based on Security Council meetings in the period from 1 April 2011 to 22 December 2011, read for supplementary insights into the debate over threshold following the decision to intervene in Libya through to the immediate aftermath of the intervention. A search of the index to proceedings of the Security Council in this period for situation-specific, regional and thematic meetings to discuss the 'protection of civilians in armed conflict'[677] revealed a total of 25 Security Council meeting records, 12 of which were held to hear briefings. The analysis therefore focused on the remaining 13 meeting records, which saw contributions from Security Council members and invited UN member states (see Table 1, below). The examined meetings comprised 7 specific meetings records on the 'Situation in Libya', 2 regional meetings under agenda item 'Peace and Security in Africa' and 4 thematic meetings records on the 'Protection of Civilians in Armed Conflict'. The latter debates, in which participating states recollected the events of and the international response to the crisis in Libya with a view to civilian protection, afforded valuable insights into how, normatively speaking, the international community should determine whether a threshold has been crossed.

[677] Index to proceedings of the Security Council: 66th year (2011). [Accessed 05 November 2019] Available at: http://research.un.org/en/docs/sc/quick/meetings/2011

Table 4.1. List of UNSC meetings which saw contributions from member states between 22 February 2011 and 22 December 2011*[678]

Meeting Record	Date	Topic	Security Council Outcome/Vote
S/PV.6674 (Resumption 1)	5 December	Peace and Security in Africa	S/RES/2023 (2011) 13-0-2
S/PV.6650 (Resumption 1)	9 November	Protection of civilians in armed conflict	-
S/PV.6650	9 November	Protection of civilians in armed conflict	-
S/PV.6647	2 November	The Situation in Libya	-
S/PV.6644	31 October	The Situation in Libya	S/RES/2017 (2011) 15-0-0
S/PV.6622	26 September	The Situation in Libya	-
S/PV.6620	16 September	The Situation in Libya	-
S/PV.6595	28 July	The Situation in Libya	-
S/PV.6566	27 June	The Situation in Libya	-
S/PV.6561	21 June	Peace and Security in Africa	-
S/PV.6531 (Resumption 1)	10 May	Protection of civilians in armed conflict	-
S/PV.6531	10 May	Protection of civilians in armed conflict	-
S/PV.6528	4 May	The Situation in Libya	-
S/PV.6498	17 March	The Situation in Libya	S/RES/1973 (2011) 10-0-5
S/PV.6491	26 February	Peace and security in Africa — Libya	S/RES/1970 (2011) 15-0-0

The Libyan crisis and international response

Joining the wave of social movements associated with the 'Arab Spring' upheavals that spread across the Middle East and North Africa in the beginning of 2011, popular protests against the government of Colonel Muamar Qaddafi erupted in Libya in mid-February 2011. The anti- Qaddafi demonstrations quickly turned violent, leading to a country-wide

[678] *Table based on Index to proceedings of the Security Council: Ibid.

rebellion with clashes between security forces and rebels. As Simon Adams notes, although reports of civilian casualties in a week of fighting offered varying estimates, and even though some of them 'were exaggerated, by 22 February it was clear that the Qadhafi regime, in its desperation to hold on to power, was willing to use extreme violence to crush the popular uprising [and] that the threat of mass atrocities was imminent and real'.[679] Indeed, in his speech of the same day, Qaddafi not only referred to protesters as 'cockroaches' and 'rats' 'serving the devil', but made his intentions clear by publicly declaring his resolve to 'cleanse Libya house by house'.[680]

Initial response to the Libyan crisis

The rapidly unfolding government-led campaign of mass violence against Libyan populations, featuring incidents of hate speech and threats by the government, prompted a strong reaction by the international community with senior UN officials and regional organisations condemning these acts of aggression. Expressing his concern for the escalation of violence, UN Secretary-General Ban Ki-moon urged Qaddafi to put an end to the use of force in a telephone conversation of 21 February, as the former's spokesperson reported.[681] Qatar's prime minister and foreign minister went even further as he called for an extraordinary meeting of the Arab League and stated that 'condemnation is not enough, [the UN Security Council] should take the responsibility and do something to help civilians in Libya'.[682]

On 22 February, in the wake of Qaddafi's unequivocal threatening statements, there was a surge of international activism. The League of Arab States (LAS) was amongst the first to condemn the excessive violence against the civilian populations and take disciplinary action against Libya by suspending it from its meetings until it met the League's demands for an immediate end to all violence.[683] Later in the day, the Security Council held a

[679] Simon Adams, 'Libya and the Responsibility to Protect', Global Centre for the Responsibility to Protect Occasional Paper Series No. 3 (New York: Global Centre for the Responsibility to Protect, 2012), p. 6.

[680] BBC news, 'Libya protests: Defiant Gaddafi refuses to quit.' 22 February 2011 [Accessed 4 February 2020] Available at: https://www.bbc.co.uk/news/world-middle-east-12544624

[681] UN News press release, 'UN chief urges restraint by Arab leaders as protests continue', 21 February 2011 [Accessed 4 February 2020] Available at: https://news.un.org/en/story/2011/02/367112-un-chief-urges-restraint-arab-leaders-protests-continue

[682] Al Jazeera news release, 'Libya protests spread and intensify', 21 February 2011 [Accessed 06 February 2020] Available at: https://www.aljazeera.com/news/africa/2011/02/2011221133557377576.html

[683] Security Council Report, 'Update Report No. 3: Libya', 25 February 2011 [Accessed 4 February 2020] Available at: https://www.securitycouncilreport.org/update-report/lookup_c_glkwlemtisg_b_6586331.php

private meeting to discuss the issue of 'Peace and Security in Africa', where the Under-Secretary-General for Political Affairs, Lynn Pascoe, delivered a briefing on the situation in Libya.[684] In a statement to the press following the informal consultations, UN Security Council member states welcomed the Arab League's statement, urged the Libyan government to put an end to the violence and 'to meet its responsibility to protect its populations'.[685] Also on the same day, the UN Secretary-General's Special Advisers on the Prevention of Genocide and the Responsibility to Protect and the UN High Commissioner for Human Rights, Ms Navi Pillay, expressed their concern that the 'widespread and systematic attacks against the civilian population may amount to crimes against humanity'.[686] In this regard, they specifically underscored that the reported use of military forces, mercenaries, machine guns, snipers and aircraft against civilians were grave violations of international law.[687] Going one step further, the Special Advisers explicitly framed their response in R2P-terms by reminding Libya of its 2005 Word Summit commitment to protect populations from the four atrocity crimes and their incitement.[688] On 23 February, the Secretary-General reiterated the concern expressed by his Special Advisers that 'the reported nature and scale of the attacks on civilians are egregious violations of international humanitarian and human rights law' and underscored that 'the government of Libya must meet its responsibility to protect its people'.[689] According to Bellamy, early activism on the part of the UN Secretariat played a pivotal role in framing the Libyan crisis in terms of human protection and mass atrocity

[684] S/PV.6486 (2011).

[685] UNSC press statement, 'UN Security Council Press Statement on Libya', SC/10180-AFR/2120, 22 February 2011 [Accessed 4 February 2020] Available at: https://www.un.org/press/en/2011/sc10180.doc.htm

[686] OHCHR press release, 'Pillay calls for international inquiry into Libyan violence and justice for victims', 22 February 2011. [Accessed 4 February 2020] Available at: https://newsarchive.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=10743&LangID=E; United Nations press release, 'Statement by the UN Secretary-General Special Adviser on the Prevention of Genocide, Francis Deng, and the Special Adviser on the Responsibility to Protect, Edward Luck, on the situation in Libya', 22 February 2011.

[687] Ibid.

[688] Ibid.

[689] UN News press release, 'Ban strongly condemns Qadhafi's actions against protesters, calls for punishment', 23 February 2011 [Accessed 4 February 2020] Available at: https://news.un.org/en/story/2011/02/367342-ban-strongly-condemns-qadhafis-actions-against-protesters-calls-punishment; For a video of the full statement see: BBC News, 'UN Secretary-General Ban Ki-moon condemns Libya violence', 23 February 2011 [Accessed 4 February 2020] Available at: https://www.bbc.co.uk/news/av/world-middle-east-12561784/un-secretary-general-ban-ki-moon-condemns-libya-violence

prevention, thus setting the tone for the ensuing Security Council dialogue,[690] and in part provided the impetus for the adoption of Resolution 1970 three days later.[691]

Also on 23 February, the African Union's Peace and Security Council 'strongly condemn[ed] the indiscriminate and excessive use of force and lethal weapons against peaceful protestors, in violation of human rights and International Humanitarian Law', and urged the Libyan authorities to put an end to the violence and protect their populations.[692] Soon after, on 25 February, UN Human Rights Chief, Navi Pillay, reminded the Council of its pledge at the 2005 World Summit to uphold the R2P, and invoked the 'manifest failing' threshold when recalling the international community's subsidiary responsibility to intervene by taking action to protect populations at risk.[693] The same day, the UN Human Rights Council made an unprecedented request to suspend Libya's membership of the Council, paving the way for its effective suspension by the UN General Assembly on 1 March.[694] Concurrently, the Council decided to dispatch an international commission of inquiry to investigate the suspected human rights violations in Libya.[695]

On 25 February, the UN Security Council held a meeting to consider the issue of 'Peace and Security in Africa', at which the Secretary-General briefed member states on the situation in Libya and the permanent representative of Libya called upon the international community to act.[696] The Secretary-General highlighted the 'continuing reports of violence and the indiscriminate use of force' in Libya in the three days that passed since the Council was last briefed on the situation.[697] He voiced his 'concerns about the nature and scale of the conflict…includ[ing] allegations of indiscriminate killings, arbitrary arrests, the shooting of peaceful demonstrators and the detention and torture of the

[690] Bellamy, 'Libya and the Responsibility to Protect', p. 265; See also: Bellamy and Williams, 'The new politics of protection?', p. 839.

[691] Bellamy, 'Libya and the Responsibility to Protect', p. 265.

[692] African Union Peace and Security Council, Communique on Libya, 23 February 2011 [Accessed 5 February 2020] Available at: http://www.peaceau.org/uploads/psc-communique-on-the-situation-in-libya.pdf

[693] OHCHR, 'Situation of Human Rights in the Libyan Arab Jamahiriya: Statement by Navi Pillay', 25 February 2011.

[694] OHCHR press release, 'UN Human Rights Council recommends suspension of Libya', 25 February 2011. [Accessed 4 February 2020] Available at: https://www.ohchr.org/EN/NewsEvents/Pages/HRCSpecialSessionLibya.aspx

[695] Ibid.; United Nations General Assembly, A/RES/65/265, 1 March 2011.

[696] S/PV.6490 (2011), p. 2.

[697] Ibid.

opposition and the use of foreign mercenaries,'[698] as well as 'house-by house searches and arrests', 'go[ing] into hospitals to kill wounded opponents', 'the killing of soldiers who refused to fire upon their countrymen' and 'threaten[ing] citizens with a civil war and the possibility of mass killing if the protests continue'.[699]

In addition to enumerating of specific acts of violence evidencing human rights violations taking place in Libya, the Secretary-General underscored the gravity of these abuses. Specifically, he made references to the reported death toll of a thousand people being killed and high casualties in recent violent clashes.[700] He also provided ample evidence of irregularities in the movement of populations, noting that there were 'serious indications of a growing crisis of refugees and displaced persons' with reports of a cumulative toll of 37, 000 people fleeing Libya since 22 February and potentially 'much larger numbers of residents and migrant workers… trapped and unable to leave for safety'.[701] Secretary-General Ban also underscored reports of violence against vulnerable groups, 'of indiscriminate attacks on foreigners believed to be mercenaries, as well as … of women and children being among the victims', and 'of refugees being harassed and threatened with guns and knives'.[702] Related indicators of a hostile policy against victim populations included 'dangerous impediments to medical treatment and access for humanitarian workers'.[703]

Together, the aspects of the situation in Libya underscored by the Secretary-General made for a convincing case that the 'egregious violations of human rights taking place in Libya' were of a gravity commensurable with the urgent need to act swiftly and do 'everything possible to ensure' the protection of Libyan 'civilians at demonstrable risk'.[704] To highlight the urgency of the situation, this call for action came with the stipulation that if further evidence was needed, 'it should be sought simultaneously with measures to afford protection'.[705] Leaving little doubt as to the severity of the situation and the need to act, the Secretary-General quoted Ms Pillay's statement of the same day to stress that 'when a State manifestly fails to protect its population from serious international crimes, the

[698] Ibid. p.3.
[699] Ibid. p.2
[700] Ibid.
[701] Ibid. p. 3.
[702] Ibid.
[703] Ibid.
[704] Ibid.
[705] Ibid.

international community has the responsibility to step in and take protective action in a collective, timely and decisive manner'.[706] Thus, a day before the UN Security Council voted on the first draft resolution on Libya, the Secretary-General invoked the 'manifest failing' threshold requirement to remind member states of their subsidiary obligations under the R2P.

The Secretary-General's attempt to convince member states of the Security Council that the situation in Libya warranted urgent action and the international discourse that preceded it underscore aspects of the situation that scholars such as Gallagher, Rosenberg and Strauss have associated with the 'manifestly failing' threshold, namely evidence of grave human rights violations, death toll, weapons used, social dislocation, and the deliberate targeting of vulnerable groups (see Chapter I). Moreover, senior UN officials invoked the 'manifest failing' threshold when they wished to remind the international community of their protection responsibilities. The extent to which claims of Libya's 'manifest failing' and the relevant factors professed by regional organisations and UN officials resonated with Council members will be assessed later in this chapter through an examination of P3 and UN Security Council discourses on collective action in response to the Libyan crisis.

Resolution 1970

The day after Secretary-General Ban's briefing, on 26 February, the UN Security Council unanimously adopted Resolution 1970, 'condemn[ing] the violence and use of force against civilians' and 'the gross and systematic violation of human rights'.[707] Echoing the warnings of the UN Secretariat, the Council recognized that crimes against humanity might have been committed in light of the violent attacks against civilian populations in Libya and welcomed regional organisations' condemnation of such reported violations of international law, as well as the Human Rights Council's decision to establish a commission of inquiry to investigate the latter.[708] Notably, the resolution encompassed a comprehensive package of targeted sanctions under Chapter VII (art. 41) of the UN Charter, including an arms embargo, travel bans and asset freezes, that gave practical effect to the Council's formal demands from the Libyan government for 'an immediate end to the violence' and respect for human rights and international law, coupled with a

[706] Ibid.
[707] S/RES/1970 (2011).
[708] Ibid.

referral of the case to the International Criminal Court (ICC).[709] Although the Chapter VII deterrent measures established against the Libyan regime fall within the remit of R2P's reactive dimension, Resolution 1970 made no reference to the latter and only invoked 'the Libyan authorities' responsibility to protect its population'.[710] Whilst Resolution 1970 was 'relatively uncontroversial', as Bellamy notes, some Council members communicated their reluctance towards adopting more coercive legal measures during the informal consultations preceding the resolution.[711] For instance, Russia drew a clear line between 'restrictive' sanctions, such as the ones imposed by Resolution 1970, and 'sanctions, even indirect, for forceful interference in Libya's affairs, which could make the situation worse'.[712] According to Bellamy, the first resolution on Libya 'bundled together a variety of punitive measures when slower-moving events might have facilitated a more graduated approach to coercive inducement'.[713] The severity of the concrete measures aside, the very demand for an end to the violence and respect for human rights and international humanitarian law means that the Security Council had already started taking the necessary steps to determine whether a 'manifest failing' was underway in Libya. Similarly to the 'unable or unwilling' requirement of 'asking the host state to address the situation', which was discussed in Chapter III, factoring the host state's response to international demands that it takes action to mitigate the risk of atrocities is a key step towards ascertaining the host state's role in attenuating the risk of the four crimes and thus determining whether it is 'manifestly failing'. Part II of this chapter will evidence the significance that states attributed to requests that the host state address the threat of atrocities in determining whether the threshold for strong international action had been met.

The period between the two resolutions

In response to Resolution 1970, on 2 March, the Libyan government submitted a letter to the UN Security Council, protesting the premature denouncement of the regime and requesting the suspension of the Resolution until the truthfulness of the accusations could

[709] Ibid.

[710] Ibid.

[711] Bellamy and Williams, 'The new politics of protection', p. 840.

[712] S/PV.6491.

[713] Bellamy, 'Libya and the Responsibility to Protect', p. 266.

be confirmed.[714] Libya's request, however, was not matched by a demonstrable commitment to put an end to the violence and work towards a peaceful resolution of the crisis. On the contrary, the weeks following the adoption of Resolution 1970 were marked by reports of unrelenting violence, which prompted the international community to ramp up the pressure on the Qaddafi regime and deliberate further action.

The UK started considering the possibility of a no-fly zone at the end of February,[715] and so did France,[716] but it was not until the Gulf Cooperation Council (GCC) expressed its support for a no-fly zone in a statement of 7 March, when the debate started to gain traction and the impetus for the regional support that followed was generated.[717] The next day, the Organization of the Islamic Conference (OIC) also released a statement of 'support for the international calls to establish a no-fly zone over Libya under UN supervision' but stressed its opposition to 'any form of military intervention to [sic.] Libya'.[718] On 10 March the GCC denounced Qaddafi's regime for having 'lost its legitimacy' and reiterated its support for a no-fly zone, whilst drawing attention to factors evidencing the severity of atrocities committed against the Libyan population, namely the weapons used (i.e. heavy weapons and live ammunition), the recruitment of mercenaries, and the large civilian death toll.[719] The same day, the African Union's Peace and Security Council highlighted the same aspects of the situation, expressing its 'unequivocal condemnation of the indiscriminate use of force and lethal weapons, whoever it comes from, resulting in the loss of life, both civilian and military'.[720] Nonetheless, instead of

---

[714] Security Council Report, 'Update report No. 1: Libya', 14 March 2011 [Accessed 7 February 2020] Available at: https://www.securitycouncilreport.org/update-report/lookup_c_glKWLeMTIsG_b_6621881.php

[715] David Cameron, 'Prime Minister's statement on Libya', Statement on the situation in Libya to the House of Commons, 28 February 2011 [Accessed 30 January 2020] 28 February 2011. Available at: https://www.gov.uk/government/speeches/prime-ministers-statement-on-libya--2

[716] Alain Juppé, 'Libya', Reply given by Alain Juppé to a question in the National Assembly, 1 March 2011. [Accessed 17 January 2020] Available at: http://www.assemblee-nationale.fr/13/cri/2010-2011/20110128.asp

[717] The Guardian news blog, 'Libya uprising – Monday 7 March', 7 March 2011 [Accessed 31 March 2020] Available at: https://www.theguardian.com/world/blog/2011/mar/07/libya-uprising-live-updates

[718] OIC, 'Final Communique Issued by the Emergency Meeting of the Committee of Permanent Representatives to the Organization of the Islamic Conference on the Alarming Developments in Libyan Jamahiriya', 8 March 2011 [Accessed 31 March 2020] Available at: https://www.oic-oci.org/topic/?t_id=5022&ref=2110&lan=en

[719] Al Jazeera, 'GCC: Libya regime lost legitimacy', 11 March 2011 [Accessed 16 Aril 2020] Available at: https://www.aljazeera.com/news/middleeast/2011/03/201131021173060618l.html; Security Council Report, 'Update report No. 1: Libya'.

[720] African Union Peace and Security Council, Communique of the 265th Meeting of the Peace and Security Council, PSC/PR/COMM.2 (CCXLV), 10 March 2011 [Accessed 4 April 2020] Available at: https://www.gov.za/communiqué-265th-meeting-african-union-peace-and-security-council

calling for international action to address the situation, the AU reaffirmed its 'rejection of any foreign military intervention, whatever its form' and its 'strong commitment to the respect of the unity and territorial integrity of Libya'.[721] On 11 March, at an extraordinary summit of the European Council in Brussels, European Union (EU) heads of state reached an agreement to 'examine all necessary options [to protect civilian populations] provided that there is a demonstrable need, a clear legal basis and support from the region'.[722] However, the Council did not announce its support for the Franco-British calls for a no-fly zone due to the opposition they were reportedly met with by other NATO members, particularly Germany and the US.[723] In contrast with the cautious stances of the AU and EU, on 12 March the League of Arab Nations (LAS), known for its opposition to humanitarian intervention, issued a declaration demanding that the UN Security Council establish a no-fly zone over Libya, and safe areas to protect the civilian populations from shelling.[724] Glanville, Bellamy and Williams underscored LAS' unprecedented verdict as 'decisive' and 'crucial' for generating the political consensus amongst key players 'to push for a resolution authorizing the use of force and also in tempering the opposition among other states to the passage of such a resolution'.[725]

By 10 March Qaddafi's retaliation against rebel forces in the form of heavy weapons attacks by air, land and sea turned the tide in his favour, as he started to recover territory taken by the opposition.[726] Qaddafi's forces had cleared all obstacles on their way to the rebel stronghold of Benghazi by then.[727] As Simon Adams argues, the unwavering behaviour, 'nature and structure' of the Libyan regime 'closed off other diplomatic

---

[721] Ibid.

[722] European Council, 'Declaration adopted by the extraordinary European Council', EUCO 7/1/11 (Rev 1), 11 March 2011, Brussels, para. 6. [Accessed 16 April 2020] 20 April 2011. Available at https://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ec/119780.pdf

[723] Camille Grand, 'The French experience: Sarkozy's War?', in Karl P. Mueller et al. (eds.) *Precision and Purpose: Airpower in the Libyan Civil War* (Santa Monica: RAND Corporation, 2015): 183-202, p. 187.

[724] UNSC, 'Council of the League of Arab States, Resolution No. 7360, extraordinary session, 12 March 2011', S/2011/137, 15 March 2011, Annex, paras. 1-2. Glanville, 'Intervention in Libya', p. 334.

[725] Glanville, 'Intervention in Libya', pp. 334-335; Bellamy and Williams, 'The new politics of protection', p. 841.

[726] Reuters news article, 'Gaddafi launches counter-offensive on Libyan rebels', 6 March 2011 [Accessed 2 July 2020] Available at: https://www.reuters.com/article/us-libya-protests/gaddafi-launches-counter-offensive-on-libyan-rebels-idUSTRE71G0A620110306

Chris McGreal, 'Libyan rebels in retreat as Gaddafi attacks by air, land and sea', 10 March 2011 [Accessed 7 February 2020] Available at: https://www.theguardian.com/world/2011/mar/10/libya-gaddafi-takes-ras-lanuf

[727] France 24 news article, Gaddafi forces pound key rebel-held town, 15 March 2011 [Accessed 7 February 2020] Available at: https://www.france24.com/en/20110315-gaddafi-forces-pound-key-rebel-held-town-ajdabiya-libya-airstrikes

possibilities' for 'further mediation and accommodation'.[728] A testimony to the latter was the UN Secretary-General's failed attempt to convince Qaddafi to comply with the demands of Resolution 1970 in a telephone conversation with the Libyan leader.[729] It was in the context of the Libyan authorities' rejection of the demands of the international community, the reversal of the situation on the ground, and regional support for international action that the UN Security Council proceeded to debate and vote on the draft of Resolution 1973, submitted by France, Lebanon, the UK and the US.[730]

Resolution 1973

Resolution 1973 was adopted on 17 March, with ten states voting in favour of the resolution (Bosnia and Herzegovina, Colombia, France, Gabon, Lebanon, Nigeria, Portugal, South Africa, United Kingdom of Great Britain and Northern Ireland, United States of America) and five abstaining (Brazil, China, Germany, India, and the Russian Federation).[731] 'Deploring the failure of the Libyan authorities to comply with resolution 1970', the Council voiced its concern over the 'deteriorating situation, the escalation of violence, and the heavy civilian casualties'.[732] Once again, it reiterated the Libyan state's primary 'responsibility to protect' and reaffirmed that 'the widespread and systematic attacks [against civilians in Libya] may amount to crimes against humanity'.[733] The Council also took note of the abovementioned recent activity of regional organisations, including acts of condemnation and calls for a no-fly zone. Taking action under Chapter VII of the UN Charter, the UN Security Council demanded an immediate ceasefire, a 'complete end to all violence', and increased efforts to find a diplomatic solution to the Libyan crisis. [734] Having thus framed the situation as a 'threat to international peace and security', the Council sanctioned the use of 'all necessary measures [excluding a foreign occupation force] to protect civilians'.[735] In addition to strengthening the asset freeze and arms embargo sanctions established by the preceding resolution, the Council imposed a

[728] Adams, 'Libya and the Responsibility to Protect', p. 5.
[729] Bellamy and Williams, 'The new politics of protection', p. 840.
[730] UNSC, S/2011/142, 17 March 2011.
[731] S/PV.6498.
[732] S/RES/1973 (2011).
[733] Ibid.
[734] Ibid.
[735] Ibid.

no-fly zone to be administered through a 'ban on all flights' in Libya's airspace, with the exception of those needed to enforce it and those providing humanitarian relief. [736]

Despite the Libyan authorities' announcement they would honour the demand for an immediate ceasefire on 18 March, their continued advance on Benghazi suggested otherwise.[737] The next day, Secretary-General Ban reported that Libya's compliance with the ceasefire demand could not be confirmed.[738] On 19 March, France, UK and the US proceeded to enforce the no-fly zone by commencing military operations with the launch of air strikes against Libya.[739]

From an R2P perspective, the Security Council's decision to respond to the events unfolding in Libya in anticipation of an impending massacre in Benghazi would have had to be premised not only on evidence of one of the four atrocity crimes being committed, but also, that the Libyan authorities were 'manifestly failing' to protect their populations. The above overview of the key events in the first month of the Libyan crisis, with specific regards to the risk of atrocity crimes and the threshold for international action, suggests that, objectively, UN member states would have had good reasons to find the Libyan state to be 'manifestly failing' in its responsibility to protect. In addition, senior UN officials and regional organisations underscored factors indicative of the scale and gravity of the atrocities committed against Libyan civilians, explicitly invoked the 'manifestly failing' threshold and reminded the international community of its subsidiary responsibility to protect populations abroad in a bid to prompt Security Council action. The next step in this chapter is, therefore, to ascertain the extent to which UN member states were receptive of this portrayal of the Libyan state as 'manifestly failing', with a view to determining if UN member states actually contemplated Libya's 'manifest failing' in discussions over collective action in Libya (as they should have in theory).

[736] Ibid.

[737] The Guardian news blog, 'Libya military action - Friday 18 March', 18 march 2011 [Accessed 8 February 2020] Available at: https://www.theguardian.com/world/blog/2011/mar/18/libya-military-action-live-updates

[738] Security Council Report, 'Chronology of Events: Libya', 2 August 2019 (revised) [Accessed 8 February 2020]. Available at: https://www.securitycouncilreport.org/chronology/libya.php

[739] Ibid.

PART I: THE DEBATE OVER THRESHOLD IN LIBYA

As detailed in Chapter II, the mapping of the debate over threshold is a necessary first step in order to come to terms with the presence and breadth of relevant state discourses, the language used and any inconsistencies and ambiguities therein. Specifically, this part of the present chapter will shed light on the language and scope of the discourse over threshold in states' justifications for their decision to employ coercive measures in Libya and the wider debate over the 2011 crisis. Specifically, it will highlight the phrases states used in these debates when referring to the notion of threshold in a bid to gain a better insight into states' awareness and understanding of the 'manifest failing' threshold as spelled out at the 2005 World Summit. Thus, the first part of the analysis of the Libyan crisis will not only lay the groundwork for the second part of the chapter, but also lend some general insights into states' understanding of the threshold for collective action, which will inform the framework for 'manifest failing' inquiries articulated in the next chapter.

The discussion is structured thematically in accordance with the different ways in which states invoked the threshold for collective action, distinguishing between instances when states drew on concepts found in R2P documents – 'manifestly failing' and 'unable or unwilling' – and when they did not. In each of the two sections, the initial analysis of the P3 discourse within and without the Security Council will be succeeded by an analysis of the statements made by other UN member states.

Deliberating the threshold for collective action in the language of 'manifest failing' and 'unable or unwilling'

Although none of the P3 invoked the exact phrases 'manifestly failing' or 'unable or unwilling', some referred to the related notions of a 'failing' or 'unable' state when making the point that the Libyan state was not fulfilling its primary responsibility to protect its populations. For instance, in the UN Security Council discussion following the adoption of Resolution 1970, France's Permanent Representative to the UN, Mr. Gerard Araud, stated: '[t]he text, unanimously adopted today, recalls the responsibility of each State to protect its own population and of the international community to intervene when States fail in their duty'.[740] Paraphrasing Paragraph 139 of the WSOD, Araud referred not only to the host state's protection responsibilities, but also to the international

[740] S/PV.6491, p. 5.

community's subsidiary responsibility to protect and the trigger for interventionary action. Although the wording 'should… national authorities manifestly fail to protect' found in Paragraph 139 was replaced with the phrasing 'when States fail in their duty', there is little doubt that the latter constitutes a direct reference to the former, given the context in which it was invoked. Interestingly, Mr. Araud also referred to the Libyan authorities' 'inability' to protect their populations in a press statement of the same day: 'First, [the] responsibility to protect: in this very strong text we have recognized it, under Chapter VII. You know what it means: if a government *is not able* to protect its own population, it means the international community has the right and the duty to step in' (emphasis added).[741] Once again, this is a clear case where a reference to R2P's reactive dimension is coupled with a reference to the threshold for coercive action. However, on this occasion the French representative opted for an alternative wording, citing the Libyan state's 'inability' to protect. As was the case with Mr. Araud's use of the 'manifest failing' language, he did not make a verbatim reference to the 'unable or unwilling' formula. However, there is no evidence to suggest that the omission of the word 'unwilling' from the phrase implies a substantive objection to the assessment of a state's 'willingness' and an exclusive reliance on assessments of a state's 'ability'. It rather appears that the French ambassador merely employed abridged formulae of the two widely used terminologies, 'manifestly failing' and 'unable or unwilling' to relay the idea that the threshold for the adoption of non-pacific measures had been crossed. Although the interchangeable use of the 'unable or unwilling' and 'manifest failing' vocabularies is markedly inconsistent, the way in which they were both employed to paraphrase the text of the WSOD, in statements of the same day, indicates that they were understood to be identical in meaning with and representative of the 'manifest failing' threshold of the 2005 Outcome Document.

By contrast, whilst the US government was consistent in the language it used, the manner in which it discussed Libya's failing in its primary responsibility was more nebulous. In a statement of 23 February, Barack Obama made an explicit reference to the host state's responsibility to protect, followed by a reference to the failing of the Libyan government:

> Like all governments, the Libyan government has a responsibility to refrain from violence, to allow humanitarian assistance to reach those in need, and

[741] Gérard Araud, 'Libya/adoption of UNSCR 1970', Remarks to the press by Gérard Araud, France's Permanent Representative to the United Nations, 26 February 2011, New York [Accessed 28 January 2020] 1 March 2011. Available at: https://franceintheus.org/spip.php?article2193

> to respect the rights of its people. It must be held accountable for its failure to meet those responsibilities, and face the cost of continued violations of human rights.[742]

The wording 'failure to meet those responsibilities' invoked in the context of the R2P appears to be a nod to the 'manifest failing' threshold, despite the notable absence of an acknowledgement of the international community's responsibility to step in when this test is met. Whilst this early statement stops short of invoking R2P's reactive dimension, it still implies that a threshold will be crossed and 'costs' incurred by the Libyan government, should the violence persist. A similar reference to the Libyan regime's failing to fulfil its primary responsibility was made three days later by the US Permanent Representative to the United Nations, Ms. Susan Rice, in a statement to the Security Council after Resolution 1970 was passed:

> 'The Security Council has acted today to support the Libyan people's universal rights. These rights are not negotiable. They cannot be denied. Libya's leaders will be held accountable for violating these rights and for failing to meet their most basic responsibilities to their people'.[743]

Again, the prospect of international action is hinted at through a reference to the Libyan state's 'failing' in its protection responsibilities, but not openly recognised in the way that the French representative did. Despite the marked reluctance of the US government to explicitly justify the coercive actions taken by the international community with reference to R2P's reactive dimension, it remained consistent in its use of the 'failing' state trope. Consequently, the underlying message read that the Libyan government's failing to uphold its primary responsibility, manifested in the continuation of violence, will have (unspecified) ramifications. Despite falling short of citing the 'manifest failing' formulation verbatim and invoking the international community's responsibility to protect, the US government's references to the Libyan state's 'failing' in its protection responsibilities towards its people are suggestive of a very specific kind of state failing taking place – the regime's 'manifest failing' to protect its populations.

---

[742] Barack Obama, 'Remarks by the President on Libya', 23 February 2011 [Accessed 23 September 2019] Available at: https://obamawhitehouse.archives.gov/the-press-office/2011/02/23/remarks-president-libya

[743] S/PV.6491, pp. 3-4.

Although neither France, nor the US employed the 'manifest failing' or 'unable or unwilling' language consistently in government statements and UN meetings throughout the Libyan crisis, they both referred to Libya's 'manifest failing' at a crucial moment, when justifying their decision to adopt Resolutions 1970. Unlike its P3 counterparts, the UK government invoked neither of the two formulations in its contributions to the discourse surrounding the Libyan crisis. Despite this, the UK's justification and votes in favour of coercive action were explicitly tied to the notion that a threshold had been crossed in ways that will be detailed in the next section.

With the exception of France's and US's statements on Resolution 1970, none of the other Security Council member states employed the language of 'manifest failing', 'unable or unwilling' or related concepts when justifying their vote on Resolutions 1970 and 1973 in front of the Council. Notably, the two formulations were completely absent from deliberations surrounding the adoption of Resolution 1973. So were explicit references to the protection responsibilities attached to the international community. Offering an explanation for this notable absence, Jennifer Welsh purports that the coercive aspect of the R2P 'is still contested by some members of the Security Council'.[744] Likewise, Morris cites the controversy surrounding R2P's reactive dimension as a conceivable explanation for member states' reluctance to invoke it in a Chapter VII context in Libya, even though the R2P may have 'served as a motivating factor' or 'provided the conceptual framework through which [they] framed their policy options'.[745] Indeed, it would have been imprudent to appeal to any contentious concepts when speaking at the vote on Resolution 1973, mandating the use of 'all necessary measures' to protect civilians, given some member states' (i.e. Brazil, China, Germany, India and Russia) reservations over the use force in Libya. As demonstrated below, subsequent UN deliberations on the Libyan crisis also lend credibility to this supposition. In particular, the broader discourse on Libya, where securing the vote on resolutions authorising collective action was not at stake, saw an increase in the frequency of both explicit and implicit references to the external responsibility of the international community and the 'manifest failing' threshold.

[744] Jennifer Welsh, 'Civilian Protection in Libya: Putting Coercion and Controversy Back into RtoP', 25 (3) *Ethics and International Affairs* (2011): 255-262, p. 255. See also Tim Dunne and Jess Gifkins, 'Libya and the state of intervention', 65 (5) *Australian Journal of International Affairs* (2011): 515-529, pp. 521-522.

[745] Morris, 'Libya and Syria', p. 1274.

UN member states returned to the debate over the threshold for coercive action in discussions of the situation in Libya under agenda item 'the protection of civilians in armed conflict'. Whilst some only invoked the R2P in relation to Libya's host state responsibilities, the majority of participating states reaffirmed the international community's responsibility to take collective action and made explicit references to the 'manifest failing' or 'unable or unwilling' tests.[746] In the first of these meetings in May 2011, the most accurate reference to the international community's responsibility to protect and the 'manifest failing' threshold, as outlined in the 2005 WSOD, came from Croatia: 'We thus support the international community's increased preparedness to take collective action through the Security Council when national authorities manifestly fail to protect their populations from violations of humanitarian law'.[747] In the second meeting in November 2011, Brazil made an equally strong statement highlighting 'the responsibility of the international community to act collectively, through the United Nations, should national authorities manifestly fail to protect their populations'.[748] Both statements were made following briefings addressing the situation in Libya and in the context of a meeting, where the Libyan crisis received considerable attention. Even though neither of the two states specifically referred to Libya's 'manifest failing', they expressed blanket support for the international community's readiness to live up to their commitment to 'take timely and decisive' action in situations calling for the protection of civilian populations at the time, one of which was the atrocity crisis in Libya. Importantly, in so doing, Brazil and Croatia closely adhered to the 'manifest failing' language and did not shy away from invoking R2P's reactive dimension.

Other states explicitly referred to the 'failing' of Libyan authorities to indicate that the threshold for international action against Qaddafi's regime had been crossed. Germany made a particularly strong statement with reference to the situation in Libya and other cases where the international community faced the challenges of preventing and halting violence against civilians. In this context, it expressed its unwavering commitment to what was agreed at the 2005 World Summit, citing both the protection responsibilities of national governments and the international community, as well as the substantive threshold requirement for coercive action:

[746] UNSC, S/PV.6531 (Resumption 1) 10 May 2011, p. 2.
[747] Ibid. p.15.
[748] S/PV.6650, p. 16.

> ‘Germany firmly supports the principle of the responsibility to protect, including the responsibility of the international community, through the Council, to take appropriate action should the authorities concerned fail in their duty to protect civilians. Let me add that now is not the time for us to begin to step back from, or compromise on, the commitments that all of us have undertaken in endorsing the principle of the responsibility to protect’.[749]

To the same effect, Portugal offered a more generous interpretation of the text of the Outcome Document, following its condemnation of military attacks against civilians in Libya and elsewhere: ‘When civilians are targeted and national authorities or the conflicting parties fail in their obligation to protect them, the United Nations, and especially the Security Council, have the duty to speak up and the obligation to act’.[750] Likewise, citing the Security Council’s ‘responsibility to authorize international protection when States fail and betray their obligations to the extreme of widespread and indiscriminate killing of their own people’, Norway expressed its support for the ‘Council’s decisiveness in taking necessary measures under Chapter VII to protect civilians in both Libya and Côte d’Ivoire’.[751] Having condemned ‘the indiscriminate and excessive use of force’ against civilians in Libya and reaffirmed the host state’s primary protection responsibilities, Turkey introduced a variation of the ‘manifest failing’ formula by referring to international community’s ‘responsibility to help protect civilians in situations where States *openly fail* to do so’ (emphasis added).[752] What is more, Turkey made a straightforward declaration that ‘This is essentially what happened in Libya’.[753] By attaching the label of ‘failing state’ to Libya, the above UN members unequivocally acknowledged that the ‘manifest failing’ threshold had been crossed in this case – a judgement reinforced by the explicit affirmation of the international community’s responsibility to step in and protect persecuted populations.

The same language was also employed by states who were otherwise critical of the Libyan intervention, namely South Africa and India. Despite its open criticism of NATO activities in Libya for exceeding and ‘abusing’ the mandate of Resolution 1973,

[749] Ibid. p.27.
[750] Ibid. p.3.
[751] S/PV.6531 (Resumption 1), p. 11.
[752] Ibid. p.12.
[753] Ibid.

South Africa proceeded to reaffirm the concept of the international community's responsibility to protect.[754] In doing so, it 'underline[d] that [while] it remains the responsibility of States to protect civilians within their borders [,] [a]rmed opposition groups also bear a responsibility for ensuring that unarmed civilians are protected'.[755] Consequently, '[f]ailure by both State and non-State actors to uphold that responsibility should not go unpunished'.[756] South Africa went on to impress that 'accountability must… be sought at the national level', where the responsibility for determining the judicial mechanisms through which perpetrators are held account lies with the host state.[757] Ultimately, South Africa concluded that when a national government fails to ensure accountability, 'the international community has a collective responsibility to act with the utmost adherence to international law, in accordance with the purposes and principles of the Charter and as prescribed in the 2005 outcome document and the Constitutive Act of the African Union'.[758] While the WSOD does not recognise non-state actors as bearers of responsibilities to protect civilian populations in the way South Africa does, the failing of the host state to seek accountability for atrocities committed by state or non-state actors has surfaced as one of the indicators of a 'manifest failing' in Chapter I and 'unwilling or unable' to meet the threat posed by a non-state actor in Chapter III. In this sense, South Africa reaffirms the relevance of measures taken by the host state to ensure accountability as an indicator of the failing of the host state to protect its populations.

Whereas India also invoked the 'failing state' trope, it did so only to draw attention to the insufficient measures taken against non-state actors: 'There are also instances when the Council is expected to quickly criticize national Governments for failing in their responsibility to protect civilians, while little or no accountability is enforced on armed groups indulging in violence'.[759] However, this demand for non-state accountability is not suggestive of or linked to triggering the subsidiary protection responsibility of the international community. As such, it can only be seen as a reference to the primary responsibility of national governments and a nod to the 'manifest failing' threshold by

[754] S/PV.6650, pp. 22-23.
[755] Ibid., p. 23
[756] Ibid.
[757] Ibid.
[758] Ibid.
[759] S/PV.6650, p. 18

virtue of its reference to the 'failing' of national governments. While the two BRICS members are right to highlight the role of non-state actors in the commission of atrocities, it is important to reiterate that it is the role of the host state in mitigating the threat of atrocities that underlies the determination of whether the threshold for collective action has been met. Neither the R2P, nor international law recognise protection responsibilities on the part of non-state actors, which means that if the host state cannot respond effectively to the atrocity crimes perpetrated by the latter, it may be deemed to be 'manifestly failing' to protect its populations.

Despite the preponderance of references to the 'manifest failing' / 'failing' of host states, some opted for the 'unable or unwilling' language to signify that the threshold for collective action had been crossed. Before turning its attention to the civilian protection strategies employed in response to the crisis in Libya, Nigeria stated that when states 'are unable or unwilling to fulfil [their] obligation [to protect civilians], the international community, in particular the Security Council, must respond to the plight of civilians in armed conflict.[760] Australia, on the other hand, employed both the 'unable or unwilling' and 'manifestly failing' terminology in the following statement to the Security Council:

> Unfortunately, as we know, there are cases in which States are unwilling or unable to act. […] Victims and affected communities must be confident that when their own State fails in its responsibilities, because it is either unwilling or unable, the Council will act as guardian of their interests, wherever they happen to live.[761]

The above explanation that a 'state fails in its responsibilities, because it is either unwilling or unable' makes it quite clear that Australia sees the two formulas as equivalent in substance.

On balance, the above discussion of the discourse over threshold in Libya indicates that states were familiar with the conceptual toolkit introduced in the 2005 World Summit Outcome Document. The majority of states adhered to the 'manifest failing' terminology to refer to the threshold for coercive action. Some of them did so by

[760] UNSC, S/PV.6531, 10 May 2011, p. 19.
[761] S/PV.6650, p. 33.

referring to the host state's 'failing' in their responsibility to protect whilst reaffirming the protection responsibilities of national governments and/or the international community. References to the 'unable or unwilling' test were few and far between, but when made constituted definitive affirmations of the threshold for non-pacific action at the heart of R2P's reactive dimension. Some states used the 'manifest failing' and 'unable or unwilling' terminologies interchangeably as though they mean the same thing. Regardless of the language they chose to employ, states did not seek to alter or stretch the meaning and interpretation of the threshold for coercive action beyond the text agreed at the 2005 Summit.

Deliberating the threshold for collective action in the absence of the 'manifest failing'/'unable or unwilling' language

Even in the absence of explicit references to the phrases 'manifestly failing' or 'unable or unwilling', the notion of threshold for international action featured prominently in the discourse on the Libyan crisis and was often the centrepiece of state's professed justifications for coercive measures to protect civilians. A prominent case in point was Colombia's statement to the UN Security Council after the vote on Resolution 1973. Dunne and Gifkins named Colombia as the sole member of the Security Council to 'us[e] pillar three language to frame resolution 1973 when speaking at the vote', similarly to the way in which France did so at the meeting on Resolution 1970 (as noted above).[762] To illustrate, Colombia justified its position on Resolution 1973 as a vote 'in favour of measures that are aimed at protecting the civilian population from imminent attacks by a Government that, through its actions and statements, has shown that *it is not up to the international responsibility of protecting its population*' (emphasis added).[763] However, contrary to what Dunne and Gifkins suggest, Colombia's recognition of R2P's reactive dimension was not the straightforward reference to the international community's responsibility to protect it appeared to be. Rather, it was an inaccurate use of the phrase 'international responsibility' to refer to Libya's host state responsibility to protect. Further evidence of this can be found in the very next sentence of Colombia's statement, which cites resolution 1970 to the effect that it reminded Libya of its primary responsibility to protect its populations, but did not affirm the subsidiary responsibility of the international community. Put simply, the above reference to Libya's 'international responsibility of

[762] Dunne and Gifkins, 'Libya and the state of intervention', p. 521.
[763] S/PV.6498, p. 7.

protecting its populations' is plainly a reference to Libya's failing in its internal responsibility as a national government, rather than a reference to the external responsibility of the international community. All things considered, it can be said that Colombia certainly alluded to R2P's reactive dimension insofar as it clearly conveyed the idea that Libya was 'manifestly failing' by suggesting its behaviour had demonstrated that it was 'not up to the…responsibility of protecting its populations'. In this sense, Colombia's statement exemplifies a trend in the broader discourse on the Libyan crisis illustrated in this section, namely UN member states making the argument that the threshold for collective action was crossed without employing the 'manifest failing' or 'unable or unwilling' terminologies.

This tendency is notably exhibited in President Obama's justification for the US decision to intervene in his address to the nation two days after the commencement of military operations in Libya. Having painted a vivid picture of the human rights violations against Libyan civilians, the US President implied that a transfer of responsibility from the host state to the international community had taken place due to Libya's surrender of its primary responsibility towards its populations: 'Gadhafi has forfeited his responsibility to protect his own citizens and created a serious need for immediate humanitarian assistance and protection, with any delay only putting more civilians at risk'.[764] This way of relaying that the threshold for international action had been crossed can also be observed in the Security Council meetings at the vote on Resolutions 1970 and 1973, which saw states' justifications for their position on sanctioning coercive action in Libya focus on the nature and seriousness of the human rights violations and the escalation of violence. Specifically, Security Council member states proffered a range of factors that have been associated with the 'manifest failing' threshold to communicate that the gravity and magnitude of the human rights violations in Libya had reached a threshold that warranted the adoption of coercive measures. The exact manner in which they did so will be illustrated in the second part of this chapter, which focuses on identifying key determinants that will underwrite the framework articulated in the next chapter, based on the aspects of the situation in Libya that states invoked when making the case that the

[764] Barack Obama, 'Letter from the President regarding the commencement of operations in Libya', 21 March 2011 [Accessed 23 September 2019] Available at: https://obamawhitehouse.archives.gov/the-press-office/2011/03/21/letter-president-regarding-commencement-operations-libya

threshold for international action to protect vulnerable populations from the threat of atrocities had been crossed.

The British and French discourse on Libya also featured a variety of ways in which the idea of threshold was inserted into justifications for coercive action in Libya without drawing on the 'manifestly failing' or 'unable or unwilling' language. For instance, days after the adoption of Resolution 1973, David Cameron made a strongly-worded statement explaining why the UK 'could not have waited longer to use force':

> 'The United Nations gave Gaddafi an ultimatum and he completely ignored it. To those who say we should wait and see, I would say we have waited and we've seen more than enough... Gaddafi has had every conceivable opportunity to stop massacring his own people. The time for red lines, threats, last chances is over.'[765]

Appealing to constructs that are synonymous with the notion of a 'threshold', namely the Libyan authorities' flagrant disregard for the UN's 'ultimatum' and the crossing of 'red line', Cameron sent a strong message that the time for further action had come. This 'underlying logic of a threshold [, characterised by Gallagher as] a line [being] crossed in that what was tolerated yesterday cannot be today and action has to be taken',[766] was also illustrated in the French Prime Minister's justification in front of the National Assembly for 'what happened' that prompted France to act:

> 'Gaddafi was banking on the international community being powerless. Admittedly, we nearly descended into an endless cycle of appeals and warnings, whose sole consequence was offended speeches. France refused to accept this fate. The President of the Republic has chosen to act'.[767]

Although, neither the 'manifest failing' / 'unable or unwilling' terminologies, nor the concept of the 'responsibility to protect' played a role in this impassioned 'refusal to

---

[765] David Cameron, 'PM Statement to the House on Libya: Prime Minister David Cameron has secured comprehensive support from MPs for the military intervention in Libya', 21 March 2011 [Accessed 14 January 2021] Available at: https://www.gov.uk/government/speeches/pm-statement-to-the-house-on-libya

[766] Gallagher, 'What constitutes a "Manifest Failing"?', p. 436.

[767] François Fillon, 'Statement by the Government on the intervention of the Armed Forces in Libya for the implementation of UN Security Council Resolution 1973, and the debate on this statement', National Assembly (France), 22 March 2011 [Accessed 15 January 2020] Available at: http://www.assemblee-nationale.fr/13/cri/2010-2011/20110144.asp#INTER_0

accept the unacceptable',[768] it clearly conveyed the understanding that the situation in Libya demanded collective international action because Qaddafi's government had crossed the symbolic line of tolerable host state behaviour towards its own populations.

Other French government officials were more forthright in expressing this sentiment. In a meeting of the Security Council during the active intervention phrase, Gerrard Araud did so by paraphrasing the text of the 2005 WSOD:

> 'The responsibility to protect civilians belongs first, we all know, to national Governments. But when they do not fulfil that duty, and when serious violations of international humanitarian law and human rights — war crimes, crimes against humanity — are planned or committed, it is then the duty of the Security Council to intervene to protect civilian populations. There is no other choice when these atrocities are committed by Governments against their own populations'.[769]

In this generous rewording of the relevant paragraphs of the Outcome Document, Araud replaced the phrase 'manifestly failing' with the alternative formulation 'when states do not fulfil their duty'. In the given context, the latter is an unmistakable reference to the threshold for coercive action as prescribed in the WSOD. This commonly used formula of 'when x (insert threshold requirement(s)) occurs, then the international should intervene' was also employed by Chile in the following statement delivered at an earlier Security Council meeting: '*When States cannot protect their civilians*, the international community, through the United Nations, cannot remain indifferent to the fate of those whose rights are being seriously, systematically and repeatedly violated' (emphasis added).[770] Likewise, in another post-factum justification of the Libyan intervention the French representative stated that:

> 'When a Government attacks civilian populations instead of protecting them; when the atrocities committed sear the human conscience; and when the stability of an entire region is affected, the international community has

[768] Ibid.

[769] S/PV.6650, p. 19.

[770] S/PV.6531 (Resumption 1), p. 9.

> a responsibility to intervene and to protect civilians. That is what we done [sic.] in Libya'.[771]

The above statement is an apt example of how the idea of a figurative line being crossed is communicated through highlighting certain aspects of the situation that are indicative of the host state's manifest failing, namely the persecution of civilian population by the host state and the 'conscience-searing' gravity of the violent treatment they are subjected to.

Another prominent way in which states indicated that a threshold had been crossed without referring to the host state as either 'manifestly failing' or 'unable or unwilling' was through framing the host state as 'illegitimate' by virtue of the way it treated its populations. The French foreign minister Alain Juppé did so in his reply to a question in the National Assembly on 1 March: 'the Libyan regime, by suppressing popular protest movements with extreme violence, has lost all legitimacy'.[772] Juppé reiterated the same point in a session of the National Assembly on the following day. Again he suggested, that on account of the 'extreme brutality' with which Qaddafi quelled those who challenged him, he 'has disqualified himself, to such an extent that he has lost all legitimacy'.[773] Significantly, on this occasion the claim was directly followed by a reminder of what the UN recognised with the adoption of the R2P principle, namely the responsibility of individual states to protect their populations and the license for the international community to act in their stead 'if they do not do that'.[774] Within this context, invoking the 'loss of legitimacy' trope served to drive the point that the Qaddafi government had crossed the red line that marks the transfer of the responsibility to protect the Libyan populations form the host state to the international community.

Although references to the Libyan state's 'illegitimacy' / 'loss of legitimacy' as a result of atrocity crimes perpetrated against its civilian population were a prominent fixture in the discourse on the Libyan intervention, they were not as explicitly linked to the notion that a threshold had been passed as seen in the above example. Still, the phrase was frequently employed in a relevant, albeit more general, sense to give weight to the gravity

[771] S/PV.6528, p. 13.

[772] National Assembly (France), 2nd session of 1 March, 1 March 2011 [Accessed 17 January 2020] Available at: http://www.assemblee-nationale.fr/13/cri/2010-2011/20110128.asp

[773] National Assembly (France), 1st session of 2 March, 2 March 2011 [Accessed 17 January 2020] Available at: http://www.assemblee-nationale.fr/13/cri/2010-2011/20110130.asp

[774] Ibid.

of the situation and imply that it had crossed the threshold of acceptable host state behaviour. Notably, Nicolas Sarkozy justified the UN Security Council's decision to intervene on the grounds of the regime's illegitimacy 'to protect civilians from the murderous madness of a regime which, in killing its own people, has lost all legitimacy'.[775] This argument was reiterated in a joint letter by the UK and France, which underscored the continued violent attacks against civilians 'with aircraft and helicopters', in spite of international calls and the demands of Resolution 1970 for an end to the violence.[776] The same points were raised by the UK after the vote on Resolution 1973, where it declared Libya was 'a violent, discredited regime that has lost all legitimacy'.[777] Germany was equally categorical that the Qaddafi regime 'has lost all legitimacy and can no longer be an interlocutor for us',[778] whereas Portugal stated that it had 'lost all its credibility and legitimacy vis-à-vis its own population and the international community'.[779] The language of 'illegitimacy' was also used by Lebanon[780] and Colombia[781] as a means to highlight that the human rights violations against Libyan civilians were unacceptable. What these instances illustrate is that the very idea of a threshold being crossed is inherent in all arguments that a state has 'lost all legitimacy' on account of the grave and systematic violations of the human rights perpetrated against its civilian populations, even when such claims are not explicitly linked to the threshold for coercive international action. Put simply, claiming that the brutality with which a state treats its civilian populations is so extreme as to render the state in question utterly illegitimate is one way to convey the notion that a threshold has been crossed.

It is worth noting that denouncing Qaddafi's legitimacy as a leader also extended beyond the concern with the plight of innocent populations in Libya and the latter's failing in its responsibility to protect. Specifically, the 'loss of legitimacy' argument was frequently succeeded by calls for Qaddafi to step down from power. As early as 26 February, President Obama stated that 'when a leader's only means of staying in power is to use

[775] Nicolas Sarkozy, 'Paris Summit for the Support of the Libyan People – Statement by Nicolas Sarkozy, President of the Republic', 19 March 2011 [Accessed 23 September 2019] Available at: https://franceintheus.org/spip.php?article2241

[776] David Cameron and Nicolas Sarkozy, 'Letter from David Cameron and Nicolas Sarkozy to Herman Van Rompuy', 10 March 2011 [Accessed 14 February 2020] Available at: https://www.theguardian.com/world/2011/mar/10/libya-middleeast

[777] S/PV.6498, p. 4.

[778] Ibid. p.4

[779] Ibid. p.8

[780] Ibid. p.3

[781] Ibid. p.7

mass violence against his own people, he has lost the legitimacy to rule and needs to do what is right for his country by leaving now'.[782] In an interview on 17 March, Juppé went as far as to summon the Qaddafi regime's loss of legitimacy 'by using violent force against its populations' in support of the claim that the civilian protection military campaign in Libya could extend to Qaddafi's fall from power, should he fail to comply with the demands of UN Security Council Resolutions.[783] However, even in these instances where states employed the 'loss of legitimacy' trope to justify professed objectives that were not purely humanitarian, they still premised their claims of Libya's illegitimacy on the notion of a gravity threshold being crossed with reference to acts of violence or mass atrocities against civilian populations in Libya.[784] In sum, the discourse over the adoption of coercive action to protect civilian population in Libya suggests that declaring a state 'illegitimate' can be used to openly argue that the 'manifestly failing' threshold has been crossed or more implicitly to suggest that a 'manifest failing' is under way. In fact, whilst claims that a state is losing its legitimacy rarely amounted to explicitly framing the threshold for coercive action in terms illegitimacy, they were often employed to indicate that a threshold has been crossed by virtue of the gravity of the human rights violations taking place in the host state. As evidenced in the statements cited above, this line of argumentation was frequently pursued not only by the P3, but also by the rest of the Council.

---

[782] The White House Office of the Press Secretary, 'Readout of President Obama's Call with Chancellor Angela Merkel of Germany', 26 February 2011 [Accessed 14 February 2020] Available at: https://obamawhitehouse.archives.gov/the-press-office/2011/02/26/readout-president-obamas-call-chancellor-angela-merkel-germany; See also Statement by Ms Rice (US): S/PV.6491, p.3.; Barack Obama, 'Remarks by President Obama and President Calderón of Mexico at Joint Press Conference', 3 March 2011 [Accessed 14 February 2020], Available at: https://obamawhitehouse.archives.gov/the-press-office/2011/03/03/remarks-president-obama-and-president-calder-n-mexico-joint-press-confer

[783] Alain Juppé, 'Press conference given by Alain Juppé, Ministre d'Etat, Minister of Foreign and European Affairs', 17 March 2011 [Accessed 23 September 2019] 21 March 2011. Available at: https://franceintheus.org/spip.php?article2235

[784] There were a number of instances in the discourse over Libya, in which Qaddafi's 'loss of legitimacy' was foregrounded on arguments other than the grave violations of human rights taking place within the country. For instance, David Cameron suggested that Qaddafi's regime was 'illegitimate' because he had 'lost the consent of his people'.: Cameron, 'Prime Minister's statement on Libya', 28 February 2011. Likewise, President Obama repeatedly argued that Qaddafi had 'lost the legitimacy to lead' because he had 'lost the confidence of his people'.: Barack Obama, 'Remarks by the President on the Situation in Libya', 18 March 2011 [Accessed 25 January 2020] Available at: https://obamawhitehouse.archives.gov/the-press-office/2011/03/18/remarks-president-situation-libya; Barack Obama, 'Weekly Presidential Address: The Military Mission in Libya', 26 March 2011 [Accessed 25 January 2020]. Available at: https://obamawhitehouse.archives.gov/the-press-office/2011/03/26/weekly-address-president-obama-says-mission-libya-succeeding; Barack Obama, 'Remarks by the President in Address to the Nation on Libya', 28 March 2011 [Accessed 21 January 2021] Available at: https://obamawhitehouse.archives.gov/the-press-office/2011/03/28/remarks-president-address-nation-libya

That states often judged the illegitimacy of the Qaddafi government based on the seriousness of the human rights abuses the Libyan populations were subjected to is not surprising. At the heart of the R2P concept lies the principle that 'sovereignty implies responsibility', a responsibility of each state to protect its own populations from the four crimes, which limits its power 'to do what it wants to its own people'.[785] This contemporary understanding of sovereignty as 'limited by reciprocal responsibilities between the state and its citizens' is hardly novel.[786] It is prominently expressed by John Locke through the conception of the 'social contract', which frames sovereign legitimacy as originating from an implicit societal agreement, by which an individual surrenders a measure of their freedom to a larger community in return for 'whoever has the legislative or supreme power of any common-wealth [ensuring] the peace, safety, and public good of the people'.[787] In simpler terms, the protection responsibilities that a state owes its populations are weaved into the social contract. And just as an individual who fails to act responsibly and abide by established societal rules may lose certain rights guaranteed by this contract, so can a state who breaks the social contract by failing to fulfil its reciprocal responsibility to protect its populations lose its legitimacy as a sovereign. As Marks and Cooper conclude, 'when the state manifestly fails to fulfil its responsibilities to its citizens, it loses legitimacy derived from the fulfilment of the social contract'.[788] In other words, the 'manifest failing' threshold embodies the idea that the primary protection responsibility of states limits their sovereignty, whereby should the state fail to fulfil its obligations towards its populations, it forfeits its legitimacy as a sovereign. Thence, R2P's reactive dimension entails that a state which has thus compromised its sovereignty (and with it the associated protection against external intervention under international law) may 'be liable to enforcement action'.[789] As Michael Doyle puts it, 'States are still sovereign, independent in their domestic affairs, but they are no longer free to commit one of [the] four [atrocity] crimes without risk of legitimate international constraint'.[790]

In this line of argumentation, the 'manifest failing' threshold can be understood as synonymous with a state's loss of legitimacy. Hence, it is hardly surprising that states

[785] ICISS, 'The Responsibility to Protect', p. XI, para. 1.35.

[786] Marks and Cooper, 'The Responsibility to Protect', pp. 93-94.

[787] John Locke, *Second Treatise of Government* (1690). C. B. Macpherson (ed.). (Indianapolis: Hackett Publishing, 1980) para. 131. Quoted in Marks and Cooper, 'The Responsibility to Protect', p. 94.

[788] Marks and Cooper, 'The Responsibility to Protect', p. 94.

[789] See Doyle, 'The Politics of Global Humanitarianism', pp. 15-16.

[790] Ibid. p. 16.

framed the Libyan state as 'illegitimate' in a way that suggested that a substantive threshold for interventionary action had been crossed. Whilst the R2P does not condone violent regime change, some human rights theorists advocate that when a state forfeits its legitimacy in the way described by social contract theory, a reciprocal right to overthrow the illegitimate government arises.[791] This explains why calls for regime change and the suggestion that a threshold for coercive action had been crossed often appeared together in state discourses when UN members labelled Qaddafi's government as illegitimate. As illustrated above, the context in which such claims are made can provide clues as to whether references to a host states' loss of legitimacy can be seen as suggestive of its 'manifest failing' or not. This will be taken into account when determining which references to 'illegitimacy' can help us identify the criteria states saw as relevant to determining whether Libya was 'manifestly failing' in its 'responsibility to protect'.

The evidence provided in this section lends further support to the argument that the notion that a threshold had been met featured prominently in state discourses on the Libyan crisis and state justifications for the need to take collective action. Even in the absence of the 'manifest failing' or 'unable or unwilling' formulations and/or explicit references to the international community's collective responsibility to protect, the underlying logic of a threshold being crossed was clearly discernible at the heart of high-level state discourses on the Libyan intervention.

Conclusion

To conclude the first part of this chapter, it is worth returning to one of Bode's unsubstantiated claims discussed in Chapter I, namely that states 'kept alluding to the "manifest failure" of Libya and what this means for the responsibility of the international community'.[792] With regards to the latter question, the above analysis of state discourses on Libya suggests that states generally recognised that a 'manifest failing' on the part of the host state triggers the international community's responsibility to protect. As for the former, there is ample evidence that the UN Security Council did much more than just 'allude to' Libya's 'manifest failing'. Although word-perfect references to the 'manifestly failing' test were rare, a number of states invoked R2P's threshold for coercive action by employing abbreviated or alternative formulations, namely the 'failing state' trope or

[791] Ibid.

[792] Bode, '"Manifestly Failing" and "Unwilling or Unable"', p. 8.

'unable or unwilling' terminology. Both the 'manifest failing' and 'unable or unwilling' formulations, as well as affirmations of R2P's reactive dimension, were notably absent from deliberations on the vote on Resolution 1973 for reasons detailed above. However, states cited these phrases or related concepts freely in other Security Council meetings where the Libyan crisis was discussed, including the one at the vote on Resolution 1970 authorising Chapter VII measures short of the use of force. Whilst states occasionally drew on the 'unable or unwilling' formulation instead of 'manifestly failing' or used both phrases interchangeably, there was no evidence to suggest that their preference for one formulation over the other was based on a perceived substantive difference between the two terminological siblings. Whether they stayed true to the wording of the WSOD or not, states appeared to be talking about one and the same thing – a substantive threshold requirement for international action in response to atrocity crimes.

In addition to these straightforward references, there were many instances in which the P3 and other UN members invoked the notion of threshold without employing the 'manifest failing' or 'unable or unwilling' language. These included but were not limited to: simply paraphrasing (more or less liberally) the idea of a threshold for coercive action expressed in Paragraph 139 of the WSOD, professing a range of factors which together convey the argument that a threshold had been crossed, claiming that the host state had lost its legitimacy on account of the way it treats its populations or otherwise implying that a transfer of responsibility from the host state to the international community had taken place. This shows that a simple search for the phrases 'manifestly failing,' 'unable or unwilling', and all imaginable variations of these formulations in UN documents is unlikely to produce an accurate picture of the presence, prominence and significance of the debate over threshold. However, when one looks at the broader discourse on coercive action in Libya in more depth, it is clear that states (including all of the P3) frequently invoked the notion of threshold or the underlying logic of threshold in various ways.

Even though there are marked inconsistencies in the language used by states to represent the 'manifest failing' threshold, generally speaking, UN members demonstrated awareness of this notion. Overall, the analysis of the Libyan case thus far provides evidence of a strong presence of the debate over threshold, indicating that the latter was one of the key considerations that states felt the need to address when deliberating on and justifying the adoption of forcible measures, as well as in post-factum reflections on the appropriateness of coercive action. The second part of this chapter will add to this host of

evidence by demonstrating that in addition to invoking the notion of a substantive threshold requirement in debates over international action, states also had concrete ideas in mind as to what it entailed. It is these more concrete notions of threshold that will provide additional foundations for the composite framework presented in the next chapter.

PART II: THE DETERMINANTS AND INDICATORS OF LIBYA'S 'MANIFEST FAILING'

Having demonstrated that states appealed to a substantive threshold for international action in discussions over the appropriateness of coercive measures in response to the atrocity crisis in Libya, attention now turns to the concrete factors that states professed when making the case that the host state is not fulfilling its primary responsibility to protect its populations adequately. Specifically, the second part of this chapter will draw on instances of justificatory discourse when states sought to defend their position on the adoption of Chapter VII enforcement measures against Libya in order to extract the factors (indicators and determinants) of a 'manifest failing' from discernible trends in the information / evidence states tend to offer when making the case that the host state is failing (or not) in its responsibility to protect. It is these determinants and indicators that will be triangulated with those identified in Chapters I and III to form the basis for the detailed normative framework for 'manifest failing' inquiries presented in Chapter V.

The second part of the chapter is organised in four sections, one for each of the P3 states and one focusing on the justifications offered by the other 12 members of the Security Council. This separation is maintained with a view to allowing for comparisons to be drawn amongst the P3 states and between the P3 states (collectively) and the rest of the Council. In doing so, this chapter will lend an insight into the following questions. Was there convergence or dissonance / tension amongst the P3 with regards to the understanding of the threshold for collective action, including the associated factors they professed? Did the rest of the Security Council echo or criticise the P3's justifications for the adoption of coercive measures and legitimacy claims based on the notion that a threshold has been crossed? What factors concerning the notion of threshold did the P2 and E10 states underscore in their statements to the Council and to what extent did these intersect with the ones offered by the P3?

FRANCE

It is undeniable that France played a pivotal role in marshalling international support for the Libyan intervention. France was the first P3 state to publicly acknowledge the situation in Libya as early as 21 February with a communique by the Presidency of the Republic censuring the use of force against Libyans and calling for an end to the

violence.[793] President Sarkozy followed up swiftly with a statement to the Council of Ministers on 23 February, in which he reiterated the demand for an immediate halt to the violence in the face of 'massive human rights violations' and asked the foreign minister to propose the prompt adoption of sanctions against those responsible for the continuing violence to France's European Union partners.[794] The French initiative continued with a request that the UN Human Rights Council met to discuss the situation in Libya and that the resolution negotiated at this meeting requested Libya's suspension from the Council.[795] Not long after, France, along with Germany, the UK and the US, participated in the drafting of Resolution 1970, proposing the adoption of sanctions against those involved in the violence against Libyans.[796]

On 25 February, France's Permanent Representative to the UN, Gérard Araud, provided the first clear indication that the situation in Libya had passed the threshold for robust international action. Specifically, Mr. Araud unequivocally stated that '[t]he seriousness of the situation in Libya requires a strong response from the international community'.[797] Whilst he did not call for the adoption of specific measures, he made it clear that the situation was of a gravity that warranted triggering the international community's involvement. The next day, as noted in the first part of this chapter, Mr. Araud cited the international community's responsibility to protect and the threshold for collective action, by referring to Libya's 'failing' to protect its populations in the Security Council debate following the adoption of Resolution 1970 and declaring the Libyan state 'unable' to fulfil its protection responsibilities in a press statement of the same day.[798] It is clear that at this point the French government was of the opinion that the threshold for international action had been crossed. Their activism in promoting the adoption of Chapter VII measures short of the use of force under Resolution 1970 was a material manifestation of this view.

---

[793] Nicolas Sarkozy, 'Situation in Libya', Communiqué issued by the Presidency of the Republic, 21 February 2011, Paris [Accessed 17 January 2020]. Available at: https://franceintheus.org/spip.php?article2161

[794] Nicolas Sarkozy, Statement at the Council of Ministers on the crisis in Libya, 23 February 2011, Paris [Accessed 17 January 2020] Available at: https://franceintheus.org/spip.php?article2176

[795] French Ministry of Foreign and European Affairs, 'Situation in Libya', Communiqué issued by the Ministry of Foreign and European Affairs, 24 February 2011, Paris [Accessed 17 January 2020] Available at: https://franceintheus.org/spip.php?article2174

[796] Gérard Araud, 'Libya (UN/Sanctions/Libya)', Remarks to the press by Gérard Araud, France's Permanent Representative to the United Nations, 25 February 2011, New York [Accessed 23 September 2019] 1 March 2011. Available at: https://franceintheus.org/spip.php?article2192

[797] Ibid.

[798] Araud, 'Libya/adoption of UNSCR 1970', 26 February 2011.

Even at this early stage of international involvement, when the use of force was 'not an option' for the French government,[799] several determinants and indicators of a 'manifest failing' started to crystallise in its justificatory discourse over collective action in Libya. As already mentioned, the gravity / seriousness of the situation was cited as a catalyst for international response by the French representative to the United Nations. Elsewhere, foreign minister Juppé defined the situation in Libya as one of 'exceptional gravity', before proceeding to argue that Qaddafi had 'lost all legitimacy' and invoke the international community's responsibility to protect.[800] The trigger for international action found a more concrete expression in the debate following the vote on Resolution 1970, where Mr. Araud cited 'the continued brutal and bloody repression and the threatening statements made by the Libyan leadership' as the grounds for UN demands to put an end to the violence before reminding the Council of the international community's responsibility in case a 'state fails' to protect its populations.[801] The evidence of ongoing violence and declarations of intent to harm are both indicators that point to the risk of imminent atrocities. Prospective assessments of the likelihood of further violence were pivotal to determining not only whether there was a cause for international action in Libya in late February, but also whether the situation warranted a stronger coercive response thereafter. In a bid to justify further coercive action against Libya at the National Assembly, Juppé spoke of the prospective death toll in case of inaction, i.e. 'the life and death of thousand men'.[802] 'The prospect of an imminent massacre', together with the escalation of violence against civilians 'which only grew fiercer' in the aftermath of Resolution 1970, was the very justification the French Foreign Ministry offered for the 'urgent adoption of Resolution 1973', the day after the UN Security Council vote.[803]

Naturally, as the crisis in Libya deteriorated, the French position over the issue of threshold came increasingly into focus. Specifically, following the meeting of the European Council on 11 March, the French president explicitly specified the threshold conditions that would inform the government's consideration of stricter and more

[799] Araud, 'Libya (UN/Sanctions/Libya)', 25 February 2011.

[800] National Assembly, 2 March 2011.

[801] S/PV.6491, p. 5.

[802] National Assembly, 1st session of 8 March, 8 March 2011[Accessed 23 August 2020] Available at: http://www.assemblee-nationale.fr/13/cri/2010-2011/20110134.asp

[803] Ministry of Foreign and European Affairs, 'France/Libya/UNSCR 1973', Statements issued by the Ministry of Foreign and European Affairs Spokesperson (excerpts), 18 March 2011, Paris [Accessed 29 January 2020] 21 March 2011. Available at: https://franceintheus.org/spip.php?article2236

intrusive measures, including the use of force. Rephrasing the text of the resulting European Council declaration in a statement to the press, Sarkozy underscored that for the purpose of civilian protection the Council had agreed to 'consider all the necessary options, provided there is a *demonstrable need for action*, *a clear legal basis* in the United Nations, and *regional support*, by which I mean the Arab League' (emphases added).[804] The president later unpacked the three threshold criteria when asked to elaborate on the Franco-British proposal for defensive operations directed at military targets in Libya if Qaddafi was to use air force against civilian populations.[805]

Sarkozy first noted that the question of 'what happens' in this scenario has been 'raised for a number of countries'.[806] Significantly, the French president revealed: 'the question Mr Cameron and I asked was: at that point, do we stand back and watch or must we react?'[807] This statement provides unambiguous evidence that world leaders purposefully discussed the question of whether (and when) the threshold for international action is crossed in the event of 'defenceless civilian populations being massacred in a large-scale military action against them'.[808] Sarkozy detailed the French position on this matter as follows:

> 'We set certain conditions: a United Nations mandate, from the Security Council, the support of partners – we were thinking, of course, of the Arab League and also those Libyan authorities whom we recognize as political interlocutors – and of course large-scale military attacks on unarmed, non-violent, civilian populations. It's in that context that the question [whether to act or not] arises'.[809]

It is clear that the above conditions correspond to and expand on the three prerequisites for considering 'all necessary options' that the French president had already invoked. Sarkozy's statement reveals that a Security Council mandate for coercive action is what provides *a clear legal basis* for international response, whereas the backing by regional

---

[804] Nicolas Sarkozy, 'Situation in Libya and the Mediterranean – Extraordinary European Council – Press conference given by Nicolas Sarkozy (excerpts)', 11 March 2011, Brussels [Accessed 12 November 2019]. Available at: https://franceintheus.org/spip.php?article2223; For the text of the declaration see: European Council, EUCO 7/1/11 (Rev 1), 11 March 2011.

[805] Ibid.

[806] Ibid.

[807] Ibid.

[808] Ibid.

[809] Ibid.

organisation and internationally recognised host state authorities characterises the *regional support* requirement. Hence, 'large-scale military attacks on…civilian populations' could only epitomise the so-called *demonstrable need* for action.

The president went on to elaborate on the latter requirement by stating that an external intervention, albeit not envisioned at the time, could be contemplated in 'a specific scenario in which systematic attacks were made on civilian populations who were massacred by military means'.[810] This clarification of what the *demonstrable need* requirement stands for suggests that it encompasses an assessment of the gravity of the situation. Support for the existence of a 'demonstrable need' in line with Sarkozy's depiction of this condition as being manifested in systematic or large-scale violence against peaceful populations inflicted by military means is ample in the French government's justificatory discourse on the Libyan intervention. The manifestations of these indicators of 'demonstrable need' include assertions of 'extreme violence',[811] 'crimes against humanity',[812] 'war crimes'[813] (which variously attest to the presence of large-scale or systematic violence), as well as explicit justifications of the military intervention in Libya and its compliance with the UN mandate with reference to the weapons used (aircraft and ships) by Qaddafi's forces for attacks on civilians and cities.[814]

Following the adoption of Resolution 1973, the demonstrable need requirement saw another iteration in the French prime minister's statement to the National Assembly on 22 March, informing parliament of the decision to deploy troops in Libya. In particular, Fillon claimed that the French position had always been '[c]lear that any intervention should be conditional upon [the] prerequisites' of demonstrable need, regional support and a solid legal basis, which he argued had now been fulfilled:

> 'A demonstrable need on the ground. Who has not noticed it?
> Support from the countries of the region. The call of the Arab League gives us that.

[810] Ibid.

[811] National Assembly, 1 March

[812] S/PV.6491, p. 5; S/PV.6498, p. 2.

[813] Alain Juppé, 'Libya/Côte d'Ivoire', Interview on Libya to "France 2" television, 19 March 2011, Paris [Accessed 28 January 2020] 21 March 2011. Available at: https://franceintheus.org/spip.php?article2238

[814] Alain Juppé, 'London Conference on Libya/Consensus', Press conference given by Alain Juppé (excerpts), 29 March 2011 [Accessed 6 November 2019] Available at: https://uk.ambafrance.org/Alain-Juppe-s-press-conference,18821

> A solid legal basis. We have it with the adoption of Security Council Resolution 1973'.[815]

In this initial description of the four conditions, the 'demonstrable need' requirement stands out as the only one that is not backed up by a definition, explanation or concrete evidence as to how it had been met in the Libyan context. Nonetheless, its content can be gleaned from Fillon's subsequent outline of the reasons for the French involvement in Libya. Namely, the prime minister first underscored that the decision to intervene was not made 'overnight' but followed a succession of diplomatic measures enacted to supress the violence, including the adoption of Resolution 1973 by the UN Security Council, the reaction of the European Council and the G8, as well as condemnations of the violence in Libya and calls for action by the African Union (AU), the Organisation of Islamic Cooperation (OIC, formerly the Organisation of the Islamic Conference) and the League of Arab States (LAS).[816] Against this backdrop, Fillon argued that 'the use of force became the only solution' when it was clear that Qaddafi's regime was not deterred by any of these warnings, sanctions and calls for an end to his campaign of violence against civilian populations in Libya.[817] A similar justification for international action was offered by Juppé at the vote on Resolution 1973, when he stated that Qaddafi's disregard for the demands of Resolution 1970 and increase in the violence against civilian populations were an 'intolerable provocation' which prompted a sweeping reaction from the international community with regional organisations calling for an end to the violence.[818] In a bid to justify the need for Security Council action (and urge member states to vote in support of the draft of Resolution 1973), the French foreign minister proceeded to stress that '[d]espite these calls for peace, the situation in Libya today is more alarming than ever'.[819] The above statements reveal that France's judgement as to whether the substantive threshold for coercive action had been crossed was also premised on a determination of host state conduct in response to demands and actions taken by the international community, which also surfaced as a determinant of a 'manifest failing' in Chapter III.

---

[815] National Assembly, 1st session of 22 March, 22 March 2011 [Accessed 15 January 2020] Available at: http://www.assemblee-nationale.fr/13/cri/2010-2011/20110144.asp#INTER_0

[816] Ibid.

[817] Ibid.

[818] S/PV.6498, p. 2.

[819] Ibid.

In sum, the French government specified three conditions that need to be met for coercive action to take place (demonstrable need, clear legal basis and regional support). In so doing, it made clear that questions of legality, legitimate authority, political will, and consequences of such action only arise after the need for action has been established, hence the inclusion of the *demonstrable need* requirement. In the three-prong framework of the French government, the latter represents the substantive threshold requirement for collective action, whose consideration is necessary to determine whether or not states should act to protect populations abroad as a matter of principle. In this sense, the 'demonstrable need' requirement, as depicted in French justificatory discourse over protective action in Libya, fulfils the same function as the 'manifest failing' threshold does in the R2P framework specified in the WSOD. The subsequent analysis of the UK justificatory discourse on the adoption of coercive measures in Libya lends additional support to this argument.

Thus, the French government invoked the notion of threshold in various ways to justify its position to take action against Libya, be it by referring to the host state's 'failing' or 'inability', the 'demonstrable need on the ground', or otherwise raising and discussing the question of the circumstances under which international action is warranted. Most importantly, the above analysis of French justificatory discourse over collective action in response to the Libyan crisis shed light on a range of factors that the French government professed to make the case that the substantive threshold for international involvement had been crossed. Notably, the adoption of coercive measures was justified with reference to evidence of atrocity crimes taking place in Libya, the gravity/seriousness of human rights violations against innocent populations, the imminent risk of further atrocities, evidence of public statements of intent and systematic perpetration, and the irresponsiveness of the Libyan authorities to international requests to quell the threat of atrocities. These determinants and the indicators associated with them will be reflected in a table at the end of this chapter (Table 4.2), which brings together the determinants and indicators of Libya's 'manifest failing' cited by states in UNSC debates on the Libyan intervention.

UK

The UK was comparatively slow to take a leading role in the international response to the Libyan crisis. David Cameron's initial reluctance to commit to any forcible action at the outset was exhibited in his unenthusiastic reception of Sarkozy's call of 23 February for

the European Union to impose sanctions against Qaddafi's regime.[820] Whist the prime minister did not exclude the prospect of sanctions in the future, should the situation in Libya deteriorate, he made it clear that the threshold for such actions had not been reached at the time: 'I do not think we are at that stage yet. We are at the stage of condemning the actions Colonel Gaddafi has taken against his own people'.[821] This talk of different 'stages' of a crisis being linked to particular consequences or actions taken by the international community suggests that the question of whether a certain threshold had been crossed or not was active in the prime minister's mind. To be precise, in his view, the situation in Libya warranted acts of public condemnation but was not of a sufficient gravity to meet the threshold for adopting Chapter VII coercive measures. Despite Cameron's initial lukewarm response to the French initiative to pursue such actions in Libya, on 24 February he followed Obama's lead in condemning the violence against civilians and issuing warnings of prospective consequences in the UN Security Council, including sanctions, and holding perpetrators to account.[822] The UK prime minister condemned the unacceptable violence authored by the Libyan government, underscoring the weapons used to inflict it: 'this meting out of vicious, brutal repression, including using aeroplanes, including troops on the streets, including live ammunition'.[823] It was not long before the UK started to play an active role in marshalling the international response to the crisis, notably evidenced by its participation in the drafting of Resolution 1970.[824]

Moving beyond public condemnation of the violence in Libya and warnings of further action, the adoption of Resolution 1970 on 26 February signalled that the UK government's position regarding the threshold for international action had hardened in a matter of days. In a statement in front of the Council following the vote on Resolution 1970, the UK permanent representative to the UN cited 'the appalling situation in Libya', underscoring the unacceptable 'violence … and the incitement to further violence by Colonel Al-Qadhafi', as the reason behind the UK's support for coercive measures against

[820] Nicholas Watt and Patrick Wintour, 'Libya no-fly zone call by France fails to get David Cameron's backing', 23 February 2011 [Accessed 29 January 2020] Available at: https://www.theguardian.com/world/2011/feb/23/libya-nofly-zone-david-cameron

[821] Ibid.

[822] David Cameron, 'Transcript of the PM's YouTube interview in Oman', 24 February 2011. [Accessed 29 January 2020] 26 February 2011. Available at: https://www.gov.uk/government/speeches/transcript-of-the-pms-youtube-interview-in-oman

[823] Ibid.

[824] S/PV.6491, p. 2.

Libya and its participation in presenting the text of the resolution.[825] In a statement to the UK parliament of 28 February, Cameron spoke of ramping up the pressure on the Qaddafi regime through further sanctions, suspending Libya's membership from international organisations, and did not 'in any way rule out the use of military assets'.[826] The prime minister's warning of a sharper international response culminated in the statement: 'We must not tolerate this regime using military force against its own people'.[827] Up to that point, the UK's broader understanding of the threshold for collective action was comparable to that of the French government at the same juncture of the crisis in Libya. Specifically, the British government deemed the threshold for coercive measures such as travel bans, asset freezes and an arms embargo to have been met on account of the seriousness of the existing human rights violations against Libyan civilians, incitement to violence and the weapons used, pointing to the government's role in the commission of atrocities. Thus, the UK's understanding of what determines whether the threshold for collective action has been met overlaps with and reaffirms the determinants of a 'manifest failing' identified in Chapter I, specifically the gravity of human rights violations and host state intent, manifested in both public statements of intent and indicators of systematic preparation.

The position of the UK government on the threshold for international action crystallised further at a press conference of 11 March, when David Cameron reiterated the commitment of the European Council to 'examine all necessary options [in order to protect the civilian population in Libya] provided there's demonstrable need, a clearly good basis and support from the region'. [828] The importance of contingency planning, alongside the objectives of holding Qaddafi to account and stepping up the pressure against his regime, was raised in the context of what Cameron portrayed as a 'dangerous moment' in Libya, characterised by 'barbaric acts, with Qadhafi brutally repressing a popular uprising led by his own people and flagrantly ignoring the will of the international community'.[829] Adding to these worrying aspects of the situation in Libya was the palpable risk that 'things may be getting worse, not better on the ground'.[830] Cameron put

[825] Ibid. p. 3.

[826] Cameron, 'Prime Minister's statement on Libya', 28 February 2011.

[827] Ibid.

[828] David Cameron, Speech at a press conference in Brussels, 11 March 2011, Brussels [Accessed 12 November 2019]. Available at: https://www.gov.uk/government/speeches/press-conference-in-brussels

[829] Ibid.

[830] Ibid.

the risk of further mass violence into perspective by turning to the Qaddafi regime's track record of violent conduct, specifically its continuous support for terrorism globally and 'the biggest mass murder ever on British soil, the Lockerbie bombing, as well as being associated with the deaths of many innocent people around the world'.[831] In addition to the underlying humanitarian imperative to prevent further violence, Cameron framed this risk as a threat to national security and regional stability to showcase why the problems in Libya 'matter to all of us'.[832] Regardless of the fact that it was employed to serve multiple lines of reasoning, the reminder that historically Qaddafi's regime had displayed a behavioural pattern of an agent of large-scale violence effectively served to underscore the imminence of large-scale atrocity crimes, should the international community fail to act when needed. At the close of his statement, Cameron stressed that despite the decisive response of the international community seen in the adoption of Resolution 1970 'in record time' and the ensuing 'strong EU sanctions', 'the truth is this: Qadhafi is still on the rampage, waging war on his own people. Hundreds of thousands of people have been displaced and right now there is no sign of this ending'.[833] Thus, the prime minister's statement of 11 March illustrates that determining the need for collective action is based on an assessment of the nature and gravity of present human rights violations, a prospective assessment of the imminence of the threat of atrocities, and the host state's response to the demands of the international community. These are all familiar determinants of the 'manifest failing' and 'unable or unwilling' thresholds identified in Chapters I and III and the French discourse on collective action in Libya.

On 7 March, the UK Secretary of State for Foreign and Commonwealth Affairs, William Hague, shed light on the government's position on the 'conditions [that] should be attached to trying to implement a no-fly zone', namely: '*a demonstrable need* that the whole world can see;… *a clear legal basis* for such a no-fly zone; and… *clear support from the region* – from the middle east and north African region – as well as from the people of Libya themselves'.[834] Hague's statement marked perhaps the first public mention of these criteria in the context of the Libyan crisis, predating their articulation by the European Council on 11 March and a statement of the French president of the

[831] Ibid.

[832] Ibid.

[833] Ibid.

[834] Hansard HC Deb. vol.524 col.649, 7 March 2011. [Online]. [Accessed 12 July 2020]. Available at: https://publications.parliament.uk/pa/cm201011/cmhansrd/cm110307/debtext/110307-0001.htm#11030711000002

same day.[835] Although he did not provide further explanation of the *demonstrable need* requirement, the Secretary of State mentioned it seven times in response to questions about the threshold for military action in Libya. For instance, when quizzed about whether the potential effectiveness of a no-fly zone for resolving the 'situation on the ground' would be taken into account 'before making such an operational decision', Hague retorted that 'one of the reasons why… one of the criteria should be demonstrable need [was ensuring that implementing a no-fly zone] would actually make a difference to the situation'.[836] Despite drawing on the connection between the chances of success in resolving the situation and the need on the ground, Hague later made it clear that they are two separate considerations: 'For a no-fly zone to be implemented, it would clearly have to be effective, as well as to have the demonstrable need'.[837] Notably, he suggested that the British public might support military involvement in Libya, despite their perceived distaste for such operations following Iraq and Afghanistan, if there was a demonstrable need to protect the civilian populations there.[838] However, it was not until after the vote on the second resolution on Libya that the UK prime minister elaborated on the three requirements for the use of force discussed so far and in so doing established an explicit connection between the 'demonstrable need' threshold requirement and the aforementioned aspects of the situation in Libya invoked to justify the need for a stronger coercive response, including scenarios which involve the use of force.

Before broaching the subject of threshold in his address to the House of Commons on Resolution 1973 on 18 March, Cameron took a step back to reflect on Resolution 1970 and the context within which its adoption took place.[839] Specifically, he echoed the Secretary-General's call of 23 February urging Libya to '"meet its responsibility to

[835] The French foreign minister seemingly alluded to three requirements for the first time a day after Hague explicitly introduced them (in front of the House of Commons) as part of a clearly defined three-partite framework for determining whether the threshold for coercive action was met (on 8 March). Having argued that the 'bloody crackdown' pursued by the Qaddafi government had to be stopped (demonstrable need), the minister underscored that an intervention in Libya with others to protect civilian populations from aircraft attacks would only be conceivable under an United Nations mandate (a clear legal basis) and 'in full coordination with the Arab League and the African Union' (regional support).: Alain Juppé, 'Situation in Libya', Reply given by Alain Juppé to a question in the National Assembly, 8 March [Accessed 15 January 2020] Available at: http://www.assemblee-nationale.fr/13/cri/2010-2011/20110134.asp#P302_53875.

[836] Hansard HC Deb. vol.524 col.650, 7 March 2011.

[837] Ibid. col.653.

[838] Ibid. col.654.

[839] David Cameron, 'Statement on UN Resolution 1973', 18 March 2011 [Accessed 12 July 2020] Available at: https://www.bbc.co.uk/news/uk-politics-12786225

protect its people"', amidst reports of the 'egregious violations of international and human rights law', 'credible and consistent reports of arrests, detention and torture', and a death toll of a thousand and 'many more injured'.[840] In so doing, the prime minister not only explicitly framed the situation in Libya in R2P-terms from the start, but also highlighted concrete human rights violations that have been associated with atrocity crimes and quantitative indicators of the gravity of these abuses found to be linked to a host state's 'manifest failing' (i.e. death toll and number of injured persons). Even though Cameron did not employ the 'manifest failing' terminology, the notion of a substantive threshold being crossed was implicitly discernible in his justification for international action. Specifically, it involved highlighting aspects of the situation which together paint a picture of a 'manifestly failing' state, including the enumeration of mass-atrocity related human rights violations committed in the lead up to Resolution 1970 and drawing attention to the 'reported nature and scale of attacks on civilians' with reference to the quantitative dimensions of violence against civilians as well as its qualitative aspects, such as the means by which Qaddafi orchestrated his murderous campaign using 'the full might of [his] armed forces…, backed up by mercenaries'.[841] Cameron followed a similar line of argumentation, underscoring aspects of the situation indicative of a host state's 'manifest failing', in a statement following the adoption of Resolution 1973 and the commencement of the military intervention in Libya, where he explained once again why the UK 'pushed for' Resolution 1970 in the first place: 'Gaddafi's response to his people taking to the streets in peaceful protest was utterly brutal. He used the full might of his armed forces as well as mercenaries against them'.[842]

By contrast, when it came to justifying plans for actions involving the use of force, the UK prime minister squarely addressed the issue of threshold by remarking that 'intervening in another country's affairs should not be undertaken save in quite exceptional circumstances' and therefore 'would require [the] three tests [(demonstrable need, clear legal basis and regional support)] to be met'.[843] Crucially, Cameron offered an extensive explanation as to how the 'demonstrable need' requirement was met, thus shedding light on a range of indicators relevant to the

[840] Ibid.

[841] Ibid.

[842] Cameron, 'PM Statement to the House on Libya', 21 March 2011.

[843] Ibid.

threshold for coercive action. The first half of the prime minister's justification focused on the Libyan government's conduct following the adoption of Resolution 1970:

> [O]n demonstrable need, Gaddafi's regime has ignored the demand of UN Security Council resolution 1970 that it stop the violence against the Libyan people. His forces have attacked peaceful protesters, and are now preparing for a violent assault on a city, Benghazi, of 1 million people… They have begun air strikes in anticipation of what we expect to be a brutal attack using air, land and sea forces. Gaddafi has publicly promised that every home will be searched and that there will be no mercy and no pity shown'.[844]

Similarly to the French prime minister, Cameron pointed to the Libyan regime's disregard for the UNSC's mandate for an 'immediate end to the violence' against its civilian populations.[845] To lend support to this claim the prime minister cited the continuation of violence perpetrated by the Libyan government post Resolution 1970 and evidence of an imminent prospective large-scale military attack. The latter was reinforced by Qaddafi's threats, indicative of the government's intent to harm his people. In so doing, Cameron clearly indicated that determining whether the 'demonstrable need' threshold requirement has been met involves a composite assessment of evidence of both present and prospective large-scale violence.

The second half of his discussion of this threshold reinforces this understanding:

> 'If we want any sense of what that might mean we have only to look at what happened in Zawiyah, where tanks and heavy weaponry were used to smash through a heavily populated town with heavy loss of life. We do not have to guess what happens when he has subdued a population. Human Rights Watch has catalogued the appalling human rights abuses that are being committed in Tripoli. Now, the people of eastern Libya are faced with the same treatment. That is the demonstrable need.'[846]

---

[844] Cameron, 'Statement to the House on Libya', 21 March 2011.

[845] S/RES/1970 (2011).

[846] Hansard HC Deb. vol.525 col.611, 18 March 2011. [Online]. [Accessed 12 July 2020]. Available at: https://publications.parliament.uk/pa/cm201011/cmhansrd/cm110318/debtext/110318-0001.htm#11031850000878

The prime minister suggested that in order to assess the nature and scale of prospective attacks one need only look at the evidence of the gravity of previous attacks. In the examples provided by Cameron, the latter assessment relies on indicators such as the use of heavy weaponry and death toll, as well as a 'catalogue' of human rights violations reported by NGOs. It is hardly a coincidence that the former two aspects of the situation are also two of Gallagher's 'manifest failing' indicators, whereas the latter embodies Rosenberg and Strauss' 'determination of relevant human rights violations' (which should inform 'manifest failing' assessments).[847] In light of these parallels, the UK prime minister's explanation of what the 'demonstrable need' requirement entails adds compelling evidence to the claim made in the context of the French justificatory discourse on Libya, namely that the phrase in question is an alternative formulation of the 'manifest failing' threshold.

It is worth noting that the UK's justification for the adoption of 'all necessary measures' under Chapter VII at the vote on Resolution 1973 relied neither on the 'demonstrable need' nor the 'manifest failing' formulation, but on the Libyan governments 'loss of legitimacy' to indicate that the threshold for coercive action had been met. Nonetheless, the explanation offered by the UK representative as to why this threshold was crossed was nearly identical to the abovecited justification offered by Cameron for the existence of a 'demonstrable need':

> '[A] violent, discredited regime that has lost all legitimacy is using weapons of war against civilians. Al-Qadhafi's regime has ignored this Council's demand in resolution 1970 (2011) that it stop the violence against the Libyan people. It is now preparing for a violent assault on a city of 1 million people that has a history dating back 2,500 years. It has begun air strikes in anticipation of what we expect to be a brutal attack using air, land and sea forces. Al-Qadhafi has publicly promised no mercy and no pity. We have also seen reports today of a grotesque offer of amnesty — this, from a regime that has advertised its determination to continue persecuting and killing those Libyans who want only to take control of their own future'.[848]

[847] Rosenberg and Strauss, 'A Common Approach', p. 72.
[848] S/PV.6498, p. 4.

Again, the Libyan government's disregard for the demands by the international community comes to the fore, supported by evidence of air strikes, preparations for a large-scale military attack, and Qaddafi's threatening statements towards those who oppose him, which are also indicative of the government intent. Likewise, when Cameron revisited the demonstrable need requirement in another crucial statement to the House of 21 March, following the commencement of military action against Libya, he reasoned that the demonstrable need requirement had been met 'because Gaddafi had so flagrantly ignored the demands of two UN Security Council Resolutions to end the violence against his people'.[849] Four days after the adoption of Resolution 1973, the Libyan leader's intransigence was plain as he had not met the UNSC's demand for 'the immediate establishment of a cease fire and a complete end to violence and all attacks against, and abuses of, civilians'.[850]

The evidence of Qaddafi's lack of compliance with international demands was most clearly presented in Cameron's justification for taking military action against Libya to enforce the mandate of Resolution 1973. The P3 expanded on these demands in a joint statement issued on the day after the resolution, which elaborated the 'non-negotiable' measures that Qaddafi ought to comply with pursuant to Resolution 1973, namely to put an immediate end to all attacks against civilians, to cease his 'advance on Benghazi and withdraw [his troops] from Ajdabiya, Misurata and Zawiya', as well as to restore the supply of gas, electricity and water through the county and allow the passage of humanitarian aid to populations in need.[851] It is against this backdrop that the British prime minister provided a thorough assessment of Qaddafi's compliance with the outlined measures to make the case that the threshold for implementing Resolution 1973 through military means had been crossed:

> 'Gaddafi responded to the United Nations resolution by declaring a ceasefire, but straight away it was clear that he was breaking that promise. He continued to push his tanks towards Benghazi as quickly as possible, and to escalate his actions against Misrata. On Saturday alone, there were reports of dozens of people killed in Benghazi and dozens more in Misrata. Gaddafi lied to the

[849] Cameron, 'PM Statement to the House on Libya', 21 March 2011.

[850] S/RES/1973 (2011).

[851] France in the United States website, 'France/UK/US/message to Gaddafi', Joint message addressed to Colonel Gaddafi by France, the United Kingdom and the United States, with the support of Arab countries, 18 March 2011 [Accessed 23 September 2019] Available at: https://franceintheus.org/spip.php?article2237

> international community, he continued to brutalise his own people and he was in flagrant breach of the UN resolution, so it was necessary, legal and right that he should be stopped'.[852]

From the above paragraph, it is clear that an understanding of state conduct with a focus on the host state's response to concrete demands made by the international community and their lack of compliance with UNSC resolutions played a vital role in assessments as to whether the host state is 'manifestly failing'.

Qaddafi's lack of compliance with the demands of the international community stood out as a key determinant of the threshold for the use of force in Cameron's reaction to a proposed amendment to the motion he put forward on 21 March (United Nations Security Council Resolution 1973). Although the amendment tabled by Labour MPs Jeremy Corbyn and John McDonnell was not selected for debate, the prime minister felt the need to address it and did so with reference to the notion of threshold.[853] In response to 'the suggestion that there was somehow time for further consultation' before undertaking military action against Libya, the prime minister was categorical that the threshold for such action had been crossed because Qaddafi had ignored the ultimatum the UNSC gave him.[854] In Cameron's own words, 'The United Nations gave Gaddafi an ultimatum and he completely ignored it. To those who say that we should wait and see, I say that we have waited and we have seen more than enough'.[855] He further stressed that the use of force could have been avoided 'if Gaddafi had complied immediately and fully with the requirements of the resolution' but 'the fact… that he did not' amidst threats that he 'would show the people of Benghazi no mercy' justified the decision to use force.[856] The prime minister returned to the lack of compliance of Qaddafi's government later when he explicitly made the case that the threshold for coercive action had been crossed as follows: 'Gaddafi has had every conceivable opportunity to stop massacring his own people. The time for red lines, threats, last chances is over'.[857] In response to questions raised over the

---

[852] Cameron, 'PM Statement to the House on Libya', 21 March 2011.

[853] For the text of the amendment see: House of Commons, 'United Nations Security Council Resolution 1973', Order of Business Monday 21 March 2011, Session 2010-11, 19 March 2011 [Online] [Accessed 21 January 2021] Available at: https://publications.parliament.uk/pa/cm201011/cmagenda/ob110321.htm

[854] Hansard HC Deb. vol.525 col.705, 21 March 2011. [Online]. [Accessed 12 July 2020]. Available at: https://publications.parliament.uk/pa/cm201011/cmhansrd/cm110321/debtext/110321-0001.htm#1103219000001

[855] Ibid.

[856] Ibid.

[857] Ibid. col.714.

'reasonableness' and 'proportionality' of the intervention, Cameron added that 'urgent need to take action to stop the slaughter' was clear at the time Resolution 1973 was adopted, given that 'the population of Benghazi was under heavy attack, that civilians were being killed in significant numbers, and that an exodus from the town had begun' and reaffirmed by a final analysis of the situation on the ground following its adoption.[858]

All in all, the UK's understanding of the substantive threshold for international action in the Libyan context suggests that it encompasses an assessment of the nature, gravity and imminence of the human rights violations against targeted populations (premised on extant evidence of widespread or systematic violence as well as evidence of large-scale prospective harm), host state intent (manifested in evidence of systematic preparation and public statements of intent) and the conduct of the host state in the face of atrocities (defined by Libya's lack of compliance with the demands made by the international community and UNSC resolutions). To elaborate, when describing the recent and ongoing violations of human rights in Libya, the UK enumerated concrete human rights violations that have been associated with atrocity crimes, but also made broader references to the 'brutal repression' and 'violence' against Libyans and the 'appalling human rights violations' committed in Libya. In order to convey the seriousness of these violations, the UK government referred to relevant indicators familiar from the literature on the 'manifest failing' threshold, such as death toll, and the mass movement of populations. It conveyed the imminence of the threat of atrocities by referring to the continuation, escalation of violence, the preparations for an attack on Benghazi and the commencement of air strikes. In addition, the UK prime minister cited indicators of host state intent, in particular Qaddafi 'using military violence against his people', preparations for large-scale attacks against civilians and statements of intent to harm the Libyan populations. Since all of these factors evidence the the host state's responsibility for the commission of atrocities, they were also invoked to showcase the Libyan government's inadequate response to the demands made by the international community to put an end to the violence and thus, the demonstrable need to act in its stead.

US

Similarly to the French government, it took a matter of days for the Obama administration to voice its serious concern over '"multiple credible" reports that

[858] Ibid. col.707.

"hundreds of people" had been killed and injured in several days of unrest' and condemn 'Libya's use of lethal force against peaceful demonstrators'.[859] Three days later, on 23 February, the President issued a statement pointing to the 'suffering and bloodshed' as well as the 'threats and orders to shoot peaceful protesters and further punish the people of Libya' and expressed support for the Libyan population's universal human rights, 'includ[ing] the rights of peaceful assembly, freed speech, and [their] ability… to determine their destiny'.[860] Notably, Obama acknowledged that Libya was failing in its host state responsibility to 'refrain from violence, to allow humanitarian assistance to reach those in need, and to respect the rights of its people'.[861] Nonetheless, the President did not push for any form of collective action on the part of the international community at this stage, but merely acknowledged that the 'violence must stop' and the Libyan government be held to account for failing in its primary responsibility to protect.[862] Two days later, Obama proceeded to announce the implementation of unilateral sanctions against Libya as a measure to address the threat that the risk of instability posed to US national security and foreign policy.[863]

The US president hinted that this threshold had been met on 26 February (prior to the adoption of Resolution 1970) by 'stat[ing] that when a leader's only means of staying in power is to use mass violence against his own people, he has lost the legitimacy to rule'.[864] The same point was reiterated by Ms. Rice when she justified the US support for international sanctions against Qaddafi at the vote on Resolution 1970.[865] Her statement also echoed Obama's warning of 23 February that the Qaddafi government would be held accountable for violating the 'non-negotiable' universal rights of Libyan

---

[859] Eric Schmitt, 'U.S. "Gravely Concerned" Over Violence in Libya', 20 February 2011 [Accessed 29 March 2020] Available at: https://www.nytimes.com/2011/02/21/world/middleeast/21diplomacy.html

[860] Obama, 'Remarks by the President on Libya', 23 February 2011.

[861] Ibid.

[862] Ibid.

[863] Barack Obama, 'President Obama on Libya: "These Sanctions Therefore Target the Qaddafi Government, While Protecting the Assets that Belong to the People of Libya"', 25 February 2011 [Accessed 23 September 2019] Available at: https://obamawhitehouse.archives.gov/blog/2011/02/25/president-obama-libya-these-sanctions-therefore-target-qaddafi-government-while-prot?utm_source=wh.gov&utm_medium=shorturl&utm_campaign=shorturl; Barack Obama, 'Executive Order 13566–Libya', 25 February 2011 [Accessed 23 September 2019] Available at: https://obamawhitehouse.archives.gov/the-press-office/2011/02/25/executive-order-13566-libya

[864] The White House Office of the Press Secretary, 'Readout of President Obama's Call with Chancellor Angela Merkel of Germany', 26 February 2011.

[865] S/PV.6491, p. 3.

civilians and 'for failing to meet their most basic responsibilities to their people'.[866] What is more, the US representative went a step further by noting right from the outset that with the adoption of Resolution 1970, the international community did what it was supposed to do by speaking 'with one voice' and acting collectively under Chapter VII when 'atrocities are committed against innocents', thus signalling that it 'will not tolerate violence of any sort against the Libyan people'.[867] Whilst this is not an outright recognition of the international community's responsibility to protect, it acknowledges the sentiment underlying this commitment, namely that the international community 'must' do something together when civilian populations are being subject to mass atrocity crimes. In this context, citing the Libyan authorities' failing and 'loss of legitimacy' on account of brutalising peaceful populations expressly conveyed the idea that a substantive threshold for international action had been crossed.

Almost a week after the adoption of Resolution 1970, the US president announced that, with regards to military engagement, he had 'instructed the Department of Defense as well as [the US] State Department… to examine… a full range of [military and non-military] options', so that a decision could be made based on based on 'what's going to be best for the Libyan people in consultation with the international community'.[868] Nonetheless, there was a notable absence of discussions of the threshold for international action from US justificatory discourse for the better part of the period between the two resolutions. This can be put down to the fact that as of 14 March the US had still not taken a clear stance on Libya and seemed to be embroiled in assessing the consequences of different courses of action.[869] On 15 March, two days after France announced its intention to introduce a UN resolution authorising a no-fly zone over Libya, a White House meeting was held to discuss the issue, at which 'the President and his national security team reviewed the situation in Libya and options to increase pressure on Gaddafi' with a focus on potential collective action through the Security Council.[870]

---

[866] Ibid. pp. 3-4.

[867] Ibid. p. 3.

[868] Obama, 'Remarks by President Obama and President Calderón of Mexico at Joint Press Conference', 3 March 2011.

[869] What's in Blue article, 'Insights on Libya', 14 March 2011 [Accessed 7 February 2020] Available at: https://www.whatsinblue.org/2011/03/insights-on-libya-3.php; Michael Lewis, 'Obama's Way', 11 September 2012 [Accessed 31 May 2020] Available at: https://www.vanityfair.com/news/2012/10/michael-lewis-profile-barack-obama;

[870] The White House Office of the Press Secretary, 'Readout of President Obama's meeting with his national security team on Libya', 15 March 2011 [Accessed 7 January 2020] Available at:

Obama recalled how things stood at the time in a subsequent White House-approved interview:

> '"We knew that Qaddafi was moving on Benghazi, and that his history was such that he could carry out a threat to kill tens of thousands of people. We knew we didn't have a lot of time – somewhere between two days and two weeks. We knew they were moving faster than we originally anticipated. We knew that Europe was proposing a no-fly zone."'[871]

Recently published reflective accounts of the White House meetings late on 15 March, shed more light on the US' reservations over military action in Libya and what prompted the US to change their policy within twenty-four hours, pivoting from scepticism towards further measures against Libya to taking the lead in drafting a UNSC resolution including provisions for a no-fly zone and additional military action.[872]

The US opposition towards the Franco-British draft for a resolution imposing a no-fly zone over Libya was rooted in their strong conviction that it would not make a difference to stopping Qaddafi's offensive on Benghazi. As Obama explains in his recently published memoir, *A Promised Land*, '[s]ince Gaddafi was using ground forces almost exclusively, the only way to stop an assault on Benghazi was to target those forces directly with air strikes'.[873] Hence, he was categorical that the US was 'not going to participate in some half-assed no-fly zone that won't achieve [its intended] objective'.[874] Yet, he was also aware that unless the US 'took the lead [,] Gaddafi's troops would lay siege to Benghazi', which would result in a 'protracted conflict', in the best case scenario, or 'at worst, [in] tens of thousands or more… starved, tortured, or shot in the head'.[875] Ultimately, Obama made the decisions that the US 'had to try to prevent a massacre in Libya while minimizing the risks and burdens on an already overstretched U.S. military'.[876] His plan for action was discussed at a second meeting on 15 March and

---

https://obamawhitehouse.archives.gov/the-press-office/2011/03/15/readout-president-obamas-meeting-his-national-security-team-libya

[871] Lewis, 'Obama's Way'.

[872] For a more detailed explanation of this change of policy see Simon Chesterman, '"Leading from Behind": The Responsibility to Protect, the Obama Doctrine, and Humanitarian Intervention after Libya', 25 (3) *Ethics & International Affairs* (2011): 279-285, pp. 282-283.

[873] Barack Obama, *A Promised Land* (New York: Penguin Random House, 2020), p. 656.

[874] Ibid., p. 657.

[875] Ibid., p. 658.

[876] Ibid., p. 659.

involved, convincing the UK and France to retract their proposal for a no-fly zone, in order to allow the US 'to put an amended resolution before the Security Council, asking for a broader mandate to halt attacks by Gaddafi's forces in order to protect Libyan civilians'.[877] In addition to securing a robust UNSC civilian protection mandate, Obama premised US military action on the LAS' participation in the intervention and limiting the US role to 'launch[ing] the attacks, [to] take out Qaddafi's air defenses, and establish[ing] air superiority' and then transferring 'the bulk of the operation' to European partners, especially the UK and France, and participating Arab states.[878]

According to William Burns' account of events, chief among the factors that tipped the scales in favour of taking military action in what Obama described as a '"51-49" call' in Libya 'was the likelihood of a bloodbath, and the risks for the United States, moral as well as political, of not acting to prevent it'.[879] Similarly to the UK and France's criteria for military intervention, among the other factors were the support of Arab states, which meant that the US did not have to worry about backlash from the region and the fact that a UNSC resolution sanctioning 'intervention seemed achievable, and would give the operation the international stamp of legality and legitimacy that America's second Iraq war lacked'.[880]

Notably, Obama's memoir reveals that the substantive threshold for military intervention was on the President's mind, even if he did not explicitly cite the 'manifest failing' formulation.[881] In his book, Obama reflects on the US' 'obligation to prioritize the prevention of atrocities in its foreign policy' in the context of the Libyan crisis as follows:

> '[A]s much as I shared the impulse to save innocent people from tyrants, I was profoundly wary of ordering any kind of military action against Libya, for the same reason that I'd declined Samantha's [Power] suggestion that my Nobel Prize address include an explicit argument for a global "responsibility to protect" civilians against their own governments. Where would the obligation to intervene end? And what were the parameters? How many

[877] Ibid.

[878] Ibid.; Rice, *Tough Love*, p. 283.

[879] William J. Burns, *The Back Channel: A Memoir of American Diplomacy and Its Case for Renewal* (New York: Penguin Random House, 2020), p. 315.

[880] Ibid., pp. 316-317.

[881] Obama, *A Promised Land*, p. 654.

> people would need to have been killed, and how many more would have to be at risk, to trigger a U.S. military response?'[882]

The questions the former president raises about the parameters of the international 'responsibility to protect', as well as death toll and prospective death toll, are clearly questions concerning the matter of whether a substantive threshold for international action has been crossed. As Samantha Power recounted in her memoir:

> From the start of the meeting [on 15 March], Obama had asked us to gauge the likelihood of mass killings. He demanded that we not use vague phrases like "the town will fall" without specifying what such a development would mean for Libyans.[883]

In the aftermath of the vote on Resolution 1973 Obama offered a detailed insight into how the substantive threshold for international action was met in Libya. Much like the explanations offered by the French and British governments, the Obama administration's justification for stronger and more intrusive coercive measures was intimately linked to the actions taken by the international community up to that point, the reasons why they were necessitated, and the Libyan government's reaction to and behaviour in the aftermath of these moves. These aspects of the situation that have been shown to commonly underlie P3 arguments that a substantive threshold for international action has been crossed in Libya, were reflected concisely in Ms Rice's statement at the vote on Resolution 1973, which highlighted Qaddafi's disregard for the demands of Resolution 1970 as follows: 'Colonel Al-Qadhafi and those who still stand by him continue to grossly and systematically abuse the most fundamental human rights of Libya's people'.[884] The same determinants of a 'manifest failing' were brought to the fore and unpacked in more detail in subsequent presidential statements.

On 18 March, in his first statement on Resolution 1973, the US president first reflected on the international community's initial response to the Libyan crisis and justified the adoption of Resolution 1970 with reference to the human rights violations and brutality that Libyan populations suffered at the hands of Qaddafi: 'Innocent civilians were beaten, imprisoned, and in some cases killed. Peaceful protests were forcefully put

[882] Ibid.

[883] Samantha Power, *The Education of an Idealist* (London: HarperCollins, 2019), p. 300.

[884] S/PV.6498, p. 5.

down. Hospitals were attacked and patients disappeared. A campaign of intimidation and repression began'.[885] Obama also pointed out that, in addition to sanctions and humanitarian assistance, the international response to the mistreatment of Libyans involved an international 'cal[l] for an end to violence', in light of which he declared that '[a]mple warning was given that Qaddafi needed to stop his campaign of repression, or be held accountable'.[886] The president then underscored that, despite this clear warning, Qaddafi's subsequent actions and words unambiguously evidenced not only that he failed to meet the demands of the international community but that he deliberately acted and intended to act against them: 'Once again, Qaddafi chose to ignore the will of his people and the international community. Instead, he launched a military campaign against his own people. And there should be no doubt about his intentions, because he himself has made them clear'.[887] In support of the latter claim, Obama emphatically stressed that just the day before, Qaddafi had made fresh threats that he would have 'no mercy on his own citizens', when he spoke of the city of Benghazi with a population of approximately 700, 000.[888] In addition to highlighting evidence of host state intent, and similarly to the UK prime minister, the US president highlighted Qaddafi's historical track record of violence against his own people and others to cast suspicion on his prospective conduct. Namely, he noted that '[f]or decades [Qaddafi] has demonstrated a willingness to use brute force through his sponsorship of terrorism against the American people as well as others, and through the killings that he has carried out within his own borders'.[889]

Having thus described the early developments in the situation in Libya which led to the adoption of Resolution 1970, Qaddafi's complete disregard for international demands, and his negative prior interactions with the international community in which he acted as an agent of mass violence, Obama got to the point of 'why this matters to all of us'. According to the US president, the above aspects of the situation in Libya indicated the imminence of the threat of further atrocity crimes:

> 'Left unchecked, we have every reason to believe that Qaddafi would commit atrocities against his people. Many thousands could die. A

[885] Obama, 'Remarks by the President on the Situation in Libya', 18 March 2011.
[886] Ibid.
[887] Ibid.
[888] Ibid.
[889] Ibid.

> humanitarian crisis would ensue…The calls of the Libyan people for help would go unanswered… Moreover, the words of the international community would be rendered hollow'.[890]

Even though the president also cited the threat to regional stability and 'democratic values' the US stands for, in addition to the prospect of mass atrocities taking place in case the international community fails to act, the core justification offered for US activism in 'shap[ing] a strong international response at the United Nations' was humanitarian.[891] Obama unequivocally attested to this in the same statement by clarifying that 'the focus [of the US' work at the UN, together with allies and partners,] has been clear: protecting innocent civilians within Libya, and holding the Qaddafi regime accountable'.[892] Likewise, the decision to adopt coercive measures under Resolution 1973 'in the face of actions that undermine global peace and security' was described succinctly as a 'decision…driven by Qaddafi's refusal to respect the rights of his people, and the potential for mass murder of innocent civilians' and, moreover, a 'decision [taken] with the confidence that action is necessary' and premised on a 'cause [that] is just'. [893] On this occasion, Obama's talk of 'necessity' in conjunction with the abovementioned aspects of the situation, which he cited as giving rise to this need to act, is reminiscent of the language of 'demonstrable need' employed by the French and UK governments to indicate that a substantive threshold for international action had been crossed.

In this regard, the US president sent a clear message to Qaddafi that by acting in breach of the non-negotiable demands made in Resolution 1973 to 'immediately' stop 'all attacks against civilians,…stop his troops from advancing on Benghazi, pull them back from Ajdabiya, Misrata, and Zawiya, and establish water, electricity and gas supplies to all areas; and allow the passage of humanitarian assistance', he should be prepared to face dire consequences.[894] Namely, '[i]f Qaddafi does not comply with the resolution, the international community will impose consequences, and the resolution will be enforced through military action'.[895] As previously noted, this indirect way of

[890] Ibid.
[891] Ibid.
[892] Ibid.
[893] Ibid.
[894] Ibid.
[895] Ibid.

insinuating that certain actions would mean crossing a red line – a threshold upon which collective action is premised (in this case activating the trigger for military enforcement) – was also exhibited in Obama's justification in the wake of the adoption of Resolution 1970.

Following up on his warning of 18 March, the next day Obama announced the commencement of 'a limited military action in Libya in support of an international effort to protect Libyan civilians'.[896] His justification for authorising the involvement of US Armed Forces was as follows:

> 'Even yesterday, the international community offered Muammar Qaddafi the opportunity to pursue an immediate cease-fire, one that stopped the violence against civilians and the advances of Qaddafi's forces. But despite the hollow words of his government, he has ignored that opportunity. His attacks on his own people have continued. His forces have been on the move. And the danger faced by the people of Libya has grown'.[897]

By underscoring the aspects of the situation which indicated that the Libyan government had ignored the demands of Resolution 1973, i.e. the continuation of violence against civilian populations and the advancement of Qaddafi's military in preparation for further attacks, Obama made a convincing case that the conduct of Libyan authorities put civilian populations at an increased and imminent risk of mass violence. These factors of the situation in Libya, which very much align with the determinants of 'demonstrable need' put forward by France and the UK, formed the backbone of the US president's argument that the threshold for the enforcement of the 'all necessary measures' mandate of Resolution 1973 had been passed. Relatedly, he went on to reiterate that even though 'the use of force is not our first choice… we cannot stand idly by when a tyrant tells his people that there will be no mercy, and his forces step up their assaults on cities like Benghazi and Misurata, where innocent men and women face brutality and death at the hands of their own government'.[898] Significantly, the above justifications for the use of force were coupled with an unambiguous reference to the international community's responsibility to protect, as Obama stated that he was 'proud

[896] Barack Obama, 'Remarks by the President on Libya', 19 March 2011 [Accessed 23 September 2019] Available at: https://obamawhitehouse.archives.gov/the-press-office/2011/03/19/remarks-president-libya

[897] Ibid.

[898] Ibid.

that [the US is] acting as part of a coalition that includes close allies and partners who are prepared to meet their responsibility to protect the people of Libya and uphold the mandate of the international community'. [899] This lends support to the notion that the aspects of the situation enumerated by Obama in defence of coercive action in Libya are precisely the factors he deemed relevant to determining whether the 'manifest failing' threshold for international action under the R2P had been met.

The US president made yet another implicit suggestion that this threshold had been crossed two days hence, in a letter updating the US Congress on military actions in Libya. Reflecting on the commencement of the military operations 'to prevent a humanitarian catastrophe and address the threat posed to international peace and security by the crisis in Libya', Obama once again reiterated the specific demands made in Resolution 1973 in a way that was almost identical to his statement of 18 March.[900] Specifically, international action in line with the mandate of Resolution 1973 was justified on the basis of a clearly outlined juxtaposition between the demands made in the Resolution and the subsequent actions of the Qaddafi government as follows:

> 'Although Qadhafi's Foreign Minister announced an immediate cease-fire, Qadhafi and his forces made no attempt to implement such a cease-fire, and instead continued attacks on Misrata and advanced on Benghazi. Qadhafi's continued attacks and threats against civilians and civilian populated areas…, as expressly stated in U.N. Security Council Resolution 1973, constitute a threat to the region and to international peace and security. His illegitimate use of force not only is causing the deaths of substantial numbers of civilians among his own people, but also is forcing many others to flee to neighboring countries... *Qadhafi has forfeited his responsibility to protect* his own citizens and created a serious need for immediate humanitarian assistance and protection, with any delay only putting more civilians at risk' (emphasis added).[901]

This portrayal of Qaddafi's regime as one that has 'forfeited' its primary host state responsibility and thus 'created a serious need for immediate…protection' to avoid

[899] Ibid.

[900] Obama, 'Letter from the President regarding the commencement of operations in Libya', 21 March 2011.

[901] Ibid.

risking the lives of more civilians is an apt way of describing a 'manifestly failing' host state. The argument that this threshold had been met was once again supported by familiar justifications pointing to aspects of the situation indicative of a 'manifest failing', namely Qaddafi's disregard for the demands made by the international community, the continuation of attacks and threats against civilian populations, a substantial death toll and number of displaced persons. Judging by his statements of 19 and 21 March, it is clear that Obama did not hesitate to justify the use of force with reference to both Libya's host state responsibility to protect and the collective responsibility of the international community to act in response to atrocity crimes abroad.

Later statements in March endeavoured to marry the language of responsibility with that of 'national interest'. This newly-added feature of the US discourse on international action against Libya is best exemplified in the following excerpt of Obama's weekly address to the nation on 26 March:

> 'Even when the US governments' justifications for acting when innocent populations are being brutalized; when someone like Qaddafi threatens a bloodbath that could destabilize an entire region; and when the international community is prepared to come together to save many thousands of lives – then it's in our national interest to act. And it's our responsibility'.[902]

Tom Keating offers one possible explanation for this post-factum introduction of 'national interest' which marks a departure from the train of thought that Obama's previous statements followed:

> 'Commentaries in the American press that emphasized the influence of values on US policy and the lack of interests in Libya eventually touched a nerve as, in a public address, US president Barack Obama offered a defence of American action on the basis of national interests'.[903]

Whatever necessitated this change in rhetoric, the professed aspects of the situation that led to authorising the Libyan intervention remained the same, namely the threats issued by Qaddafi, specifically comparing 'his own people… to rats, and threaten[ining] to go

[902] Obama, 'Weekly Presidential Address: The Military Mission in Libya', 26 March 2011.
[903] Keating, 'The UN Security Council on Libya', p. 177.

door to door to inflict punishment'; his violet conduct in the past manifested in 'hang[ing] civilians in the street, and kill[ing] over a thousand people in a single day'; and, on observing Qaddafi's forces on the outskirts of Benghazi, the imminence of a 'massacre' that could have taken place within a day (in case of inaction) and 'would have reverberated across the region and stained the conscience of the world'.[904]

What is more, while the US president initially declared that 'given the costs and risks of intervention, [the US] must always measure [its] interests against the need for action', he promptly clarified that this 'cannot be an argument for never acting on behalf of what's right'.[905] '[F]aced with the prospect of a violence on a horrific scale', the president argued, the United States could not 'turn a blind eye to atrocities in other countries' because '[t]o brush aside America's responsibility as a leader and – more profoundly – our responsibilities to our fellow human beings under such circumstances would have been a betrayal of who we are'.[906] Having thus linked the notion of responsibility to protect strangers to national values, Obama added that the US also had 'an important strategic interest in preventing Qaddafi from those who oppose him [as a] massacre would have driven thousands of additional refugees across Libya's borders, putting enormous strains on the peaceful – yet fragile – transitions in Egypt and Tunisia'.[907] Ann-Marie Slaughter later made an effort to establish a connection between interest-based and value-based justification for the Libyan intervention by suggesting that 'interests are not far from the values that are often articulated in their place' and that US interests 'in the region rested on their ability to provide support for those societal interests leading political change within states rather than for the oppressive governments seeking to contain them'.[908]

That such justifications were first introduced *ex post facto*, nine days after Obama's initial justification of Resolution 1973, indicates that they were most likely added in response to external pressures, as Keating appears to suggest. Whatever the reason for introducing more or less elaborate accounts of how values and interests intersect, it does not alter the fact that the notion of responsibility to protect consistently featured in US justifications for action in Libya and that the underlying rationale for this response had

[904] Barack Obama, 'Remarks by the President in Address to the Nation on Libya', 28 March 2011.
[905] Ibid.
[906] Ibid.
[907] Ibid.
[908] Keating, 'The UN Security Council on Libya', p. 177.

not changed. Namely, that when 'challenges that threaten our common humanity and our common security [such as] preventing genocide or keeping the peace' arise, America's 'task is… to mobilize the international community for collective action'.[909]

Although the US President did not specify a set of concrete requirements or tests for such action in the systematised manner that the heads of state and government of France and the UK did, he conveyed the idea that a threshold had been passed indirectly by pinpointing concrete aspects of the situation in Libya. In so doing, the factors put forward in his justification to intervene in Libya aligned with those professed by French and British government leaders. An apt illustration of this can be found in Obama's address to the nation of 28 March, the relevant excerpt of which is worth reproducing in full:

> 'In the face of the world's condemnation, Qaddafi chose to escalate his attacks, launching a military campaign against the Libyan people. Innocent people were targeted for killing. Hospitals and ambulances were attacked. Journalists were arrested, sexually assaulted, and killed. Supplies of food and fuel were choked off. Water for hundreds of thousands of people in Misurata was shut off. Cities and towns were shelled, mosques were destroyed, and apartment buildings reduced to rubble. Military jets and helicopter gunships were unleashed upon people who had no means to defend themselves against assaults from the air. Confronted by this brutal repression and a looming humanitarian crisis, I ordered warships into the Mediterranean'.[910]

The aspects of the situation highlighted in the above paragraph are consistent with some of the key determinants and indicators evidencing the nature, gravity and imminence of the threat that civilian populations in Libya were faced with that the UK and France tended to profess to make the argument that the threshold for coercive action had been crossed. Prominent amongst them are the enumeration of the concrete human rights violations taking place and the means by which they were inflicted, the continuation or escalation of violence, threatening statements made by Qaddafi, and the Libyan authorities' disregard for the demands of the international community. As Gallagher points out in relation to the above paragraph, although Obama did not cite the phrase

[909] Obama, 'Remarks by the President in Address to the Nation on Libya', 28 March 2011.
[910] Ibid.

'manifest failing', his argument for authorising military action in Libya was clearly premised on 'the escalation in violence in both qualitative and quantitative terms [which] signifies that a line has been crossed – a threshold passed – thus implying that the Libyan regime had "manifestly failed"'.[911] Having already highlighted the continuities between the arguments, determinants and indicators professed by the US, the UK and France with regards to the substantive threshold for coercive action in response to atrocities, the remainder of this chapter will focus on the justifications offered by the rest of the UN Security Council as to whether this threshold was met in the context of Resolutions 1970 and 1973.

SECURITY COUNCIL

Meeting 6491

UN Security Council meeting 6491, which took place on 26 February at the vote on Resolution 1970, was the first publicly recorded meeting of the Council on Libya which saw contributions from (all fifteen) member states. As detailed earlier in the chapter, the meeting was marked by notable references to the threshold for international action by the US and France. To recapitulate, whilst the US spoke of the failing of Libyan leadership and its loss of legitimacy, France invoked the protection responsibilities individual states have towards their population and the international community's responsibility 'to intervene when States fail in their duty'.[912] By contrast, the rest of the Security Council did not follow suit in employing the 'manifest failing' formulation or closely related concepts, nor did they adopt alternative terminologies to refer to the threshold for collective action. Nevertheless, the underlying argument that a substantive threshold had been crossed was discernible in the commentary of member states.

All states condemned the violence against the civilian populations in Libya and in so doing often reflected on the nature, gravity of the human rights violations taking place and the role of the state. India voiced its strong disapproval of the use of force against innocent civilians, whereas others branded it as 'disproportionate' (Nigeria), and 'indiscriminate and excessive' (South Africa).[913] Some highlighted the gravity of the situation and the role of the state in qualitative terms. Lebanon underscored the League

[911] Gallagher, 'Syria and the indicators of a "manifest failing"', p. 5.
[912] S/PV.6491, pp. 3-5.
[913] Ibid. p.3.

of Arab State's 'profound condemnation' of the violence against Libya's civilian population, 'especially the use of foreign mercenaries, live ammunition, heavy weapons and other methods against the demonstrators, all of which are grave violations of human rights and international humanitarian law'.[914] Germany stated that the message Resolution 1970 sent was that 'the international community will not tolerate the gross and systematic violation of human rights by the Libyan regime'.[915] Others, underscored the seriousness of the violence with a focus on its quantitative dimension. Notably, the general idea of a quantitative threshold being crossed is exemplified in the Brazilian position that '[t]he level of violence against the civilian population is totally unacceptable', and a similar statement by Bosnia and Herzegovina citing 'the unacceptable level of violence against Libyan civilians'.[916] Taking a slightly different approach to conveying the passing of a quantitative threshold, India drew specific attention to two oft-cited indicators of the seriousness of the threat to civilian populations, namely 'the loss of numerous lives and injuries to many more'.[917] In a similar manner, South Africa spoke of the 'untold atrocities and countless losses of civilian lives'.[918] Likewise, Russia noted 'the many lives lost among the civilian population' and condemned the 'unacceptable' use of force and violence against peaceful demonstrators, whilst Lebanon pointed to 'the hundreds killed and thousands injured among the Arab Libyan people'.[919] Whereas Nigeria spoke more broadly of 'the inflammatory rhetoric and the deplorable loss of life we are witnessing in Libya' (respectively indicative of host state intent and gravity), Bosnia and Hezegovina cited a concrete number 'the deaths of at least a thousand people'.[920] These statements, especially the latter, echo the death toll statistics 'of more than 1,000' killed and reports of 'high casualties' in clashes on 25 February, cited by the Secretary-General in his briefing to the Council of the day before.

Other noteworthy aspects of the situation in Libya that featured in member states' justifications for their vote in support of coercive measures were the escalations of violence and the social dislocation of Libyan populations. South Africa expressed its

[914] Ibid. p.4.
[915] Ibid. p.6.
[916] Ibid. p.6
[917] Ibid. p.2
[918] Ibid. p.3
[919] Ibid. p.4
[920] Ibid., pp.3, 6.

'concer[n] at the deteriorating situation in Libya' in unison with Nigeria, Colombia and Bosnia and Herzegovina who made generic references to the escalation of violence. Portugal, on the other hand, voiced its deep concern with the worsening refugee crisis.[921] Bosnia and Herzegovina also highlighted the urgency of the humanitarian crisis unfolding in Libya, expressing its concern with 'information about the flow of refugees fleeing the violence in their country and about a considerable number of internally displaced persons', food scarcity and the shortage of medical supplies.[922] As discussed in Chapter I the mass movement of populations has been singled out as an indicator of the gravity of the situation, whereas the escalation of violence / deteriorating refugee crisis have been identified as indicators of the imminence of the threat of atrocities in Chapter III.

It is worth noting that in addition to the substantive threshold for collective action, the threshold requirement of regional support, explicitly singled out as one of three prerequisites for the use of force by the P3, featured in the justifications of a third of UNSC member states at the vote on Resolution 1970 – the P2, South Africa, Nigeria, Lebanon and Brazil.[923] Notably, Lebanon stated that it was 'when the Libyan authorities did not respond to that call [of the Council of the League of Arab States to halt the violence against civilians that] Lebanon decided' to support Resolution 1970.[924] Unlike the other states who brought up the support of regional organisation to highlight its significance as a prerequisite for forcible action, Lebanon's statement reveals that assessing the Libyan state's conduct in light of the demands made by regional organisation was a key determinant that informed their decision to support collective action against Libya. In this sense, the failing of the Libyan authorities to comply with regional calls to put an end to the violence against civilian populations can be seen as a clear behaviour-based indicator that the host state is 'manifestly failing' in its responsibility to protect its populations. This is consistent with the emphasis placed by the P3 on Libya's lack of compliance with international demands and specifically those made by the UN Security Council when making the case that Libya had crossed a red line. It also provides support for the inclusion of the determinant of 'the host state's

[921] Ibid. p.5
[922] Ibid. p.6
[923] Ibid., pp. 3, 4, 6.
[924] Ibid. p.4.

response to the demands made by the international community' (identified in Chapter III), in the framework outlined in Chapter V.

Overall, Security Council member states echoed key aspects of the situation in Libya relevant to the substantive threshold for international action that featured in the Secretary-General's briefing of the previous day, cited earlier in this chapter (pp. 182-184).[925] Member states generally condemned the human rights violations taking place in Libya, although most of them did not enumerate specific acts of violence in the manner in which the Secretary-General did in his briefing. This is understandable since said briefing had already informed the Council of the reported human rights abuses the Libyan regime had allegedly inflicted upon its populations. Similarly to the Secretary-General's report, Security Council member states highlighted that the situation had surpassed a gravity threshold in qualitative and quantitative terms by making various references to death toll, the number of displaced people, weapons used, threatening statements by the host state, and the unacceptable and disproportionate use of force. The seriousness of the situation was also implicit in expressions of concern with the escalation of violence and the deterioration of the situation in Libya, indicative of an imminent threat of atrocities.

It is true that the P2 and E10 did not make direct references to Libya's 'manifest failing', as well as the international community's subsidiary responsibility to protect populations, even though member states were explicitly reminded of both by the Secretary-General the day before the vote on Resolution 1970. They also steered clear from invoking the R2P altogether, with the exception of Colombia's reference to the primary protection responsibilities of the Libyan authorities.[926] However, the above aspects of the situation in Libya underscored by the P2 and E10 evidence that the underlying argument that a substantive threshold had been crossed was present in the justifications of states in favour of a strong response to the Libyan crisis. With the exception of China, all states made references to aspects of the situation associated with the notion of a 'manifestly failing' host state based on the qualitative and quantitative dimensions of the violence against the Libyan populations, be it by referring to the type, gravity or imminence of the human rights violations taking place in Libya or the regime's lack of compliance with the demands of regional organisations and evidence of host state intent. Thus, the relevant aspects of the situation professed by the P2 and E10 aligned with those invoked by the P3

[925] See S/PV.6490 (2011).
[926] Ibid., p. 5.

and the Secretary-General and equally conveyed the idea that the situation in Libya warranted decisive action on account of the fact that a substantive threshold for international intervention had been met (in addition to other prerequisites, such as regional support).

Meeting 6498

This agreement amongst UNSC members on which substantive aspects of the situation in Libya were relevant to determining the need for international action was sustained in the context of deliberations at the vote on Resolution 1973, even though some member states were not entirely convinced that the threshold for the use of force had been met (as discussed below). Considering the developments in Libya in the three weeks since the vote on Resolution 1970, India,[927] Brazil,[928] Bosnia and Herzegovina,[929] South Africa[930] and China[931] expressed their grave concern with the continuing deterioration of the situation in Libya, whereas Portugal noted the 'escalation in violence' contrary to the demands of Resolution 1970.[932] These statements attested to the imminence of further atrocities, which has repeatedly surfaced in this chapter as one of the key determinants of a 'manifest failing'.

Similarly to the meeting at the vote on Resolution 1970, aspects of the situation indicative of the gravity of the human rights violations taking place in Libya were prominent in some states' explanations of their vote on additional coercive measures against Qaddafi's regime. For instance, India expressed its concern about 'the loss of numerous lives and injuries to many more' and 'the welfare of civilian populations', characterising the use of force against them as 'unacceptable'.[933] Lebanon also underscored the gravity of the situation in quantitative and qualitative terms with reference to 'acts of violence and the use of heavy weapons and aircraft against large swaths of the civilian population' and '[h]undreds of innocent victims… and… the displacement of hundreds of thousands of Libyan citizens'.[934] Nigeria communicated the idea that a gravity threshold had been

[927] S/PV.6498 (2011), p. 6.
[928] Ibid.
[929] Ibid., p.7
[930] Ibid., p.9
[931] Ibid., p. 10
[932] Ibid., p. 8
[933] Ibid., p. 5.
[934] Ibid., pp. 3-5.

crossed quite literally by stating that the adoption of Resolution 1973 'was necessitated by the persistently grave and dire situation in Libya' and that the 'magnitude of this humanitarian disaster is indeed what compelled Nigeria to vote in favour of resolution 1973'.[935] In spite of abstaining in the vote, Germany also underscored the seriousness of the situation, expressing 'particula[r] concer[n] about the plight of the Libyan people and the widespread and systematic attacks they are suffering'.[936]

In addition to the above determinants of a 'manifest failing', i.e. the imminence of mass atrocities and the gravity of human rights violations, the Council placed emphasis on the failure of the Libyan authorities to heed the warnings of the international community and comply with its demands. This is aptly exemplified in Nigeria's statement to the Council, which reads: 'Notwithstanding the clear expression of our common will and the comprehensive measures instituted under resolution 1970 (2011), the Libyan authorities have continued to violate the terms of the resolution and fundamental principles of international law'.[937] Likewise, Lebanon remarked that 'Colonel Al-Qadhafi's regime disregarded the demands and yearnings of his people as well as international resolutions'.[938] The same sentiment was expressed in a forceful statement by Colombia: 'The Government of Colombia deeply deplores the fact that the provisions of that resolution have been systematically violated and that our calls have gone unheeded'.[939] That Colombia saw this as an indicator of Libya's 'manifest failing' is plain, since the above comment was made in the context of a remark that the Libyan government had shown it was not fulfilling its primary protection responsibilities 'through its actions and statements'.[940] Thus, Colombia made it clear that the host state's demonstrable lack of compliance with international resolutions was pivotal to its determination that the 'manifest failing' threshold was crossed in Libya, which informed its vote in favour of Resolution 1973.

Of particular interest in this regard is also the statement made by South Africa, whose decision to align with the P3 in support of Resolution 1973 provided the decisive vote in favour of the use of force in Libya. As Nissen and Pouliot explain, in the absence of a

[935] Ibid., p. 9
[936] Ibid., p. 5.
[937] Ibid., p. 9
[938] Ibid., p. 3
[939] Ibid., p. 7.
[940] Ibid.

united AU stance, South Africa 'found itself in a very uncomfortable position, pushing for more diplomacy while being pressurized to help stop the killing of civilians'.[941] Given that South Africa was considered to be the swing state in the adoption of Resolution 1973, its favourable vote was met with censure not only by African states for conspiring with the West, but also by a number of UN member states for being 'plainly incompetent' and a result of South Africa being 'fooled' by the P3.[942] South Africa's wish to 'restore its diplomatic authority by denouncing the military intervention in Libya' is one plausible explanation for the stark shift in the country's stance from support for the intervention to strong criticism of the NATO campaign.[943] Nonetheless, it is worth exploring how it addressed the issue of threshold in its statements at the vote on Resolution 1973.

South Africa framed Resolution 1970 as an encouragement for Libya to 'act responsibly and stop committing more acts of violence against their own people', namely to fulfil its protection responsibilities towards its population.[944] By then highlighting Libya's defiance against this resolution, manifested in the continuation of its campaign of 'kill[ing] and displac[ing] numerous civilians while continuing to violate their human rights', South Africa made a convincing case that the Libyan authorities were failing in their primary 'responsibility to protect'.[945] In other words, South Africa underscored the gravity of human rights violations and the Libyan government's lack of compliance with UNSC resolutions in its justification that the 'manifest failing' threshold had been met.

So far, the analysis of UNSC member states' explanations for their vote on Resolution 1973 has shed light on the determinants and indicators associated with the 'manifest failing' threshold that states professed to make the case that this substantive threshold requirement for international action had been met. The rest of this section turns specifically to the position of states who abstained in the vote, in order to call attention to the reasons why the use of force remained a contentious matter despite the apparent consensus over the need to act. Although the grounds for these abstentions have been

[941] Rebecca Adler-Nissen and Vincent Pouliot, 'Power in Practice: Negotiating the International Intervention in Libya', 20 (4) *European Journal of International Relations* (2014): 889-911, p. 904.

[942] Ibid.

[943] Ibid. For an alternative analysis of South Africa's approach to the Libyan crisis, see Alexander Beresford, 'A responsibility to protect Africa from the West? South Africa and the NATO Intervention in Libya', 52 (3) *International Politics* (2015): 288-304.

[944] S/PV.6498, p. 10.

[945] Ibid.

widely discussed in the literature on Libya,[946] the justifications offered by the five states in question are worth revisiting with a view to the threshold for the use of force, so as to determine whether these objections reflect a disagreement about the 'manifest failing' of the Libyan state or were a product of an entirely different set of concerns.

Amongst those who abstained, China expressed the most unwavering rejection of the use of force and engaged the least with the widely-acknowledged grounds for this extreme form of coercive action – the need to protect civilians against an imminent threat of atrocity crimes. Despite professing to be 'gravely concerned by the continuing deterioration of the situation in Libya', the Chinese representative made it clear that his government did not support the provisions of Resolution 1973 pertaining to the use of force on account of its historical principled objection to non-consensual military intervention, which he reaffirmed unequivocally with the statement 'China is always against the use of force in international relations'.[947] He then underscored that China had 'serious difficulty with parts of the resolution' as many questions concerning the use of force posed during consultations preceding its adoption remained unanswered.[948] Having voiced its misgivings about military intervention, China attributed its decision not to prevent the adoption of the resolution to 'attach[ing] great importance to the relevant position by the 22-member Arab League on the establishment of a no-fly zone over Libya and to the position of African countries and the African Union', and to 'the special circumstances surrounding the situation in Libya'.[949] The latter, according to Garwood-Gowers, reflected a premeditated strategy of the Chinese government to underscore 'the unique nature of the crisis in order to avoid the situation of Libya being viewed as a precedent for future civilian protection operations'.[950]

In-depth reviews of China's approach to humanitarian intervention and the R2P also suggest that China's abstention on Libya was not indicative of a substantive change in its position on the extraterritorial use of force for human protection in the absence of host state consent. In his analysis of China's stance on humanitarian interventions since the end of the Cold War, Jonathan Davis singles out UNSC authorisation and host state

[946] See for example Keating, 'The UN Security Council on Libya', pp. 179-185.

[947] S/PV.6498, p. 10.

[948] Ibid.

[949] Ibid.

[950] Andrew Garwood-Gowers, 'China and the "responsibility to protect": The implications of the Libyan intervention'. 2 (2) *Asian Journal of International Law* (2012): 375-393, p. 387.

consent as 'China's two baselines for humanitarian intervention'.[951] Despite 'sings of greater flexibility in Beijing's position – including its willingness to consider situations of humanitarian need as threats to international peace and security and to use is growing global influence to secure target state consent', its stance on the above two prerequisites has remained consistent.[952] Likewise, Sarah Teitt observes that since the 2005 agreement on the R2P, China has 'preserved the vestiges of its once firm stance on non-interference – the requirement for host state consent for collective military deployment – without appearing to completely turn a blind eye to mass atrocities'.[953] Put simply, 'China's position is best understood as rejecting claims to a unilateral right of humanitarian intervention while, in practice, acquiescing in the exercise of a multilateral right through the Security Council'.[954] Given China's continued resistance to associating the R2P with coercive action,[955] the fact that it did not object to the Libyan intervention in practice does not mean that post-2005 China has come to recognise that the requirement of the host state's 'manifest failing' has replaced the high bar of host state consent as the threshold for the use of force. Hence, it is not surprising that China did not engage with the notion of a substantive threshold for coercive action for human protection. Such behaviour is consistent with Davis' observation that 'there are signs that China is reticent to further engage with the concept of the R2P, and Beijing continues to resist efforts to flesh out clear criteria for when and how the Security Council should act if individual states fail to live up to their responsibilities'.[956]

In light of the above observation, China's policy on Libya rather evidences another facet of its approach to the R2P, which involves circumventing open opposition to:

> 'what it perceives to be Western normative order underpinning the R2P [by] leveraging its relationship with like-minded states to limit the prospect of R2P directly undermining its resistance to non-consensual intervention, or to

[951] Jonathan E. Davis, 'From Ideology to Pragmatism: China's Position on Humanitarian Intervention in the Post-Cold War Era', in Brendan Howe and Boris Kondoch (eds.) *The Legality and Legitimacy of the Use of Force in Northeast Asia* (Leiden: Brill, 2013): 117-190, pp.119, 179-181.

[952] Ibid. p. 119.

[953] Sarah Teitt, 'The Responsibility to Protect and China's Peacekeeping Policy', 18 (3) *International Peacekeeping* (2011): 298-312, p. 304.

[954] Davis, 'From Ideology to Pragmatism', p. 179.

[955] Teitt, 'The Responsibility to Protect and China's Peacekeeping Policy', p. 304.; Garwood-Gowers, 'China and the "responsibility to protect"', p. 381.

[956] Davis, 'From Ideology to Pragmatism', p. 182.

> utterly discredit its commitment to enhancing civilian protection through political negotiations rather than enforcement measures'.[957]

Accordingly, the fact that the majority of states recognised the relevance of the substantive threshold for collective action and that it had been crossed in Libya left China with little choice but to allow Resolution 1973 to pass and not to directly challenge the arguments made by others in this regard.

Russia voiced objections similar to those raised by China. The Russian Federation abstained in the vote due to 'considerations of principle' such as the inclusion of provisions in the text of Resolution 1973 that 'could potentially open the door to large-scale military intervention'.[958] It also expressed regret that 'methods involving force' had prevailed over its own draft resolution submitted to the Security Council on 16 March, which called for 'a peaceful settlement of the situation in Libya'.[959] That this proposal did not find sufficient support is hardly surprising not only because of the traction the Franco-British draft had gained, but also 'because it seemed dangerously out of step with the rapidly evolving situation on the ground'.[960] Discussions on the draft later adopted as Resolution 1973 did not assuage Russia's concern with the excessive use of force, since the questions they posed as to 'how the no-fly zone would be enforced, what the rules of engagement would be and what limits on the use of force there would be' remained unanswered.[961] While Russia gave serious consideration to the Arab League's request for the establishment of a no-fly zone as a measure for civilian protection, it felt that the provisions introduced in the draft of Resolution 1973 'transdend[ed] the initial concept as stated by the League of Arab States' because they made an extensive military intervention possible.[962] Nonetheless, Russia explained its decision not to vote against the resolution as being guided by the 'basic principle [of the protection of civilian populations] as well as by the common humanitarian values' it shared with the rest of the Council.[963] This is also evidenced in Medvedev's statement to the press regarding the

[957] Teitt, 'The Responsibility to Protect and China's Peacekeeping Policy', p. 299.
[958] S/PV.6498, p. 8.
[959] Ibid.
[960] Bellamy and Williams, 'The new politics of protection?', p. 844.
[961] S/PV.6498, p. 8.
[962] Ibid.
[963] Ibid.

Russian vote on Resolution 1973, in the aftermath of prime minister Vladimir Putin's public criticism:

> 'the consequences of this decision [to abstain] were obvious. It would be wrong for us to start flapping about now and say that we didn't know what we were doing . . . everything that is happening in Libya is a result of the Libyan leadership's absolutely intolerable behavior and the crimes that they have committed against their own people'. [964]

Keating's analysis of Russia's policy on Libya suggests that it reflected a difference in approach within Russian leadership between prime minister Putin, who 'ma[de] a principled defence of non-intervention alongside a harsh criticism of the US's [(sic.)] tendency to resort to the early use of force and [president Medvedev, who] defend[ed] the abstention on the grounds of protecting civilians and working alongside the Arab League', with the hope this stance would enhance his chances of re-election.[965] Power, however, excludes the possibility that Medvedev did not coordinate with Putin prior to instructing the Russian representative not to veto, highlighting that 'the world was broadly united in taking military action to protect Libyan civilians'.[966] Regardless of what informed the Russian position on Libya, it is clear that it was defined by a tension between the acknowledgement of the humanitarian imperative to act to put an end to what it believed to be 'unacceptable' violence against the Libyan population and misgivings about the use of force. The same conflict is reflected in the positions of the other three states who decided to abstain, as well as in Lebanon's justification for its vote in favour of Resolution 1973.

The statement offered by the Brazilian representative is an apt illustration of the antagonism between the need for human protection owing to the 'manifest failing' of the host state and the risks associated with the use of force. On the one hand, Brazil 'condemn[ed] the Libyan authorities' disrespect for their obligations under international law and human rights law' and professed its sympathy for the Arab League's call for 'strong measures to stop the violence through a no-fly zone'.[967] On the other hand, it decided to abstain in the vote because Resolution 1973 'contemplat[ed] measures that go

[964] Power, *The Education of an Idealist*, p. 304.
[965] Keating, 'The UN Security Council on Libya', p. 179.
[966] Power, *The Education of an Idealist*, p. 304.
[967] S/PV.6498, p. 6.

far beyond' the Arab League's appeal to the UN Security Council to establish a no-fly zone over Libya.[968] Specifically, Brazil was concerned that the use of force as specified in Paragraph 4 of the resolution would not succeed in 'realizing [the] common objective' of putting an end to the violence and ensuring civilian protection and that it might 'caus[e] more harm than good'.[969] Still, despite retaining reservations about the use of force, Brazil dispelled any ambiguity that it did not recognise the humanitarian imperative to act by stating: 'Our vote today should in no way be interpreted as condoning the behaviour of the Libyan authorities or as disregard for the need to protect civilians and respect their rights'.[970] Brazil's condemnation of Qaddafi's government and the human rights violations it committed against its own populations leaves little doubt that it deemed that the substantive threshold for collective action was met, despite being somewhat dissuaded by prudential considerations concerning the consequences of the use of force, its chances of success and the scope of the measures that the text of Resolution 1973 authorised.

Respectively, India put down its abstention in the vote to the fact that the Resolution 1973 sanctioned 'far-reaching measures under Chapter VII of the United Nations Charter, with relatively little credible information on the situation on the ground in Libya'.[971] While India reaffirmed that it remained 'gravely concerned about the deteriorating humanitarian situation in Libya', the lack of 'clarity about details of enforcement measures' and 'an objective analysis on the ground' (owing to the fact that the anticipated report of the UN Special Envoy was not delivered before the decision on the use of force was made) ultimately led to its abstention.[972] The fact that India cited aspects of the situation indicating that the 'manifest failing' threshold had been crossed (as illustrated above) suggests that its position on the Libyan crisis was ostensibly informed by a substantive threshold for collective action. Its hesitation as to whether this threshold had been met in relation to military intervention stems from finding the available evidence / information inadequate in order make a sufficiently well-informed determination with regards to the use of force.

Likewise, Germany's abstention was not a product of a principled objection to the relevance of the substantive threshold requirement for international action. On the

[968] Ibid.
[969] Ibid.
[970] Ibid.
[971] Ibid.
[972] Ibid.

contrary, Germany not only stressed the gravity of the human rights violations taking place in Libya but clearly indicated that a threshold for stronger coercive action had been met, voicing its 'full support [for] the package of economic and financial sanctions in the resolution just adopted'.[973] It nonetheless rejected the use of military force on account of the risks involved, namely '[t]he likelihood of large-scale loss of life' and if it 'turns out to be ineffective… the danger of being drawn into a protracted military conflict that would affect the wider region'.[974] Consequently, its sceptical stance on international action in Libya appeared to be informed by prudential and pragmatic considerations concerning the use of force, as opposed to an understanding that the substantive threshold for such action had not been crossed or was not relevant to its position on Libya. Adler-Nissen and Pouliot's interviews with government officials with a 'central position in the multilateral negotiations over Libya' suggest that 'it appears that the German UN mission did not fully grasp the course of negotiations and, in particular, the American change of mind that occurred late on 15 March'.[975] Specifically, in assuming 'that the US would not go for the war', they felt secure in their abstention, which came as a shock to diplomats in the US and across the world, owing to its unprecedented divergence with the position of Germany's customary partners in multilateral diplomacy.[976]

Much like the five states who abstained, Lebanon was conflicted about its vote on Resolution 1973, but decided to support the resolution in spite of its doubts about the use of force. Namely, in addition to 'reaffirm[ing] the importance of and need for full respect for the sovereignty and territorial integrity of Libya', Lebanon voiced its strong principled objection to the use of force as follows: 'It is quite clear that Lebanon, which has itself experienced the atrocities of war and violence, would never advocate the use of force or support war in any part of the world – especially not in the brotherly country of Libya'.[977] Consequently, it expressed its 'hopes that the resolution adopted today will have a deterring effect, ensure that Libyan authorities move away from using all forms of violence against their own people, and avert the use of force'.[978] Despite its effort to maintain their principled opposition to the use of force while voting to authorise such

[973] Ibid., p.5.
[974] Ibid.
[975] Adler-Nissen and Pouliot, 'Power in Practice', pp. 897, 903.
[976] Ibid., pp. 902-903.
[977] S/PV.6498, pp. 3-4.
[978] Ibid. p.4.

international action against Libya, it would be unreasonable to suggest that Lebanon was unaware of the possibility that the resolution might not have the desired deterring effect on the Libyan authorities that so many other warnings by regional organisation and the international community had failed to produce. As Simon Adams reasons, '[t]hose Security Council members who voted for Resolution 1973 understood that they were voting for air strikes to protect civilians', especially since military intervention was decided upon 'after other attempts at dissuasion had failed'.[979] Despite this somewhat unconvincing justification for the use of force, premised on hopes that it would not be necessary to resort to it, the rhetorical emphasis in Lebanon's explanation of its vote was on civilian protection. This was manifested in its appeal in the Security Council for it to 'assume its responsibilities with regard to the situation in Libya, including taking the necessary measures to impose a no-fly zone' and measures to ensure the protection of civilians in Libya.[980] As discussed above, the aspects of the situation (relating to the gravity and imminence of atrocities committed against these populations and the conduct of the Libyan government in this regard) that Lebanon professed in its justification communicated the argument that the substantive threshold for such international action had been met.

All things considered, Lebanon's vote in favour of Resolution 1973 is a prime example of the need to protect vulnerable populations taking precedence over an otherwise strong opposition to the use of force. This is poignantly revealed in the concluding paragraph of Lebanon's justification for its vote, which states: 'Faced with the great suffering being experienced by the Libyan people, the loss of life and the great dangers that still exist, although this resolution falls short of our expectations, we hope that it carries a great deal of hope for a better future for Libya and its valiant people'.[981] Once again, Lebanon referenced key determinants of 'manifest failing', namely the present gravity of the situation in Libya and the imminent threat (i.e. prospective 'danger') that civilian populations were faced with, thus sending a clear message that its decision was very much informed by the Libyan government's blatant failing to fulfil its responsibility to protect its populations.[982]

[979] Adams, 'Libya and the Responsibility to Protect in Libya', p. 7.
[980] S/PV.6498, p.3.
[981] Ibid. p.4.
[982] Ibid.

The analysis of the statements made by the P2 and E10 suggests that the positions of most states were explicitly informed by the notion that a substantive threshold for collective international action had been crossed. With the exception of China, all states referenced aspects of the situation indicating that a substantive threshold had been crossed in explanations of their vote on Resolution 1973. For those who admitted to weighting their reservations regarding the use of force against the need to act, the latter took precedence over the former and led them to either vote in favour of Resolution 1973 (Lebanon) or abstain (Russia, Brazil, Germany and India). Even if the substantive threshold for international action was not of primary or even genuine concern for those who abstained, the overwhelming consensus that the degree of the Libyan government's 'manifest failing' was commensurable with the adoption of 'all necessary measures' made it difficult for them to justifiably wield their veto power.

Bellamy and Williams make a similar albeit broader argument, that the abstentions of UNSC members who had misgivings about the use of force can be attributed to their 'belie[f] that they could not legitimize inaction in the face of mass atrocities'.[983] This suggestion ties into the discussion in Chapter II on how political principles constrain the behaviour of states and how a state's political actions may be limited or hindered in the absence of plausible and acceptable justifications to legitimate them (see subsection 'Relying on the statement of intervening states'). Specifically, Bellamy claims that 'Once states accept that international society has a responsibility to protect and that the proper question should be how rather than whether to fulfill that responsibility, they have limited political room for maneuver in the face of compelling evidence of an imminent threat of mass atrocities'.[984] Here Bellamy opens the important caveat that 'most cases are sufficiently complex to allow states to accept the need for action but at the same time argue on prudential grounds about the most appropriate form of action, limiting the frequency of "timely and decisive" action'.[985] In addition, as discussed in this and previous chapters there are other threshold requirements that play a role in whether coercive international action is agreed upon and actually takes place, i.e. the absence of adequate peaceful alternatives, a UNSC mandate and regional support. Still, determining

[983] Bellamy and Williams, 'The new politics of protection?', p. 844.
[984] Bellamy, 'Libya and the Responsibility to Protect', pp. 266-267.
[985] Ibid. p. 267.

whether there is a just cause for action is undoubtedly the first step in responding collectively to suspected mass atrocities.

To return to the case under discussion, what made it difficult for states to justify inaction in Libya was that the need to act was generally agreed upon and that the other prerequisites for coercive action were also satisfied. As Bellamy puts it:

> 'In the face of Qaddafi's record and his public utterances… states could not plausibly argue that the threat of mass atrocities was not real. And in the face of Benghazi's imminent collapse, they could not plausibly argue that the threat was not imminent. With strong regional support (for whatever reason) and in the absence of reasonable alternative policies for preventing a massacre in Benghazi, those Council members that remained skeptical about the use of force believed that they could not legitimize inaction'.[986]

Similarly, Garwood-Gowers argues that when confronted with the prospect of 'attract[ing] significant criticism and damag[ing] their international reputations' should they block enforcement action in the face of the gravity and imminence of human rights violations in Libya, the five states chose to abstain despite their reservations about the use of force.[987] This is to say that regardless of the importance individual states attributed to the determination as to whether a substantive threshold had been met in Libya, the 'manifest failing' of the Libyan state played a role in shaping their position on Resolutions 1970 and 1973. Hence, the aspects of the situation in Libya indicative of the regime's 'manifest failing' that UNSC members declared to have informed their position on the adoption of coercive measures did in fact influence their decision-making on Libya in one way or another, because they could not legitimate inaction in the face of mass violence and in the absence of other reasonable excuses.

CONCLUSION

Chapter II already established that due to the constraints on state behaviour that emanate from having to provide acceptable reasons for a given course of action in line with the moral standards of a given society, examining the justifications that states profess to have informed their actions amounts to examining one of the key determinants for their

[986] Ibid.

[987] Garwood-Gowers, 'China and the "responsibility to protect"', pp. 384-385.

actions. Consequently, this chapter set out derive the determinants and indicators of a 'manifest failing' from the statements of states who participated in decision-making on the international response to the Libyan crisis. This analysis of justificatory discourses surrounding the Libyan intervention produced important findings with regards to the way in which states understand the substantive threshold for collective action in instances of mass atrocity crimes.

As far as the P3 are concerned, there were no contradictions as to the determinants that informed their understanding of whether the 'manifest failing' threshold was met in Libya. There was significant overlap in the aspects of the situation that they focused on, namely the type of human rights violations taking place, the gravity of the violations, the imminence of the threat of atrocities, evidence of host state intent, and his lack of compliance with the demands made by the international community. In addition, the US and the UK attributed great importance to Qaddafi's history of violence and negative interactions with the international community to cast a shadow on his regime's prospective conduct.

The analysis of the meetings at the vote on Resolutions 1970 and 1973 showed that the aspects of the situation that the P3 stressed to make the case that the international community ought to respond to the predicament of Libyan populations largely coincided with the factors professed by the P2 and E10 when they justified their stance on international action to protect civilians. This is to say that the broader aspects of the situation in Libya that the P3 associated with the threshold for collective action were echoed in the collection statements made by UNSC members in the meetings analysed here, but does not mean that each of the P2 and E10 invoked all determinants mentioned above in their individual statements. Excepting China, all UNSC member states touched upon at least one of these determinants – the nature, gravity, imminence of atrocities in Libya, the intent and conduct of Qaddafi's government – to indicate that the Libyan regime was 'manifestly failing' in its responsibility to protect. Whilst there is no question that the use of force was a major sticking point in Security Council debates, none of the five states who abstained rejected the substantive factors that other member states had professed when making the case that this threshold had been crossed. In other words, the misgivings about the use of force that some states retained emanated from a mix of principled, practical and prudential considerations concerning military intervention as opposed to substantive objections that the threshold for such action had not been met.

It is worth noting that this was also the case in subsequent meetings of the Council during the active phase of the military intervention in Libya, despite the fact that some voiced strong objections to the way in which the mandate of Resolution 1973 was exploited.[988] As Bellamy phrases it:

> 'In the strictest sense, therefore, the problem was not so much the use of force to protect civilians from mass atrocities – … duly authorised by the Security Council – but the facts that this use of force resulted in regime change and that this result was intended by those responsible for implementing the Security Council's decisions even though the Council itself had not specifically authorised regime change.[989]

Likewise, Kurtz and Rotman observe that 'even governments who otherwise show callous indifference to human suffering allowed themselves to be convinced of the need for intervention, both diplomatic and military', in the face of the 'brutal suppression'.[990]

Without going into detail, UNSC member states did not question the fact that a substantive threshold for international action had been met in Libya in subsequent meetings in the period from 1 April to 22 December 2011. On the contrary, some of them (e.g. Germany, France, and Lebanon) even reaffirmed their belief that this was the case at the time of the vote on Resolutions 1970 and 1973 with the benefit of hindsight,[991] while others simply reiterated their support for the Libyan intervention (e.g. the UK, Russia, and Bosnia and Herzegovina).[992] What is more, other UN members, including Gabon, Italy, Switzerland, Japan, Liechtenstein, the United Arab Emirates, Australia, Croatia, Canada, Austria, and Luxembourg, also expressed their support for the actions

---

[988] A notable case in point was Nicaragua, which condemned the 'shameful manipulation of the slogan "protection of civilians" for dishonourable political purposes, seeking unequivocally and blatantly to impose regime change, attacking the sovereignty of a State Member of the United Nations and violating the Organization's Charter.': S/PV.6531, p. 34. Venezuela was likewise critical of the pursuit of regime change in Libya, the undue civilian casualties and collateral damage, and carrying out acts of aggression 'in the name of the responsibility to protect'.: S/PV.6531 (Resumption 1), pp. 18-19; UNSC, S/PV.6650 (Resumption 1), 9 November 2011, p. 26. Similarly, Syria pointed to 'the 130, 000 Libyan civilians killed' by NATO members under the pretext of civilian protection.: S/PV.6650 (Resumption 1), p. 27.

[989] Alex J. Bellamy, 'The Responsibility to Protect and the Problem of Regime Change' in Thomas G. Weiss et al. *The Responsibility to Protect: Challenges and Opportunities in Light of the Libyan Intervention* [Online] (e-International Relations, 2011): 20-23, p. 22. [Accessed 12 August 2020] Available at: https://www.e-ir.info/wp-content/uploads/R2P.pdf

[990] Gerrit Kurtz and Philipp Rotmann, 'The Evolution of Norms of Protection: Major Powers Debate the Responsibility to Protect', 30 (1) *Global Society* (2016): 3-20, p. 17.

[991] S/PV.6531, pp. 18, 23; S/PV.6650, pp. 19-20.

[992] S/PV.6531, pp. 8, 9, 12.

the Security Council decided to take with regards to the situation in Libya,[993] with some specifically pointing to the Libyan regime's 'manifest failing' as the justification for this international response (e.g. Chile, Norway, and Turkey).[994] Hence, it can be said that the arguments made by the P3 regarding the substantive threshold for collective action were not only echoed by the majority of UNSC members, but also resonated with the wider UN membership.

Overall, the Libyan case provided a detailed insight into what a 'manifest failing' entails from the perspective of state actors who were faced with the dilemma as to whether to take coercive action against another state or not in the context of the threat of mass violence against the civilian populations of the host state in question. In addition, this chapter highlighted that the aspects of the situation that states professed when arguing that a substantive threshold for collective action had been crossed accorded with the 'manifest failing' determinants and indicators identified in Chapter I and III. In so doing, it demonstrated that there is meaningful overlap between the factors that 'manifest failing' and 'unwilling or unable' scholars have identified as relevant to determining whether these thresholds have been crossed and the factors that decision-makers placed rhetorical emphasis on when making the case this threshold had been crossed in Libya. Moreover, the analysis of discourses presented in this chapter shed light on addition indicators associated with these determinants, which will contribute to their specification as part of the 'manifest failing' framework outlined in Chapter V. The full list of the 'manifest failing' determinants and indicators identified in this chapter has been compiled in the table below.

Table 4.2. The Determinants and Indicators of Libya's 'Manifest Failing' Cited by States

DETERMINANTS	**INDICATORS**
The information of relevant human rights/acts of violence	The killing, displacement and causing grave physical harm to civilian populations, egregious violations of international and human rights law, reports of arrests, detention and torture, and other relevant human rights violations seen to amount to evidence of crimes against humanity.

[993] Ibid. pp.22, 25-26, 28, 32-35; S/PV.6531 (Resumption 1), pp. 2, 15-17; S/PV.6650 (Resumption 1). p. 9.

[994] S/PV.6531 (Resumption 1), pp. 9, 11-12. See Part I of this chapter for further details and examples.

The gravity of human rights violations	Evidence of exceptional gravity/seriousness of the situation, including death toll, number of internally displaced persons, number of injured, evidence of widespread or large-scale violations of human rights. The deliberate targeting of civilians, disproportionate violence against civilian populations, defenceless civilian populations being massacred in a large-scale military action against them/ large-scale military attacks on unarmed, non-violent, civilian populations.
The imminence of atrocities	Preparations for attack on civilian populations, ongoing violence, the continuation of violence, the escalation of violence, prospective death toll ('the life and death of a thousand men) / likelihood of mass killings, the impact of future developments on civilian populations, the continued brutal and bloody repression against populations, the prospect of an imminent massacre.
The role of the state in the planning commission and perpetuation of atrocities	*Public statements of intent*, including hate speech and the threatening statements made by the Libyan leadership. *Evidence of systematic preparations*, including systematic attacks on civilian populations who were massacred by military means, evidence of human rights violations committed by the government, including the killing and internal displacement of civilian populations, indiscriminate violence against civilian populations, the use of military force against civilian populations, the use of government weaponry, the use of foreign mercenaries, live ammunition, heavy weapons and other methods against peaceful demonstrators, constituting grave violations of human rights and international humanitarian law.
The host state's response to international requests and measures to address the threat of atrocities	Diplomatic measures failing to deter the host state from committing grave human rights violations. Lack of compliance with, defiance against or disregard for the demands made by the international community, especially those made by regional organisations and the Security Council in Resolutions 1970 and 1973, e.g. to put an end to the violence, to stop the offensive on Benghazi.

History of mass violence and negative interactions with the international community	Government history of human rights violations/acting as an agent of mass violence, implying it could carry out a threat to kill tens of thousands of people.

CHAPTER V: A NORMATIVE FRAMEWORK FOR 'MANIFEST FAILING' INQUIRIES

This chapter brings together the multiple perspectives on the 'manifest failing' and 'unwilling or unable' thresholds articulated in Chapters I, III and IV in order to synthesise determinants relevant to establishing whether a host state is 'manifest failing' in its responsibility to protect its populations from mass atrocities. Specifically, it proposes a normative framework for 'manifest failing' inquiries premised on the investigation of existing contributions to specifying the determinants and indicators of a 'manifest failing' and the scholarship on mass atrocities discussed in Chapter I, the exploration of the substantive factors of the 'unable or unwilling' test in Chapter III, and the determinants and indicators of a 'manifest failing' that states professed in the context of the Libyan intervention, specified in Chapter IV. The present chapter is structured in three parts. The first part reflects on the substantive basis of this framework, its composition and utility. The second part elaborates on each determinant and its associated indicators. The third part showcases the application of this framework to the ongoing 'war on drugs' in the Philippines. In so doing, it illustrates how the framework can be utilised to determine whether a government is 'manifestly failing' to protect its populations in the context of divergent opinions as to whether a situation ought to viewed through an R2P / mass atrocity lens.

I. FRAMEWORK COMPOSITION AND UTILITY

The content of this framework for 'manifest failing' inquiries has been informed by several key observations pertinent to the threshold for collective action in response to atrocities that have been noted in this thesis. First, the 'manifest failing' of a host state to protect populations from atrocity crimes is not synonymous with the mere presence of grave human rights violations, even if they are deemed to amount to atrocity crimes. This means that for a 'manifest failing' inquiry to be warranted the threat of atrocity crimes must be of a gravity that casts doubt on the host state's willingness or ability to protect its populations. Second, and relatedly, should such a suspicion arise, then ascertaining the role of the host sate in mitigating the risk of atrocity crimes and its capacity to do so is crucial to determining whether it is 'manifestly failing' to protect its populations. Third, the 'manifest failing' threshold requirement should be understood as a process. This

means that collective action may be justifiable when there is sufficient evidence to establish that a 'manifest failing' is underway in order to prevent an atrocity crisis from arising or escalating, rather than wait until the host state has utterly failed to protect its populations, at which point the international community would have also failed in its subsidiary responsibility to protect said populations. This resonates with what Rosenberg and Strauss term 'the normative concerns embedded within the RtoP, most of all its ability to proactively attempt to prevent imminent or on-going forms of mass atrocities based on existing legal obligations'.[995] Consequently, 'manifest failing' inquiries will often involve both prospective and retrospective investigations, in order to determine whether a sufficiently grave threat of international crimes exists to raise concerns that the host state may be 'manifestly failing' to protect its populations from these crimes and the probability that a future serious threat of atrocities will manifest itself if the host state fails to mitigate the risk of the four crimes.

The determinants specified in the framework outlined in this chapter represent general aspects of the situation or categories of information that should be taken into consideration when assessing whether the threshold for international action has been crossed, whereas the indicators are the concrete manifestations of these determinants. The lists of indicators provided for each determinant have been drawn from a variety of sources to ensure that they are comprehensive enough to aid the assessment of a variety of situations, namely the scholarship on mass atrocities, 'manifest failing' and existing tools for the assessment of atrocity crimes (see Table 1.1 in Chapter I), the scholarship on the 'unwilling or unable' test through reasoning by analogy (see Table 3.1 in Chapter III) and an independent investigation of states' understanding of the 'manifest failing' threshold in the context of the Libyan crisis (see Table 4.2 in Chapter IV).

The framework proposed in this chapter can be utilised in all situations found to be at risk of atrocity crimes, so as to determine whether the host state is 'manifestly failing' to protect its populations. In this sense, it can complement frameworks for early warning and prevention of atrocity crimes and the work of non-governmental organisations that monitor situations for the risk of atrocity crimes or analyse specific country cases. Specifically, once a situation has been deemed at risk of atrocity crimes, it can help

[995] Rosenberg and Strauss, 'A Common Approach', p. 59.

relevant stakeholders to assess the role of the state in creating, perpetuating or mitigating the risk of the four crimes.

Like the R2P itself, this framework is not prescriptive, namely it does not specify in advance what measures must be adopted in a given situation. Rather, it offers a tool for assessing situations at risk of atrocity crimes on a case-by-case basis more effectively through a continuum of steps, which can also help relevant stakeholders to determine what measures (if any) are most appropriate in a particular situation. Unlike existing frameworks for the analysis of atrocity crimes, it has assembled an array of determinants and associated indicators that can help to ascertain whether a state is 'manifest failing' in its responsibility to protect its populations from mass atrocities, namely whether the substantive threshold requirement for collective international action under the R2P has been met. Therefore, it can also aid determinations as to whether 'last resort' measures under Chapter VII of the UN Charter are justifiable or not in response to the threat of atrocity crimes that populations abroad are faced with.

It is important to note that, while determining that a state is 'manifestly failing' to protect its populations from the four crimes may indicate that there is a substantive basis or a need for a strong international response, this does not automatically legitimise coercive action. As underscored in the Introduction and Chapter IV of this thesis, it depends on a number of other conditions being satisfied. First, coercive action is premised upon the inadequacy of peaceful measures, which need to be considered first. This means that even if the 'manifest failing' inquiry suggests that a situation merits a strong international response, any form of coercive action must be taken as a 'last resort', that is, when 'first response' measures have been deemed inadequate. Second, it requires UN Security Council authorisation. Third, regional support is a key catalyst for coercive action and deemed to be one of the three prerequisites for the use of force, as highlighted by states in the context of the Libyan crisis. By virtue of the latter two requirements, the incidence and nature of international response to any given situation in which there is evidence of the host state's 'manifest failing' will depend on a range of political and prudential considerations.

Nonetheless, as outlined in the Introduction, determining whether the 'manifest failing' threshold has been met should not be understated, given that the knowledge that the international community has a substantive basis to do something to prevent or halt serious

atrocity crimes, but does nothing, inevitably prompts questions as to the reasons why this is the case. Conversely, a determination that the host state is not 'manifestly failing' in its protection responsibilities, would undermine the success of attempts to advocate for or undertake unwarranted coercive action against that state.

II. THE PROPOSED FRAMEWORK

The framework outlined here proposes that 'manifest failing' inquiries should be based on nine determinants, namely the determinations of 1) human rights violations and acts of violence associated with atrocity crimes; 2) the gravity of the threat of atrocity crimes, 3) the imminence of the threat of atrocity crimes; 4) host state intent; 5) the host state's territorial control and capacity to supress the threat of atrocity crimes; 6) measures proposed or taken by the host state to address the threat; 7) host state cooperation; 8) the host state's response to international requests to address the threat, and 9) the host state's track record of violence and gross human rights violations and its prior interactions with the international community on matters relevant to atrocity crimes. These determinants pertain to the assessment of the threat of atrocities, the capacity and conduct of the host state and its present and past interactions with the international community.

Table 5.1. A Normative Framework for 'Manifest Failing' Inquiries

DETERMINANTS	INDICATORS
1. Relevant human rights violations / acts of atrocity crimes	Killing, murder, deaths caused by induced famine, deaths caused by indiscriminate targeting, torture, inhuman treatment, mutilation, rape and sexual violence, enslavement, abduction, or otherwise deliberately causing great suffering, physical or mental harm, extra-judicial executions, arbitrary arrest and detention, deliberate military attacks or threats of attacks on civilians and civilian areas, expropriation, wanton destruction of property, looting, destruction of subsistence food supply, denial of water or medical attention, destruction or denial of medical supplies, man-made famine, redirection of aid supplies, forced population movement (e.g. forcible removal, displacement and deportation of civilian populations), and restricted movement, (e.g. confinement of civilian populations in ghetto areas), discrimination in access to work and resources, discrimination in education and lack of access to justice and redress, lack of freedom of speech/ press/ assembly/ religion, political marginalization.

2. The gravity of the human rights violations taking place	*Magnitude and scope of relevant human rights violations*: - the number of victims of violence or level of irreparable harm inflicted by the human rights violations specified under Determinant 1 (number of dead, internally displaced, severely injured, tortured, victims of sexual violence) - the portion of the victim populations/group affected by the violence *The targeting of identifiable groups*, where: victims are identified on the basis of identity criteria related to national and ethnic origin, religion, gender, race, colour, descent, place of birth, sexual orientation and other identity markers, or on grounds linked to their association with a political opinion or group, or any other group or collective as identified by the perpetrator (state or non-state actor). *The targeting of civilians and especially vulnerable groups*, such as women, children and the elderly.
3. The imminence of the threat of atrocities	*The imminence of the threat of atrocity crimes should be assessed on the basis of the gravity and the probability of the threat, by reference to all relevant aspects of the situation, including:* the probability of grave human rights violations; the likely scale and magnitude of the threat of atrocities and the loss/harm or grave human rights violations likely to result therefrom in the absence of mitigating action; whether anticipated large-scale atrocities are part of a concerted pattern of continuous gross human rights violations (i.e. continuing/ongoing violence); the escalation of violence in qualitative or quantitative terms.
4. Host state intent	*Public statements of intent*: Evidence from primary and secondary sources (including political manifests, media records, official documents) for types of hate speech known to precede or accompany episodes of mass atrocities, including statements that atrocity-related violence was directed towards a specific group, statements of intent to destroy a specific group or intimidate a group by creating the perception of an existential threat an or incitement to violence. *Systematic preparation*:

	the selective killing of members of a group with the intent of communicating to other members that their survival may be threatened, the intentional targeting of vulnerable groups (including women, children and the elderly), impunity for or tolerance of acts of atrocities or associated serious human rights violations, evidence of systematic organisation, logistical or political preparation to commit or facilitate large-scale or systematic gross human rights violations, the pattern of refugee origin and internal displacement, systematic use of sexual violence against a population, measures to prevent births including forced sterilization and separation of the sexes, sexual trauma and impregnation through gang rape, humiliation of a group publicly or in the media, denial of the existence of recognised group or parts of their identity, exclusionary ideologies that claim to justify discrimination, existing discriminatory practices, policies and legislation, destruction of cultural symbols, the separation of people on the basis of identity, clear territorial control by the perpetrator group in the affected area, support to or failure to condemn the actions of non-state groups accused of committing gross human rights violations / acts of violence associated with atrocities or failure to condemn their actions, the systematic use of government weaponry against civilian populations, the use of heavy weapons, the quantity of weapons (heavy or light) used, the use of weapons prohibited under international law, such as chemical and biological weapons, indiscriminate weapons, or weapons, methods and materials of warfare causing unnecessary harm or suffering, the proliferation of all categories of weapons, including the illicit proliferation of small arms and light weapons, the recruitment of mercenaries (by the host state) charged with the task to intimidate or attack civilian populations).
5. The host state's control and capacity	*In the presence of the following indicators, the host state may be rendered unable to address the threat of atrocities:* the host state's lack of control over relevant parts of its territory; the host state's lack of effective civilian control of security forces; security or military forces who possess the requisite capacity but are unwilling to act against a non-state actor; failure to hold perpetrators to account; weak, untrained or underresourced law enforcement; lack of capacity to ensure that means and methods of warfare comply with international humanitarian law standards; insufficient resources to implement overall measures aimed at protecting populations; lack of awareness of and

	training on international human rights and humanitarian law to military forces, irregular forces and non-State armed groups, or other relevant actors; law enforcement institutions that lack sufficient resources, adequate representation or training; lack of awareness of and training on international human rights and humanitarian law to military forces, irregular forces and non-State armed groups, or other relevant actors.
6. Measures taken by the host state to mitigate the threat of atrocities	*In addition to the specific measures taken by the state in a given case, the state of implementation of the following measures and their impact on the real risk that exceptional grave violations of human rights could occur in the future should be considered, namely whether the host state:* publicly recognises and condemns relevant violation of human rights, issues unambiguous and public orders to its police, security and military forces to respect international humanitarian and human rights law, criminalises serious human rights violations and atrocity crimes, conducts thorough investigation into reported human rights violations, enforces accountability for relevant violations of human rights by arresting and/or prosecuting the authors of those attacks, ensures humanitarian assistance and protection for victims of violence, formulates an action plan with timelines for implementing specific measures to mitigate the risk of atrocities, in cooperation with relevant stakeholders (including affected populations and prospective victims of atrocities).
7. Host state cooperation	*The host state's unwillingness to protect its populations can be evidenced by its*: inadequate cooperation of the host state with regional and international human rights mechanisms; demonstrable unwillingness on the part of the host state to engage in dialogue, make concessions and receive support from the international community; the host state does not allow/impedes the provision of humanitarian assistance to affected populations; the host state does not allow/impedes the access of international human rights monitors and investigations into the human rights situation in the country; the host state impedes the safe passage of humanitarian and medical supplies, and humanitarian agencies and workers, into the country; the absence or restricted presence of regional or international organisations / actors on the territory of the host state with access to populations; failure to meet / comply with

	concrete demands to enable or cooperate with regional or international actors to address the threat of atrocity crimes.
8. The host state's response to international requests to address the threat	*The host state's response to and compliance with international requests to address the threat is to be evaluated against specific demands made by the international community, which could be expressed in the resolutions of regional (EU, ASEN, AU, LAS etc.) and international bodies (UNGA, UNSC, UNHRC), including:* to take all measures to protect civilians and meet their basic needs; to ensure the rapid and unimpeded passage of humanitarian assistance; an immediate end to the violence; the immediate establishment of a cease-fire and a complete end to violence and all attacks against, and abuses of, civilians, to comply with their obligations under international law, including international humanitarian law, human rights and refugee law; act with restraint, respect human rights and international humanitarian law, to impose sanctions on perpetrators.
9. The host state's track record of grave human rights violations and relevant interaction with the international community	*The following could help determine whether the host state is 'manifestly failing' to protect its populations when considered in conjunction with other determinations*: past acts / human rights violations associated with atrocity crimes (specified in Determinant 1) or their incitement; a past history of violence against a perceived group; a track record of the perpetrators host state acting as an agents of mass violence or tolerating such violence against their populations; a past history of violence or incitement to violence against perceived groups; a historical record of legislation, policies or practices of discrimination against, marginalisation or segregation of perceived groups; tolerance, impunity or failure to seek accountability for past human rights violations; denial or justification of past atrocities; track record of supporting non-state actors suspected or accused of the commission of atrocity-related acts or human rights violations or failing to condemn their actions; historical lack of compliance with or failure to meet the demands of the UN and regional organisations to protect populations from the threat of atrocities; past inaction, unwillingness or refusal of the host state to take all possible measures to put an end to imminent or ongoing atrocities

Determinant 1: Human rights violations associated with atrocity crimes

The first determinant of this framework embodies the substantive dimension of the threat of mass atrocity crimes and outlines the types of serious human rights violations whose presence could raise suspicion that the host state may be 'manifestly failing' in its responsibility to protect its populations. The indicators listed under this determinant constitute a comprehensive list of acts amounting to serious human rights violations that have been associated with atrocity crimes (see Table 4.1). They are premised on the discussion of mass atrocity crimes and the acts associated with them found in the scholarship on mass atrocities and 'manifest failing' in Chapter I, drawn from past cases of atrocity crimes, the literature on mass violence and the definitions of atrocity crimes, specified in Annex I.

The determination of relevant human rights violations is an essential first step in judging a host state's a 'manifest failing' for several reasons. First, it is vital to establishing whether there are substantive grounds to suspect that the host state is 'manifestly failing' to protect its populations from the four crimes. As such, it is a prerequisite to determining the gravity and the imminence of the threat of atrocities, which offer insights into the role of host state and are indispensable to determining whether it is 'manifest failing'. What is more, as evidenced in Chapter IV, the determination of relevant human rights violations is a central aspect of establishing the need for coercive action. This is affirmed in Puri's take on Pillar III, which suggests that '[c]learly, before embarking on action under Chapter VII of the UN Charter and authorizing the use of force, the Security Council would want to satisfy itself that the alleged violations fall within one of the categories of mass atrocity crimes'.[996] As discussed in Chapter I, whereas human rights abuses cannot provide a trigger point for Chapter VII action, when gross or persistent human rights violations occur a determination would need to be made as to whether the 'manifest failing' threshold has been crossed.[997]

Determinant 2: The gravity of atrocity-related human rights violations

The determination of the gravity of gross human rights abuses associated with atrocity crimes seeks to respond to the question of 'How severe is the threat of atrocity crimes?'.

[996] Puri, *Perilous Interventions*, p. 200.
[997] Ibid. p. 199-200.

It is underpinned by the notion that a state can be deemed 'manifestly failing' at a specific level of seriousness of existing or potential human rights violations, which is set above examples of small-scale atrocity crimes and below establishing that one of the recognised atrocity crimes is taking place under international law. Given that the determination of gravity/severity of the threat has been underscored in, and thus unifies the continuum of, different perspectives on the 'manifest failing' and 'unwilling or unable' thresholds explored in this thesis, its importance cannot be overstated.

The indicators listed under this determinant (see Table 5.1) represent quantitative and qualitative aspects of the threat of atrocities, which have been found to be most relevant to determining the seriousness of current and prospective atrocity-related human rights violations. They have been divided into three thematic categories: 1) magnitude and scope of atrocity-related human rights violations; 2) the targeting of identifiable groups; and 3) the targeting of civilians, especially vulnerable groups.

It is important to stress that the gravity of the threat (concerned with the severity of the threat of atrocities) should not be confused with the probability of the threat, despite the fact they are interconnected. Assessing the prospective gravity of the threat of atrocities addresses the question of 'how severe the plight of innocent populations will be' if the threat manifests itself, rather than the question of how probable it is that the threat will manifest itself. In other words, the indicators listed under the gravity dimension in this framework are those most relevant to determining the seriousness of current and prospective atrocity-related human rights violations rather than the likelihood of these violations occurring. The relationship between the two will be elaborated on in the subsequent discussion of the third determinant of the present framework – the imminence of the threat of the threat of atrocities.

Determinant 3: The imminence of the threat of atrocities

The determination of the imminence of the threat of atrocity crimes is concerned with the probability of grave human rights violations occurring in the future. As such it is a prospective assessment of future developments of the situation based on present facts and circumstances, which involves considering the interplay between the prospective gravity and the probability of the threat of mass atrocity crimes. Chapter III has already elaborated on the debates surrounding its content and indicators relevant to its determination in the context of self-defence against non-state actors, whereas Chapter IV evidenced its

presence in discussions surrounding the threshold for coercive action in the run-up to the military intervention in Libya.

To recap, as the discussions and the indicators related to imminence identified in both chapters (III and IV) suggest, the 'imminence' requirement is concerned with the probability of the threat rather than its temporality, which belongs to the requirement of 'necessity', embodied in the condition that 'peaceful means be inadequate' in the context of Paragraph 139 of the WSOD. As discussed in Chapter III, the latter requirement ensures the last point in time at which an effective peaceful response is possible has been reached. Hence, while the temporal remoteness of the threat will have an impact on how easy / difficult it is to determine the gravity and probability of the threat, the latter determinations are ultimately made independently of the proximity of the threat. Once a sufficiently severe and probable threat is identified, temporal factors can play a role in determining whether coercive action is necessary, based on the absence of any peaceful alternatives, including later in time. In sum, similarly to prospective assessments of the threat made by domestic and international courts discussed by Rosenberg and Strauss, determining the imminence of the threat involves 'engag[ing] in balancing of the probability of an event occurring based on the evidence available at the time of the decision, with the level of harm that would occur if such a situation would develop'.[998] Hence, the imminence of the threat of atrocity crimes should be assessed on the basis of the gravity and the probability of the threat, by reference to all relevant aspects of the situation, reflected in the indicators attached to this determinant (see Table 5.1).

The relevance of temporal factors has also been questioned in debates about the just cause for humanitarian intervention and coercive action under the R2P, which give support to the above understanding of imminence and shed light on the relationship between the probability and the temporal proximity of the threat. According to Pattison, the Libyan intervention poses the question of 'why should the focus be on *impending* mass violations of basic human rights?' (emphasis added).[999] More specifically, he asks 'If the failure to intervene against Gaddafi would have been very likely to significantly increase the probability of the mass violation of basic human rights, but not imminently, is this still sufficient for just cause?'.[1000] In so doing, Pattison juxtaposes the relevance of the

[998] Rosenberg and Strauss, 'A Common Approach', p. 69.

[999] Pattison, 'Perilous Noninterventions?', p. 222.

[1000] Ibid. pp. 222-223.

temporality of the threat against the probability of threat, not unlike Akande and Liefländer (discussed in Chapter III). In a bid to resolve this tension, Patison scrutinises Puri's argument that once the presence of gross or continuous human rights violations has been established, the trigger for Chapter VII action is the 'imminent threat of mass atrocities'.[1001]

According to Pattison, even Puri contradicts his claim that the 'imminence' of the threat should be the threshold for coercive action, because his subsequent elaboration on his understanding of just cause does not incorporate a temporal dimension. Namely, Puri claims that '[o]nce the determination is made that there is enough evidence in terms of intent, and evidence that inaction by the international community will result in large-scale mass atrocities, the international community would obviously have a role to play'.[1002] Indeed, Puri only highlights host state intent and the prospective gravity and probability of the threat. Hence his understanding of 'imminence' as the trigger for international action seems to align with Akande and Liefländer's argument (see pp. 151-153) that the gravity and probability of the threat, rather than its temporality, are the true definitive elements of imminence.

While Pattison does not question the definition of 'imminence', he posits the question of whether and to what extent 'imminence', understood as a temporal factor, should be a determinant of coercive action under the R2P. Similarly to Akande and Liefländer, Pattison suggest that on the one hand, it appears that what matters is not the temporality of the threat but its probability, i.e. 'the (strong) likelihood that mass violations of basic human rights will occur, rather than when exactly they will occur…(which may not always be determined by its temporality)'.[1003] For example, at first glance, 'there does not seem to be any difference between halting an attack that is very likely to occur (1) tomorrow, compared to halting an attack that is very likely to occur (2) in one month's time'.[1004] The alternative view is that the imminence requirement should be maintained in a bid to 'reduce the risk of abuse and undermining international instability', which

[1001] Puri, *Perilous Intervnetions*, pp. 199-200.

[1002] Ibid., p. 200.

[1003] Pattison, 'Perilous Noninterventions?', p. 223.

[1004] Ibid.

would arise if a more permissible requirement of just cause that offers a pretext for justifying preventative intervention is adopted.[1005] Ultimately, Pattison takes the view:

> 'that the risks of abuse suggest that there should be a general presumption against preventive intervention – imminence should generally be required [,] [b]ut this presumption might be occasionally overridden in the face of compelling, robust evidence that demonstrates the strong likelihood of atrocities, plus safeguards against abuse, such as having UN Security Council authorisation'.[1006]

According to him, this condition was satisfied by the Libyan intervention, as far as its initial phrase is concerned.[1007]

What makes the Libyan case even more complicated, is that in addition to the seemingly very likely but more distant human rights abuses in Libya, there were also the potential, but more contested impeding abuses in Benghazi, even though claims of Benghazi's imminent fall (notably made by France) were questioned by some (notably India).[1008] Significantly, while the US was not convinced of the temporal proximity of the threat looming over Benghazi, it did emphasise the need to act based on the imminence of the threat. This understanding is evidenced in Obama's recounting of the circumstances under which the idea of military intervention gained traction in Libya, namely:

> 'But the Libyan leader was undeterred [by Resolution 1970]. Analysts forecasted that once Gaddafi's forces reached Benghazi, tens of thousands of lives could be lost.
>
> It was around this time that a chorus grew, first among human rights organizations and a handful of columnists, and then members of Congress and much of the media, demanding that the United States take military action to stop Gaddafi'.[1009]

[1005] Ibid.
[1006] Ibid.
[1007] Ibid.
[1008] Ibid. p. 224.
[1009] Obama, *A Promised Land,* p. 654.

While his statement makes no reference to the temporal proximity of the threat, it places emphasis on the prospective gravity of the threat in the event that it manifests itself as the grounds for international consensus on coercive action in Libya, which gave rise to demands the US joins the bid for a collective response to the atrocities taking place in the country. Hence, it aligns with Pattison's observation that '[t]his potential (but, on some accounts, less likely) [impeding] attack seem[ed] to provide further reason to override the general presumption against preventive action in this case'.[1010] In sum, Pattison's discussion suggests that while temporal proximity of the threat should always be considered as a general rule because it poses a limitation on unwarranted coercive intervention, it appears that when the potential threat of atrocities is grave enough, the probability of its materialisation ultimately takes precedence. This is an important observation that sheds light on the importance of considering the imminence of the threat of atrocities, as defined in this thesis, in determinations as to whether the 'manifest failing' threshold has been crossed.

Determinant 4: Host state intent

The detailed discussion of host state intent and its centrality to determining the role of the state in the planning, commission or complicity in mass atrocity crimes in Chapter I suggested that it is the first step towards determining whether the conduct of the host state suggests that it is 'manifestly failing' to protect its populations. The first chapter of this thesis also concretised the advantages of adopting a contemporary behaviour-based approach to assessing intent in this framework and identified two categories of indicators of behaviour-based intent: namely *public statements of intent* and *systematic preparation*. Public statements of intent involve the examination of primary and secondary sources (including political manifests, media records, official documents) for types of hate speech known to precede or accompany episodes of mass atrocities, including statements that atrocity-related violence was directed towards a specific group, or statements of intent to destroy a specific group or intimidate a group by creating the perception of an existential threat an or incitement to violence. Systematic preparation can help to infer host state intent (by action or omission) from the degree of organisation of violence against vulnerable populations and evidence of the government plans, policies, practices and weapons used that point to its involvement in the commission of atrocity-related human

[1010] Pattison, 'Perilous Noninterventions?', p. 223.

rights violations. As argued in Chapter I, the strongest evidence of intent would encompass indicators from both categories, i.e. perpetrators declare their plans to commit acts of atrocity against their populations and they make manifest preparations employing resources of or at the command of the government.

However, having recognised that declarations of intent are exceedingly rare, their absence does not preclude determining the role of the government in the commission or preparations to commit atrocity crimes entirely based on the degree of organisation and systematic preparation to commit large-scale violence, which reveal the government's plans or complicity in the massacre of their populations. Similarly to the UN 'Framework of Analysis for Atrocity Crimes', the list of indicators under this determinant, provided in Table 5.2, include manifestation of host state intent by action as well as omission. Evidence of either will constitute proof of the host-state's 'manifest failing' to protect its populations.

Determinant 5: The host state's territorial control and capacity to supress the threat of atrocity crimes

In a number of cases the host state's role in the commission of mass atrocities and its willingness to address them cannot be fully appreciated without reference to its capacity and territorial control. As discussed in Chapter III (see pp. 161-165), the 'territorial state's control and capacity' is not only a key factor in determining whether the state harbouring an aggressive non-state actor is 'unwilling or unable' to counter the external threat posed by the latter, but it is equally important to determining whether a host state is 'manifestly failing' in situations where a non-state actor is responsible for the commission of atrocities against the host state's populations. The understanding that the host state's capacity is an essential part of determining whether the threshold for international action has been crossed is also reflected in Puri's interpretation of Pillar III of the R2P:

> 'Under pillar three, if there is a reasonable basis to believe that the state is not shouldering its responsibility or *it does not have the capacity to shoulder the responsibility of protection*, and the intent to commit mass atrocities is evident, the international community has a definite role to play and a responsibility to assume' (emphasis added).[1011]

[1011] Puri, *Perilous Interventions*, p. 200.

As Deeks rightly observes, '[a] state that is well-known to lack control over a relevant part of its territory is quite unlikely to be "able" to suppress threats emanating from that area'.[1012] Likewise, a host state that lacks control over a part of its territory is unlikely to be 'able' to prevent or halt mass atrocity crimes occurring in that area. The same issue arises when the state's military, security or law-enforcement forces lack the requisite capacity to address the threat or the state lacks control over its forces. Based on the discussion of these and other factors affecting the host state's ability to supress the threat posed by a non-state actor in Chapter III and the 'manifest failing' indicators derived therefrom (in Table 3.1), Table 5.1 (above) provides a comprehensive list of indicators of the host state's lack of capacity and control over its territory, whose presence in a given situation can seriously inhibit the host state's ability to adequately respond to imminent or ongoing atrocity crimes. Depending on the gravity and imminence of the threat, this could mean that the national government is not able to fulfil its protection responsibilities, even if it demonstrates its willingness to take the necessary steps to do so.

It is important to note that situations involving atrocities perpetrated by non-state actors are unlikely to be straightforward. The fact that a state is unable to fulfil its protection responsibilities, does not necessarily mean that it is willing to allow foreign interference into its domestic affairs, especially of the military kind. While in some cases this could be down to the host state's wish to preserve its sovereignty, in other cases it could indicate that the 'unable' state is also 'unwilling' to protect its populations, which may be due to its complicity in the planning or commission of atrocities against them. Consequently, although this thesis recognises the importance of considering the threat posed by non-state actors, it does not delve into the complexities of providing military assistance (under Pillar II) in the contexts of intervention by invitation explored by Welsh,[1013] but investigates the role of host state capacity in determining whether that state is 'manifestly failing' to protect its populations.

Fortunately, there is considerable amount of well-established publicly available information about ungoverned and undergoverned areas within states, which can aid determinations as to whether the host state is unable to act in parts of its territory where mass atrocities by the hand of non-state actors are suspected to be taking place. Notable resources include the Fund for Peace' Fragile State Index (FSI, formerly known as he

[1012] Deeks, '"Unwilling or Unable"', p.525.

[1013] See Welsh, 'R2P's Next Ten Years', pp. 992-993.

Failed States Index), which has published annual reports since 2005 and the World Bank's World Governance Indicators (WGI), which measures, among other things, perceptions of the likelihood of political instability and/or politically motivated violence, including terrorism (and, according to Taylor, is 'the most sophisticated attempt to measure governance and have relevance for defining what failure looks like').[1014] Relevant information of the host state's control and capacity can also be found in contemporary country assessments provided by scholars, NGOs, regional organisations and UN bodies, such as the UN Human Rights Council.

As suggested above, the assessment of the host state's control and capacity will also depend on the gravity and imminence of the human rights violations taking place. In other words, the host sate's capacity and control has to be assessed in relation to the severity and imminence of the threat in order to make an accurate assessment of its ability to prevent or halt atrocity crimes. Should the state agree with the international community's assessment that it is 'unable' to adequately respond to the mass atrocity crimes being perpetrated against its populations due to lack of control over relevant parts of its territory, insufficient control over its security forces or their lack of capacity, and ask for or accept international assistance, then a 'manifest failing' inquiry will become redundant. However, in the event that the state disagrees with this valuation and objects to international involvement, then the state could be deemed 'manifestly failing' in its responsibility to protect its populations following the consideration of other relevant determinants. The same applies to more complex cases where the host state's capacity / forces are improving, but not yet adequate to meet the threat.

Determinant 6: Measures taken by the host state to mitigate the threat of atrocities

The state of implementation of measurable steps taken by the host state to mitigate the threat of atrocities or the lack thereof are key to determinations of the host state's 'manifest failing'. The indicators listed under this determinant in Table 5.1 constitute examples of such measures, based on insights from Chapters I, III and IV, which could guide determinations as to whether the host state is taking the requisite steps to address the threat. This is not an exhaustive list, because the nature and character of appropriate measures would depend on what a specific situation requires with regards to the context

[1014] Andrew Taylor, *Failed States* (Basingstoke: Palgrave Macmillan, 2013), p. 31.

in which atrocities are unfolding, their gravity, imminence, as well as the presence of a concrete action plan proposed by the host state.

The implementation of concrete measures to address the threat of atrocities is to be evaluated with reference to their impact on the gravity and imminence of the threat of grave human rights violations. For instance, provided that the host state is not deemed to be 'unable' to take the requisite measures to address the threat, evidence of the escalation of or increase in the magnitude and scope of the threat of atrocities would point in the direction of that state's 'manifestly failing' to protect its populations. Alternatively, the host state's lack of progress in attenuating the threat of atrocities could prompt the international community to reconsider the latter's capacity to deal with the threat and offer assistance, inviting the host state to cooperate with regional and international organisations. Based on Deeks' suggestions, those conducting such evaluations should consider the practical limitations that any states would be expected to face in meeting the threat at hand, including the extent to which a state with a robust military capacity is likely to succeed in the task of fully address the risk of atrocities.

The assessment of the measurable steps implemented by the host state, ensures that a national government that has taken adequate measures to protect its populations from the four crimes would not become the victim of unfair international action against it. On the other hand, inaction by the host state, judged to have the capacity to take action to prevent or halt atrocities, but not doing everything it can to do so, would constitute compelling evidence of this state's 'manifest failing' and provide grounds for the international community to justify collective action against this state under a mandate to protect affected populations. In this sense, this assessment could serve one of two purposes, either to delegitimise or legitimise international action.

Determinant 7: Host state cooperation

As discussed and exemplified in Chapter II of this thesis, the consensual use of force under Chapter VII constitutes a case of international assistance under Pillar II of the R2P, which does not involve an assessment of the host state's 'manifest failing'. In this regard, just as Deeks observes that an 'unable or unwilling' inquiry becomes obsolete once the consent of the state hosting violent non-state actors is secured, the international community need not conduct a 'manifest failing' inquiry if host state consent for an international intervention has been obtained. However, this does not mean that the host

state's willingness to cooperate with the international community to address the threat of atrocities is not relevant to inquiries as to whether it is 'manifestly failing' to protect its populations. For instance, the inadequate cooperation of the host state with regional and international human rights mechanisms, the demonstrable unwillingness on the part of the host state to engage in dialogue, make concessions and receive support from the international community, impeding the provision of humanitarian assistance to affected populations and the access of international human rights monitors and investigations into the human rights situation in the country, among other indicators listed in Table 5.1, could all prove relevant to ascertaining whether the host state is 'manifestly failing' to protect its populations.

It is worth stressing that neither the 'unwilling or unable' test nor the 'manifest failing' requirement dictate that a host state is 'unwilling' to fulfil its responsibilities because it does not accept foreign interference, as this would amount to contesting the host state's territorial sovereignty. In other words, refusing assistance or cooperation, cannot automatically render the host state 'unwilling' and justify collective action against that state. However, when taken into account in combination with other determinants specified in the framework outlined in this chapter, it can contribute to 'manifest failing' inquiries, especially when a pattern of unwillingness to cooperate with the international community to prevent or halt a serious threat of atrocity crimes in the country can be established.

Determinant 8: The host state's response to international requests to address the threat

In the event that the international community has requested that the national government takes action to address suspected atrocity crimes, the host state's response to such requests, be they from regional bodies or international bodies, can elicit important information about the state's willingness to take action in response to suspected atrocities, as well as the next steps that the international community should take. For instance, the Libyan case showed how the host state's response to a sequence of demands made by regional organisations and the Security Council drove further action by the international community and justified the adoption of progressively stronger measures against the Libyan government. It also illustrated how the actions of the host state in light of specific demands made by the international community that the national government takes certain

measures to address the threat of atrocities played a significant role in determining that the latter was 'manifest failing'.

As discussed in Chapter IV, in the Libyan case these demands included taking all measures to protect civilians and meet their basic needs; ensuring the rapid and unimpeded passage of humanitarian assistance; an immediate end to the violence, calls for steps to fulfil the legitimate demands of the population; the immediate establishment of a cease-fire; and a complete end to violence and all attacks against, and abuses of, civilians. These and other examples of international requests discussed in this thesis have been included as indicators of this determination in Table 5.1. The host state's response to these demands, if/when they are made, will naturally complements the determination as to whether the state is taking all possible measures to address the threat of atrocities (Determinant 7).

Similarly to Determinant 7, in the event that the state expresses willingness to address the threat itself, this would require considering whether the host state possesses the requisite capacity to do so. If the international community is satisfied that the host state is both willing and able to protect its populations from the threat of atrocities, then unwarranted action (including the violation of the host state's sovereignty) and expenditure of resources on the part of the international community would be avoided. Having said that, there are a number of other scenarios in which the host state's response to a request made by the international community could constitute proof of its 'manifest failing'.

Hypothetically, an international request to address the threat can trigger an outright refusal to do so, which would provide grounds for a 'manifest failing' determination. However, this appears to be extremely unlikely, especially given the fact that even regimes intent on committing atrocities would have an incentive to appear cooperative, as seen in the case of Libya, where the government indicated its willingness to comply with international requests to declare a ceasefire, whilst proceeding to act otherwise. On the other hand, if the host state's refusal is premised on a rejection of the grounds for making this request, such as questioning that mass atrocities are taking place in the country, as in the case of the Philippines, determining the host state's 'manifest failing' will not be straightforward. In such cases, international action may be precluded until the verity of reported grave human rights violations can be established.

In other scenarios, the host state could acquiescence to the international community's requests and pledge to address the threat of atrocities. In such cases the international community ought to determine, whether the state is able to do so and how much time would be appropriate to afford the host state to fulfil its promise or plan of action. As Chapter III explains in more detail in the context on the extraterritorial threat posed by non-state actors, the time that the host state should be given to take action in response to a request to address the threat of atrocities is not something that can be determined in advance, as it would depend on a number of context-specific factors, including the nature, gravity and especially the imminence of the threat of atrocities. It would also depend on an assessment of the host state's capability (Determinant 5 of this framework), as well as its progress towards the implementation of its plan to ensure the protection of its populations (Determinant 6 of this framework). Accordingly, identifying the determinations that should influence specifying a period of compliance with international requests provides sufficient guidance for conducting 'manifest failing' inquiries on a case-by-case, which is what this framework endeavours to do. Provided that the aforementioned assessments do not yield satisfactory results and that the host state proves ineffective in addressing the threat, the international community would have a substantive basis to take collective action limited to the protection of civilian populations from the threat of atrocity crimes.

Determinant 9: The host state's track record of grave human rights violations and relevant interaction with the international community

The investigation of the 'manifest failing' and 'unable or unwilling' thresholds throughout this thesis have demonstrated the relevance of the historical conduct of the host state for determining whether the threshold for collective action to prevent atrocity crimes has been met. As mentioned in Chapter I, Rosenberg and Strauss include 'a past history of violence against perceived groups' and 'a climate of impunity in which these events unfold' as indicators of the gravity of human rights violations, whereas Gallagher highlights that Milosevic's policy track record played a role in the decision to intervene in Kosovo as part of his discussion of host state intent. As evidenced in Chapter IV, Gaddafi's poor track record of atrocities and interactions with the international community was an important factor in deciding upon collective action in the context of the Libyan intervention. Deeks, on the other hand, has specified the victim state's 'prior interactors with the host state' as a separate factor in her normative framework for guiding

'unwilling or unable' inquiries. Given that the historical conduct of the host state has been brought up in relation to various other determinants of 'manifest failing' and that it does not seem to properly belong to any of them, this thesis has followed Deek's approach and included the host state's historical record of violence and interactions with the international community on matters relevant to atrocity crimes as a separate determinant.

In most basic terms, this determination combines relevant insights into how the historical conduct of host state can help to ascertain whether it is 'manifestly failing' in its 'responsibility to protect'. Specifically, it involves an assessment of the host state's history of gross human rights violations, with particular regards to cases in which it has acted as an agent of mass violence, and the way in which it has responded to previous requests to address serious violations of human rights made by the international community. In addition to specific assessment criteria identified in Chapter I, III and IV, the list of indicators attached to this determinant (see Table 5.1.) also draws inspiration from the indicators attached to Risk Factor 2 of the UN 'Framework of Analysis for Atrocity Crisis', including 'policy or practice of impunity for or tolerance of serious violations of international human rights and humanitarian law, of atrocity crimes, or of their incitement', 'past…serious restrictions to or violations of international human rights and humanitarian law' and 'past acts of genocide, crimes against humanity, war crimes or their incitement'.[1015]

Evidence of host state inaction in the face of atrocities in the past or that it has itself been the perpetrator of atrocities could raise serious concerns or suspicions that the host state may be 'manifestly failing' to protect its populations at the first signs of an atrocity crisis and alert relevant stakeholders to the possibility of an escalation of the threat of atrocity crimes. In this sense, the historical track record of states is particularly relevant to determining the prospective gravity of atrocities and the imminence of the threat of atrocities. This was seen in the case of Libya, some states suggested that the gravity of prospective atrocities can be inferred from Qaddafi's historical track record of violence. Thus, taking into consideration what Qaddafi is capable of influenced states' perceptions and portrayal of the imminence of the threat.

It is important to stress that, on the one hand, even if the host state has no history of atrocities, this should not preclude an equally demanding investigation into the human

[1015] UN, 'Framework of Analysis for Atrocity Crimes', p. 11.

rights situation in the country and assessment of the extent to which the host state is fulfilling its protection responsibilities. On the other hand, even if some of the indicators listed in Table 5.1 are manifested in a given case, no independent judgement about the host state's 'manifest failing' can be reached based on the historical conduct of the host state or its interactions with the international community in past cases of atrocities and mass violence alone. However, when considered in conjunction with other determinants, such considerations may lend support to judgements that the host state is 'manifestly failing' to protect its populations.

III. THE PHILIPPINES' 'WAR ON DRUGS'

There are several reasons to focus on the 'war on drugs' in the Philippines, in order to illustrate the application of the framework outlined above. First, as noted in the Introduction, there has been a divergence of opinions between ASEAN and 'the West' with regards to whether the 'war on drugs' should be interpreted through a mass atrocity lens. Relatedly, Gallagher et al. point to the concerning lack of evidence that either ASEAN or the P5 are considering the application of the R2P in this case.[1016] As the authors explain in detail, the P3 have 'prioritised bi-lateral trade relations and the counter-terrorism norm, whilst China has openly supported the war on drugs and, along with Russia, see any interference as a violation of state sovereignty'.[1017] Both the disparities in interpreting the situation in the Philippines and the vested interest of the P5 to turn a blind eye to serious allegations of human rights violations in the country make Duterte's 'war on drugs' a particularly apt case to illustrate the application of the framework detailed in this chapter, in order to determine whether there are sufficient grounds to view the government of the Philippines as 'manifestly failing' to protect its populations from the four crimes.

Second, this case has received little attention in the scholarly literature, despite the fact that as the first anti-drug campaign to become the focus of an ICC preliminary examination, it is a 'timely and important case study that raises questions regarding mass atrocity prevention in international relations'.[1018] Gallagher et al.'s analysis of this case sheds light on the Philippine government's 'manifestly failing' in its responsibility to

[1016] Gallagher et al., 'Failing to fulfil the responsibility to protect', p. 20.
[1017] Ibid.
[1018] Ibid., p. 3.

protect, but it remains the only contribution dedicated to this matter in a limited pool of scholarly contributions addressing the 'war on drugs' from a mass atrocity perspective.

Third, since the publication of their work two years ago (January 2019), worrying new evidence of human rights violations in the context of the government's anti-drug campaign has surfaced in the OHCHR's comprehensive report on the human rights situation in the Philippines of June 2020 and the findings from ICC's preliminary investigation into the situation published on 14 December 2020.[1019] This begs to revisit the case in light of this evidence and other recent developments concerning the willingness of the government of the Philippines to cooperate with international institutions on human rights matters.

Fourth, given that the situation in the Philippines is one that involves grave human rights violations committed by both state and non-state actors, it ties into the discussions in this thesis regarding the role of the state and non-state actors in the commission of atrocity crimes, with the latter depicted one of the key contemporary challenges for the R2P by Welsh (as discussed above).

Fifth, in light of the upcoming elections for a new president and vice president, as part a general election in 2022, this is an apt time to assess the extent to which the Philippine government is living up to its responsibility to protect its populations.

Basic context

Since President Duterte launched his nationwide campaign to eradicate illegal drugs in the Philippines in mid-2016, there have been allegations of the killing of thousands of people on account of their alleged involvement in drug use or trade.[1020] While some of these killings have reportedly occurred in the context of clashes between or within gangs, it is alleged that many of the reported incidents involved extra-judicial killings in the course of police anti-drug operations.[1021] An ICC preliminary examination of the situation in the Philippines was announced on 8 February 2018, 'analysing crimes allegedly committed since at least 1 July 2016, in the context of the "war on drugs" campaign

[1019] Human Rights Council, 'Report of the United Nations High Commissioner for Human Rights on the situation of human rights in the Philippines', Advance Edited Version, A/HRC/44/22, 4 June 2020.

[1020] Ibid.

[1021] International Criminal Court Website, 'Preliminary examination: Republic of the Philippines'.

launched by the Government of the Philippines' and is ongoing.[1022] The ICC published the results of its preliminary investigations on 14 December 2020, which examined reports that the Philippine National Police's (PNP) 'forces have reportedly conducted tens of thousands of operations to date, which have reportedly resulted in the killing of thousands of alleged drug users and/or small-scale dealers' in the context of President Duterte's anti-drug campaign.[1023] Specifically, the Court's focus was on:

> 'allegations that President Duterte and senior members of law enforcement agencies and other government bodies actively promoted and encouraged the killing of suspected or purported drug users and/or dealers, and in such context, members of law enforcement, including particularly the PNP, and unidentified assailants have carried out thousands of unlawful killings throughout the Philippines'.[1024]

On 11 July 2019, the UN Human Rights Council adopted a resolution sponsored by Iceland:

> 'expressing [the Council's] concern at the allegations of human rights violations in the Philippines, particularly those involving killings, enforced disappearances, arbitrary arrest and detention, the intimidation and persecution of or violence against members of civil society, human rights defenders, indigenous peoples, journalists, lawyers and members of the political opposition, and restrictions on the freedoms of opinion and expression, peaceful assembly and association'.[1025]

It also called for an investigation of Duterte's drug war and the deteriorating human rights situation in the country and requested that the OHCHR presents a report on the human rights situation in the Philippines.[1026] The sanctioning of this unprecedented 'resolution marked the first time the Philippines was the subject of such a measure from the Human

[1022] Ibid.

[1023] International Criminal Court – The Office of the Prosecutor (ICC-OTP), 'Report on Preliminary Examination Activities 2020', 14 December 2020, para. 181. [Accessed 23 March 2021]. Available at: https://www.icc-cpi.int/itemsDocuments/2020-PE/2020-pe-report-eng.pdf

[1024] Ibid., para. 183.

[1025] A/HRC/RES/41/2 (2019), p. 1.

[1026] Ibid., p. 3.

Rights Council'.[1027] The OHCHR report suggested that there is compelling evidence of the threat of atrocity crimes. Following its publication, the HRC passed another resolution in October 2020, which stated that the scrutiny of the human rights situation in the Philippines will continue, but did not launch an international investigation.[1028]

Human rights violations associated with atrocity crimes

Gallagher et al.'s analysis of the situation in the Philippines suggests that crimes against humanity were being committed in the country based on the role of the national government in 1) advocating the killing of drug users; 2) dehumanising drug users through harmful rhetoric which depicts them as the 'walking dead' in a bid to legitimise violence against them; and 3) exaggerating the threat posed by this group by 'portraying drug use as a disease' that poses a 'threat to the nation'.[1029] The UN and ICC reports on the situation in the Philippines published in the past year lend further support to this conclusion, on account of proof of human rights violations in the country that evidence the commission of crimes against humanity.

The OHCHR reports that despite 'important human rights gains in recent years, particularly in economic and social rights the underpinning focus on national security threats – real and inflated – has led to serious human rights violations, reinforced by harmful rhetoric from high-level officials'.[1030] The reports issued by the OHCHR and ICC, human rights groups, such as Human Rights Watch and Amnesty International, as well as a wide range of local, regional and international news outlets all point to the fact that grave human rights violations, which fall within the list of gross human rights violations / acts of atrocity listed the framework outlined in this chapter, including extrajudicial killings, arbitrary arrests, detention and the internal displacement of populations, are taking place in the Philippines.

In addition to killings as part of Duterte's anti-drug campaign and allegations of individuals being 'subjected to serious ill-treatment and abuses prior to being killed by

[1027] Human Rights Watch, 'Philippines: Events of 2019', (no date). [Accessed 19 March 2021]. Available at: https://www.hrw.org/world-report/2020/country-chapters/philippines#

[1028] United Nations General Assembly, A/HRC/RES/45/33, 13 October 2020.

[1029] Gallagher et. al, 'Failing to fulfil the responsibility to protect', p.10.

[1030] United Nations Human Rights Council, 'Report of the United Nations High Commissioner for Human Rights on the situation of human rights in the Philippines', Advance Edited Version, A/HRC/44/22, 04 June 2020, p. 3.

state actors and other unidentified assailants, such as after being arrested or abducted and while being held in custody prior their deaths', the ICC reports 'several incidents, [in which] relatives (such as spouses, parents or children) of the victims witnessed the killings, thereby sustaining *serious mental suffering*' (emphasis added).[1031] Moreover, the Court's investigation has revealed disturbing reports of incidents in which 'members of law enforcement raped women who were apparently targeted because of their personal relationships to individuals alleged to have been involved in drug activities'.[1032]

Based on the evidence it received, the Office of the Prosecutor of the International Criminal Court concluded that it:

> 'is satisfied that information available provides a reasonable basis to believe that the crimes against humanity of murder (article 7(1)(a)), torture (article 7(1)(f)) and the infliction of serious physical injury and mental harm as other inhumane Acts (article 7(1)(k)) were committed on the territory of the Philippines between at least 1 July 2016 and 16 March 2019, in connection to the WoD campaign launched throughout the country'.[1033]

While the ICC found that there is sufficient evidence to infer that grave human rights violations which amount to crimes against humanity were committed on the territory of the Philippines up until 16 March 2019, the report of OHCHR, which covers the period up until June 2020 and subsequent reports of the sustained incidence of the above-cited human rights violations suggest that the threat of international crimes in the country is ongoing. Evidence of this will be provided in the remainder of this chapter. The next two assessments will help to understand whether the threat of atrocities is severe and probable enough to raise concerns that the government of the Philippines is 'manifestly failing' to protect its populations.

The gravity of atrocity-related human rights violations

Compelling evidence of the gravity of reported atrocity-related human rights violations in the Philippines suggests that the national government is 'manifestly failing' in its responsibility to protect. While the Commission on Human Rights reports that verifying

[1031] ICC-OTP, 'Report on Preliminary Examination Activities 2020', para. 186.
[1032] Ibid.
[1033] Ibid. para. 189.

the death toll resulting from extrajudicial killings is not possible without further investigation, the information they reviewed in preparing their report suggests that 'the drug campaign-related killings appear to have a widespread and systematic character'.[1034] Specifically, [t]he most conservative figure, based on Government data, suggests that since July 2016, 8,663 people have been killed – with other estimates of up to triple that number'.[1035] The even higher number of internally displaced persons in Mindanao, amounting to 359,941 individuals as of 31 March 2020, also raises concerns of the government's failing to protect its populations.

In addition to the high death toll and number of internally displaced individuals, the number of arbitrary arrests in also considerable. 'House to house visitation to persuade suspects to stop illegal drug activities' have been a key feature of the government's campaign against illegal drugs, project 'Double Barrel', launched on President Duterte's first day of office with the publication of PNP Command Memorandum Circular No. 16-2016.[1036] According to OHCHR, this practice 'rais[es] important due process concerns as [it does not] require search or arrest warrants and could be conducted solely on the basis of a person's inclusion on a "drug watch list"', with those persons having no legal recourse to dispute it, and has 'systematically forced suspects to make self-incriminating statements or risk facing lethal force' .[1037] Police figures indicate that for the period of 1 July 2016 - 30 November 2017, out of 42,286 anti-illegal drugs operations carried out by the police only 1.2 % (507) were based on an arrest warrant.[1038] In addition, government data reveals 'that 223,780 "drug personalities" were arrested from 146 July 2016 to 31 December 2019'.[1039] While the government of the Philippines claims that '204,721 of these individuals were charged with criminal cases [it is unclear] how many may have been related to drug trade and how many to personal drug use, how many were convicted, released or remain in pre-trial detention'.[1040] According to OHCHR, this 'lack of clarity, coupled with due process irregularities, raises concerns that many of these cases may amount to arbitrary detentions'.[1041] The above figures show that the situation in the

[1034] A/HRC/44/22 (2020), para. 22.
[1035] Ibid.
[1036] Ibid., para. 16-17.
[1037] Ibid.
[1038] Ibid
[1039] Ibid., para. 32.
[1040] Ibid.
[1041] Ibid.

Philippines has crossed a quantitative threshold of severity that is indicative of the government's 'manifest failing'. What is more, there are a range of qualitative indicators of the gravity of human rights violations in the Philippines that also point to the inadequacy of the current government in fulfilling its responsibility to protect.

A particularly disturbing feature of the 'war on drugs' is the targeting of vulnerable groups. According to Human Rights Watch, '[t]he killings in the context of the 'drug war' are targeting the most vulnerable section of the Philippines' population – the urban poor, the most marginalised and voiceless people with least access to justice and redress'.[1042] In addition, there is evidence of the deliberate targeting of children as part of Duterte's anti-drug effort. The OHCHR 'documented the killing of 73 children in the context of the campaign against illegal drugs – 62 male and 11 female', including a five-month-old victim, which took place between 1 June 2016 and 21 April 2020.[1043] It should be borne in mind that these numbers may be larger, as the Commission warns that this is not an exhaustive list. The ICC also reports that allegedly, 'a significant number of minors (ranging in age from a few months old to 17 years old) were victims of apparent WoD-related killings, and in this respect, were killed in a number of circumstances, including as direct targets, as a result of mistaken of identity or as collateral victims'.[1044] The intentional targeting of children is a worrying aspect of the human rights situation in the Philippines, which is indicative of a qualitative leap in violence and a key indicator of the 'manifest failing' of Duterte's government.

The imminence of the threat of atrocities

The APR2P's assessment of the situation of February 2021 attests that the Philippines 'remains at very high risk for atrocities as the Duterte government's drug war campaign continued unabated as the country battles rising COVID-19 cases'.[1045] The Asia-Pacific Centre notes that in spite of the imposition of strict lockdown measures to prevent the spread of COVID-19 since March 2020, 'anti-drug operations by the Philippine Drug Enforcement Agency (PDEA) and the PNP continued, resulting in the arrests of over 10,000 suspects and 67 deaths between 31 March and 31 May [2020] and seizure of illegal

[1042] Rod Austin, '"War on drugs" makes Philippines fourth most dangerous country – report', 9 July 2019. [Accessed 24 March 2021]. Available at: https://www.theguardian.com/global-development/2019/jul/09/war-on-drugs-makes-philippines-fourth-most-dangerous-country-report

[1043] Ibid.

[1044] ICC-OTP, 'Report on Preliminary Examination Activities 2020', para. 188.

[1045] APR2P, 'The Philippines: Asia Pacific Regional Outlook February 2021'.

drugs worth over USD14 million'.[1046] What is more, there has been a notable escalation in drug war-related deaths since the start of the COVID-19 pandemic, as the PNP 'reported that the death toll from its anti-drug war operations climbed to 7,987 as of end of October, with over 100 people killed in just two months', while the Chief of the PNP reported 'a total of 9,240 anti-drug operations were conducted by the police that resulted in 119 deaths' in September to October alone.[1047] The PNP also recorded a death toll of 2,423 in the period since the imposition of COVID-19 restrictions from 17 March 2020 to September 2020, averaging 13 killings per day.[1048] Notably, Human Rights Watch reported that based on government statistics 'drug war deaths in the country rose dramatically by 50 percent during the pandemic' from April to July 2020.[1049] Moreover, it stated that unwarranted arrests have also been on the rise during the pandemic.[1050]

In February 2021, the APR2P also noted an increase in human rights violations, as the government of the Philippines 'used the Anti-Terrorism Act (ATA) [signed into law on 3 July 2020] to crackdown on suspected communist or leftist sympathisers, human rights defenders, indigenous people, media reporters, and other critics of the Duterte administration'.[1051] In March 2021, Duterte issued clear orders to 'shoot and kill' armed communist rebels.[1052] The subsequent death of nine activists in the first week of March in 'raids against suspected insurgents' prompted the UN to condemn the killings and 'urg[e] authorities to avoid rhetoric that could lead to human rights violations'.[1053] On

[1046] Ibid.

[1047] Ibid.

[1048] Emmanuel Tupas, '13 killed daily during quarantine', 13 September 2020. [Accessed 24 March 2020]. Available at: https://www.philstar.com/nation/2020/09/13/2042020/13-killed-daily-during-quarantine

[1049] Carlos H. Conde, 'Killings in Philippines Up 50 Percent During Pandemic: Drug War' Deaths Rise Dramatically as Country Reels from Covid-19', 8 September 2020 [Accessed 21 March 2021] Available at: https://www.hrw.org/news/2020/09/08/killings-philippines-50-percent-during-pandemic Phil Robertson, 'Another Spike in Philippines' "Drug War" Deaths', [Accessed 21 March 2021] 28 September 2020. Available at: https://www.hrw.org/news/2020/09/28/another-spike-philippines-drug-war-deaths

[1050] Human Rights Watch News Release, 'Philippines: "Drug War" Killings Rise During Pandemic', 13 January 2021. [24 March 2021]. Available at: https://www.hrw.org/news/2021/01/13/philippines-drug-war-killings-rise-during-pandemic

[1051] APR2P, 'The Philippines: Asia Pacific Regional Outlook February 2021'. Human Rights Watch news release, 'Philippines: New Anti-Terrorism Act Endangers Rights', 5 June 2020 [Accessed 24 March] Available at: https://www.hrw.org/news/2020/06/05/philippines-new-anti-terrorism-act-endangers-rights

[1052] Catherine S. Valente, 'The Manila Times', 6 March 2021 [Accessed 24 March 2021] Available at: https://www.manilatimes.net/2021/03/06/news/duterte-orders-military-to-shoot-and-kill-armed-communist-rebels/847906/

[1053] Reuters staff, 'U.N. 'deeply worried' over Philippine killings, violent rhetoric', 9 March 2021 [Accessed 24 March 2021] Available at: https://www.reuters.com/article/us-philippines-rights-un/u-n-deeply-worried-over-philippine-killings-violent-rhetoric-idUSKBN2B11JG?il=0

this occasion, UN human rights spokeswoman Ravina Shamdasani told a news briefing in Geneva: 'We are deeply worried that these latest killings indicate an escalation of violence, intimidation harassment and "red tagging" of human rights defenders', which 'refers to government and military efforts to label anyone critical of human rights abuses or fighting for labour rights as a communist'.[1054] Hence, in addition to the drastic escalation of atrocity-related human rights violations over the past year, we have also seen a recent spike in incitement to violence, which has led to more deaths.

In sum, since July 2016, there has been a continuous pattern of violence, incitement to violence and mass human rights violations that amount to atrocity crimes in the Philippines. In light of the gravity of human rights violations in the Philippines, evidence of a continuous pattern of widespread and systematic serious human rights violations, ongoing violence, a recent escalation in violence, recent examples of incitement to violence and assessments that the situation is at very high risk of atrocities, the probability of prospective grave human rights violations / atrocity crimes in the Philippines is very high, i.e. *imminent*. Thus, the evidence of the human rights violations in the country, the gravity of these violations and the ongoing character of the serious threat of atrocities suggest that Duterte's government is 'manifestly failing' to protect its populations and that the state itself bears responsibility for this threat.

Government intent

The Philippines' President has publicly incited violence against drug suspects since his election campaign, as he 'encouraged citizens to kill suspected drugs dealers or users as a "duty", and offered huge bounties to people who turn in drug dealers – "dead or alive"', in a 'speech broadcast on national television on 5 June [2016]'.[1055] Following his election, on 30 June 2016, he urged 'If you know of any addicts, go ahead and kill them yourself', during his inauguration speech.[1056] Even though the government of the Philippines, has repudiated claims that it upholds 'a policy to kill people who use drugs and states that all deaths occur during legitimate police operations', evidence of incitement to violence and

[1054] Ibid.; Al Jazeera, '"Appalled": UN urges probe into killing of Philippine activists', 10 march 2021. [Accessed 24 March 2021]. Available at: https://www.aljazeera.com/news/2021/3/10/un-urges-probe-into-killings-of-philippine-activists

[1055] Amnesty International UK, 'More than 7,000 killed in the Philippines in six months, as president encourages murder', 18 May 2020, Available at: https://www.amnesty.org.uk/philippines-president-duterte-war-on-drugs-thousands-killed

[1056] Ibid.

hatred against this group, highlighted by the UN, points to the contrary.[1057] The UN's comprehensive report of June 2020 found that 'harmful rhetoric from the highest levels of the Government has been pervasive and deeply damaging' in recent years.[1058] According to the OHCHR, 'some statements have risen to the level of incitement to violence'.[1059] The report summarises that:

> '[t]he rhetoric has ranged from degrading and sexually-charged comments against women human rights defenders, politicians and combatants – including rape "jokes" – to statements making light of torture, calling for bombing of indigenous peoples, encouraging extreme violence against drug users and peddlers – even offering bounties, calling for beheadings of civil society actors, and warning that journalists were not immune from "assassination"'.[1060]

Moreover, President Duterte did not hesitate to reassert his 'war on drugs' threats days after the publication of the UN report. On 5 June 2020, Al Jazeera reported that the Philippine President referred to drugs in a recorded address, stating 'If you destroy my country distributing 5.1 billion pesos worth of shabu … I will kill you'.[1061] In addition to championing the murder of drug users, Duterte has urged the killing of 'leftists' and 'even violators of COVID-19 quarantine or curfew orders'.[1062] In September 2020, Amnesty International asserted that Philippine government 'continues to incite a wave of extrajudicial executions and fuel a climate of near-absolute impunity for perpetrators', with the incitement intensifying, 'despite increased international scrutiny by the UN'.[1063] This is evidenced by Duterte's recent unequivocal declaration 'It is my job to scare people, to intimidate people, and to kill people'.[1064] Such public statements constitute

[1057] A/HRC/44/22, para. 19.

[1058] Ibid., para. 77.

[1059] Ibid.

[1060] Ibid. para. 78.

[1061] Al Jazeera, '"I will kill you": Philippines' Duterte renews drug war threat', 5 June 2020 [Accessed 21 March 2021] Available at: https://www.aljazeera.com/news/2020/6/5/i-will-kill-you-philippines-duterte-renews-drug-war-threat

[1062] Ibid.

[1063] Amnesty International, 'Philippines: UN must intensify pressure to end killings as impunity reigns', 25 September 2020. [Accessed 21 March 2021] Available at: https://www.amnesty.org/en/latest/news/2020/09/philippines-un-pressure-end-killings/

[1064] Ibid.

clear evidence of government intent, which reaffirms the role of the state in sanctioning and inciting grave ongoing human rights violation.

What is more, Philippine officials have routinely maintained that the thousands of deaths that occurred in acknowledged Philippine law enforcement-led anti-illegal drugs operations were 'a result of officers acting legitimately in self-defence in the context of violent, armed confrontations with suspects'.[1065] However, as the ICC remarks 'such narrative has been challenged by others, who have contended that the use of lethal force was unnecessary and disproportionate under the circumstances, as to render the resulting killings essentially arbitrary, or extrajudicial, executions'.[1066] Moreover, the Philippines' authorities have claimed that 'thousands of killings', which have been attributed 'to unidentified assailants (sometimes referred to as "vigilantes" or "unknown gunmen")', were not connected to the 'war on drugs', but took place 'in the context of love triangles or, alternatively, feuds or rivalries between drug gangs and criminal organisations'.[1067] On the contrary, the ICC finds that a number of these vigilante-style executions occurred precisely in the context of the government's anti-illegal drugs campaign or in relation to it, with information that some of these killings, allegedly carried out 'by private citizens or groups [,] were planned, directed and/or coordinated by members of the PNP, and/or were actually committed by members of law enforcement who concealed their identity and took measures to make the killings appear to have instead been perpetrated by vigilantes'.[1068] Moreover, 'submissions received by OHCHR…describe disturbing familiarity of masked perpetrators with locations and victims, suggesting possible collusion with police and local government officials in some cases'.[1069] The available information on the human rights situation in the Philippines gathered by the ICC-OTP and OHCHR suggests that there is compelling evidence of systematic perpetration, indicative of the Philippine government's intent. As noted in the above discussion of government intent, the combination of public statements of intent and systematic preparation constitutes strong evidence of host state intent, as the latter has not only declared its wish to eliminate a specific group, but has taken observable steps to execute this plan employing the resources at their command, despite efforts to conceal and deny

[1065] ICC-OTP, 'Report on Preliminary Examination Activities 2020', para. 184.
[1066] Ibid.
[1067] Ibid. para. 185.
[1068] Ibid.
[1069] A/HRC/44/22, para. 30.

their direct involvement. As the previous section elucidated, Duterte's continuous threats, incitement to violence and hatred against alleged drug users and dealers have shown that he has no intention of slowing down his campaign of mass murder waged against his own populations.

The Philippines' territorial control and capacity

The Philippines ranks in the bottom 54 countries on the Fragile State Index, positioning it among states with a high warning of state fragility, but above states falling in the categories of 'alert', 'high alert' and 'very high alert'. This is not surprising, given that in addition to the government's 'war on drugs', there are non-state actors, such as 'combatants from communist and Moro rebel groups, ISIS-affiliated militants, and political clans who rule with impunity and use political violence in Mindanao', which are also known to have committed atrocity crimes against civilians.[1070] In this regard, the OHCHR reports that 'there are serious concerns that the persistent lack of security and economic development in Mindanao, including insufficient progress in rebuilding Marawi following the siege in 2017, reported violations of international humanitarian law, and the lack of progress in transitional justice and reconciliation provide fertile ground for radicalization'.[1071] This is evidenced by the number of internally displaced (many of which protractedly displaced) people in Mindanao (359, 941), as a result of armed conflict, natural disasters and the 2017 siege of Marawi.[1072] Whereas no hostilities have been reported between MILF (Moro Islamic Liberation Front) and AFP (Armed Forces of the Philippines) since the establishment of BARMM (Bangsamoro Autonomous Region in Muslim Mindanao) in 2018, armed clashes are ongoing in other parts of Mindanao with ISIL-affiliated non-state armed actors: the Abu Sayyaf Group, the Bangsamoro Islamic Freedom Fighters and the Maute group.[1073] Ongoing clashes between AFP and NPA (New People's Army), as well as the activities of other armed groups, continue to cause internal displacements in eastern and northern Mindanao.[1074]

[1070] Asia-Pacific Centre for the Responsibility to Protect, 'Atrocity Crimes Risk Assessment Series: The Philippines', Volume 7, September 2018 [Accessed 24 March 2021] Available at: https://r2pasiapacific.org/files/2497/Risk%20Assessment%20The%20Philippines%20FINALwith%20images%20opti%281%29.pdf

[1071] A/HRC/44/22, para. 62.

[1072] Ibid., para 61.

[1073] Ibid., para. 62.

[1074] Ibid., para. 63.

Nonetheless, a recent study into the violence perpetrated by non-state actors in the context of the COVID-19 pandemic has revealed that 'compared to the previous four years, the Southern Philippines as a whole experienced significantly less violence in 2020 in terms of both the number of violent incidents and the number of casualties resulting from those incidents'.[1075] Besides, on the whole, at present the situation in the Philippines does not suggest that the host state is unable to address the threat posed by non-state actors on account of a lack of territorial control or capacity, as it has received considerable support from the international community, especially the US' Trump administration, in combatting terrorism. While Gallagher et al. point out that '[i]f the international community… refused all international assistance [making it conditional upon the Philippines addressing human rights issues in the country] then this could leave the government without the capacity to protect its citizens from the threat posed by nonstate armed groups'.[1076] In the unlikely event that such a scenario unfolds, the Philippines may be rendered unable to protect innocent populations in parts of its territory, which would make it an 'unable' state. However, as long as the government of the Philippines continues to act as an agent of mass atrocities, it would still be 'manifestly failing' to protect its populations, regardless of whether it is just 'unwilling' or both 'unwilling' to stop committing atrocities against its people and 'unable' to meet the threat of non-state actors committing atrocities in other parts of the country.

As Gallagher et al. elucidate, 'the Philippines is a complex case in that non-state armed groups are perpetrating mass violence in certain parts of the country whilst the government is directing a policy of extrajudicial killings in other areas'.[1077] However, while the role of non-state actors in the commission of atrocities and the risks they pose to vulnerable populations should not be understated, ongoing grave human rights violations are primarily carried out by state law enforcement agencies, as the evidenced in the above four sections suggests. This renders the Philippines an 'unwilling' rather than 'unable' state, which means that the current situation in the Philippines constitutes a clear case in which the host state is 'manifestly failing' to protect its populations on account of

[1075] Luke Lischin, 'Surviving or thriving? COVID-19 and violent non-state actors in the Southern Philippines', 9 September 2020 [Accessed 5 February 2021] Available at: https://www.newmandala.org/surviving-or-thriving-covid-19-and-violent-non-state-actors-in-the-southern-philippines/

[1076] Ibid.

[1077] Gallagher et al., 'Failing to fulfil the responsibility to protect', p. 13

being primarily responsible for orchestrating the threat of mass atrocities looming over its populations.

The measures taken by the Philippines to mitigate the risk of atrocities

The evidence of the gravity and imminence of the threat of atrocities in the Philippines, especially the escalation of violence over the last year, is perhaps the most compelling evidence that Duterte's government has not taken adequate measures to mitigate the risk of atrocities. This is reflected in the most recent APR2P report on the situation in the Philippines, which identifies the need for the government to:

> 'Uphold its primary responsibility to protect by complying with international norms on human rights protection and humanitarian law [by taking steps to] hold accountable law enforcers and other members of the security sector for violations of human rights in relation to the war on illegal drugs and campaign against terrorism'. [1078]

As the findings of OHCHR's comprehensive report indicate, in spite of 'credible allegations of widespread and systematic extrajudicial killings [as part of the government's drug war], there has been near impunity for such violations'.[1079] A testimony to this is the failing of the PNP's Human Rights Affairs Office (HRA), established 'in response to allegations of extrajudicial killings' over a decade ago, which has since been criticised 'for its lack of power and effort – often limited to internal human rights campaigns only – to keep up with the breakneck speed with which allegations of rights violations arise in the Duterte administration'.[1080] Furthermore, out of 4, 583 investigations into deaths that took place over the course of police operations between July 2016 and May 2019, reportedly launched by the Police Internal Affairs Service, the government of the Philippines 'has cited only one case – that of 17-year-old Kian delos Santos – where three police officers were convicted of a drug campaign-related killing', largely owing to CCTV footage.[1081] The evidence collected by OHCHR 'also suggest widespread impunity for drug-related killings committed by unidentified persons' and

[1078] APR2P, 'The Philippines: Asia Pacific Regional Outlook February 2021'.

[1079] A/HRC/44/22, para. 26.

[1080] Rambo Talabong and Jodesz Gavilan, 'PNP has a human rights office, but what has it done?', 26 November 2020 [Accessed 24 March 2021] Available at: https://www.rappler.com/newsbreak/in-depth/philippine-national-police-human-rights-affairs-office-what-has-it-done

[1081] A/HRC/44/22, para. 26.

their links to the government and law enforcement.[1082] Moreover, OHCHR's interviews with (mostly female) victim's relatives, lawyers and journalists, evidenced a number of hindrances to pursuing justice and documenting cases, including:

> 'surveillance, harassment, threats, intimidation, lack of education, lack of protection of witnesses and victims, a feeling of powerlessness in the face of official statements encouraging killings, unwillingness by law enforcement to investigate, and reluctance by judges to critically examine drug-related cases'.[1083]

In addition, victims have been prevented from pursuing justice, and law enforcement agencies from investigating killings, on account of Duterte's public statements assuring the 'protect[ion of] police officers even if they killed 1,000 persons while on duty'.[1084]

Commenting on OHCHR's comprehensive report, the Deputy Director of Human Rights Watch's Asian division expressed doubts that 'national mechanisms will hold anyone responsible for the carnage of the drug war that has killed thousands of Filipinos', in light of Duterte's recent incitement to violence against drug users and other groups.[1085] These cases speak to the failings of the criminal justice system in the Philippines and government inaction in the face of allegations that it bears responsibility for the commission of mass atrocity crimes.

In sum, the Philippine government's response to international censure on the basis of evidence that crimes against humanity are taking place in the country is best defined by what the government has failed to do. When one considers the indicators listed under Determinant 6 of the framework proposed in this chapter, it is clear that the Philippines has taken none of the relevant measures specified therein that it could have conceivably taken to address the threat of atrocities. Namely, it has not publicly recognised and condemned relevant violation of human rights, issued unambiguous and public orders to its police and security forces to respect international humanitarian and human rights law, conducted thorough investigations into reported human rights violations or enforced accountability for relevant violations of human rights by arresting and/or prosecuting the

[1082] Ibid., para. 30.

[1083] Ibid., para. 29.

[1084] Ibid.

[1085] Al Jazeera, 'I will kill you'.

authors of those attacks and it has not formulated an action plan with timelines for implementing specific measures to mitigate the risk of atrocities, in cooperation with relevant stakeholders (including affected populations and prospective victims of atrocities). Given that the government is mainly responsible for the atrocities taking place in the country and their incitement, it is clear that while it could have conceivably employed any of the measures cited above, some of which align with concrete international demands (discussed below), it remains intent on continuing its campaign of mass murder on the premise of its 'war on drugs'. In this regard, Duterte's government has demonstrated the textbook conduct of a 'manifestly failing' host state.

The Philippines' willingness to cooperate with international organisations

The first of the two significant actions taken by the international community with regards to the situation in the Philippines – the ICC's announcement of a preliminary inquiry into the Philippine government's anti-drug campaign, was met with resounding scorn. Shortly after the ICC's statement, President Duterte issued clear orders for his officers *not to cooperate* in his characteristic oratorical style, which earned him popularity with voters at home: '"You're investigating us? Fact finding? Sorry, do not f*** with me… Who are you to interfere in the way I would run my country? You know very well that we are being swallowed by drugs'".[1086] In a surprising retaliatory move, he also announced his plans to withdraw from the ICC on 14 March 2018 'despite repeatedly daring [the court] to charge him over his war on drugs and saying he would cooperate with its investigation'.[1087] On 17 March 2019, the Philippines officially withdrew from the ICC, while the Court continues its investigation.[1088] A lawyer for a human rights activists coalition 'who had asked the Supreme Court of the Philippines for an injunction against the move', Romel Bagares, saw it as 'a terrible setback in the long fight against impunity

[1086] Joshua Berlinger, 'Philippines: UN decision to probe drug war "straight from the mouth of the Queen in Alice in Wonderland"', 12 July 2019. [Accessed 3 February 2021]. Available at: https://edition.cnn.com/2019/07/11/asia/philippines-drug-war-un-intl-hnk/index.html#:~:text=The%20Philippines%20Secretary%20of%20Foreign%20Affairs,%20Teodoro%20L.,in%202006%20by%20a%20UN%20general%20assembly%20resolution.

[1087] The Independent, 'Duterte withdraws Philippines from International Criminal Court in stunning U-turn', 14 March 2018 [Accessed 21 March 2021]. Available at: https://www.independent.co.uk/news/world/asia/duterte-philippines-icc-withdraw-war-drugs-international-criminal-court-extra-judicial-killings-a8256076.html ; Felipe Villamor, 'Philippines Plans to Withdraw From International Criminal Court', 14 March 2018. [Accessed 21 March 2021]. Available at: https://www.nytimes.com/2018/03/14/world/asia/rodrigo-duterte-philippines-icc.html

[1088] Jason Gutierrez, 'Philippines Officially Leaves the International Criminal Court', 17 March 2019. [Accessed 21 March 2021]. Available at https://www.nytimes.com/2019/03/17/world/asia/philippines-international-criminal-court.html

in the country', in a climate in which Philippine institutions 'have grievously been failing in the last two years, with apparent government inaction on thousands of deaths arising from the president's drug war'.[1089] Shortly before the departure of the Philippines took effect, the APR2P suggested that '[t]he withdrawal of The Philippines from the International Criminal Court means risks of further atrocity crimes occurring, and the impunity with which they will be carried out, will most likely increase'.[1090]

The second notable step taken by the international community with regards to the situation in the Philippines was the adoption of a UN Human Rights Council resolution, which explicitly:

> 'Call[ed] upon the Government of the Philippines *to cooperate* with the Office of the United Nations High Commissioner for Human Rights and the mechanisms of the Human Rights Council, including by facilitating country visits and preventing and refraining from all acts of intimidation or retaliation' (emphasis added).[1091]

The request of the Council was initially met with what appeared to be a less combative stance by President Duterte. Speaking to the press on 11 July 2019, he stated that 'he may consider allowing the United Nations Human Rights Council to investigate the human rights situation in the country' ('Let them state their purpose and I will review'), but also added that 'Kung dagdag lang sila sa intriga [If they'll just add to the intrigue], they better go to the media. And the media will tell them the truth'.[1092]

Despite this ambiguous response, the Philippine government's opposition to action by the international community was evident from the start. To begin with, it made considerable efforts to block the adoption of the abovementioned HRC resolution calling for its cooperation, 'including through an extensive misinformation campaign and wide-reaching diplomatic pressure'.[1093] Duterte's government maintained this course of action

---

[1089] Gutierrez, 'Philippines Officially Leaves the International Criminal Court'.

[1090] Asia-Pacific Centre for the Responsibility to Protect, 'The Philippines Baseline Assessment of R2P Implementation', October 2019 [Accessed 24 March 2021] Available at: https://r2pasiapacific.org/files/4214/Philippines_annex_assessment.pdf

[1091] A/HRC/RES/41/2 (2019), p. 2.

[1092] CNN Philippines Staff, 'Duterte on UN rights probe: State your purpose first', 12 July 2019 [Accessed 21 March 2021] Available at: https://www.cnnphilippines.com/news/2019/7/12/Rodrigo-Duterte-UN-human-rights-probe.html,.

[1093] Human Rights Watch, 'Philippines: Events of 2019'.

in the immediate aftermath of the adoption of the resolution. In late August, it issued a memorandum ordering agencies of the Philippine government to spurn financial assistance from the 18 countries, which voted in favour of the Council's resolution.[1094] The memorandum 'suspended all new talks and deals for foreign loans and grants' from these countries, due to 'the administration's strong rejection of the resolution of the UN Human Rights Council'.[1095]

While the OHCHR reported in June 2020 that it had 'had several exchanges with representatives of the Government of the Philippines, including detailed discussions on 13 and 14 February 2020 in Bangkok, Thailand' and expressed its gratitude for the government's openness to dialogue, this appeared to be the limit of the Philippines' willingness to cooperate, as it ultimately refused to grant OHCHR permission to conduct a visit to the country.[1096] Likewise, despite the Philippine government's readiness to engage in talks and technical cooperation with the HRC, acknowledged in the latter's resolution of 7 October 2020,[1097] the reaction of Duterte's administration to steps taken by the international community and actions to undermine its efforts to establish the facts about the human rights situation in the country evidence its demonstrative unwillingness to meaningfully cooperate with the international community. This provides further evidence of the Philippines' 'manifest failing' to uphold its primary protection responsibilities towards its populations.

The Philippines' response to international requests to address the threat

In addition to its unfavourable response to international requests for cooperation, the Philippines has failed to comply with other concrete demands made in the HRC resolution adopted on 11 July 2019. Specifically, the Council:

> 'Urg[ed] the Government of the Philippines to take all necessary measures to prevent extrajudicial killings and enforced disappearances, to carry out impartial investigations and to hold perpetrators accountable, in accordance

[1094] Ben O. de Vera, 'Duterte order shuns all loans, grants, aid from 18 countries backing probe of PH killings', 20 September 2019 [Accessed 21 March 2021] Available at: https://business.inquirer.net/279344/duterte-order-shuns-all-loans-grants-aid-from-18-countries-in-favor-of-probe-of-ph-killings

[1095] Ibid.

[1096] A/HRC/44/22, 29 June 2020, para. 4.

[1097] A/HRC/RES/45/33, 13 October 2020.

> with international norms and standards, including on due process and the rule of law'.[1098]

As the above discussions of the imminence of the threat of atrocities and host state intent suggest, the Philippine government has not only failed to take steps to put an end to extrajudicial killings, but was found responsible for grave human rights violations which have continued to take place in the country since July 2019 and have even escalated in the past year.

Moreover, the government of the Philippines' lack of compliance with the demands of the international community, in response to the publication of the OHCHR report in June 2020, was starkly exposed in the beginning of this year. In June 2020, the Philippines' Justice Secretary Menardo Guevarra announced the creation of a review panel set to look into more than 5,000 police 'drug war' operations that resulted in deaths, while assuring the UN that the Philippine Commission on Human Rights (CHR) 'will be involved as an independent monitoring body', and that its 'continued, unhampered function underpins our strong position against calls for an independent investigative mechanism'.[1099] However, on 11 January 2021, CHR Commissioner Karen Gomez Dumpit stated:

> 'We regret that the [CHR] was not involved in the review, contrary to the commitments and assurances made by the government during the 44th session of the [United Nations] Human Rights Council. This is an unfulfilled promise to Filipinos and the entire community of nations'.[1100]

Ultimately, a report resulting from the review of anti-drug operations was produced following a three-day summit held December 2020 led by the Philippines' Department of Justice under the theme 'Peace is the Work of Justice', which was one of the projects in a joint program on technical cooperation between the Philippine government and the United Nations pursuant to the resolution adopted by the HRC in June 2020. Following his attendance at the summit, the president of the International Association of Democratic Lawyers, Jeanne Mirer, stated that the exclusion of the CHR from the review once again

[1098] A/HRC/RES/41/2 (2019), p. 2.

[1099] Jodesz Gavilan, 'Despite UN promise, PH gov't excludes CHR from 'partial' drug war probe', 11 January 2021 [Accessed 24 March 2021] Available at: https://www.rappler.com/nation/chr-excluded-drug-war-review-panel-department-justice

[1100] Jodesz Gavilan, 'CHR still out of the loop in much hyped DOJ drug war panel' 8 October 2020, [Accessed 24 March 2021] Available at: https://www.rappler.com/nation/commission-human-rights-still-looped-out-doj-panel-drug-war-killings

evidenced that Duterte's government had 'no interest in investigating themselves and finding themselves guilty for thousands of extra-judicial killings'.[1101] In light of the aforementioned information, it is apparent that the government of the Philippines has neither responded favourably to international requests, nor made an effort to comply with them or fulfil the commitments it has made in order to avoid increased scrutiny from the international community. All of this is suggestive of the 'manifest failing' of Duterte's government to uphold its protection responsibilities towards its populations.

The Philippines' track record of violence and gross human rights violations

The post-colonial history of the Philippines, which became independent in 1946, was marked by a record of serious human rights violations. The rule of President Ferdinand Marcos, from 1965 to 1986, was marked by the imposition of martial law in 1972 and subsequent grave human rights violations, including 3,257 extrajudicial killings and the torture of 35, 000 people.[1102] Drawing on Singman, Gallagher et al. clarify that 'Duterte idolizes the Marcos authoritarian style' and has 'downplayed the atrocities' committed by his regime.[1103] Based on these facts and impressions, the authors conclude that 'a track record in which the government of the Philippines is directing the perpetration of mass human rights violations' can be observed.[1104] Significantly, Gallagher et al.'s historical analysis of mass human rights violations in the country suggests that extrajudicial killings in the Philippines are not a novel phenomenon (as of June 2016 when President Duterte took office), given that 'such violence has endured, albeit at lower levels, since the period of martial law ended in 1986', which suggests that politics in the country is permeated by 'a culture of violence and impunity'.[1105] The Philippines' history of violence and the nation's warm reception of the current president create a climate of impunity in which present grave human rights violations are taking place and support the inferences made so far in relation to other determinants of the framework outlined in this chapter, namely that Duterte's government is 'manifestly failing' to protect its populations.

[1101] Kristine Joy Patag, 'Duterte government has no interest in probing themselves — global group', 28 January 2021 [Accessed 21 March 2021].s a Available at: https://www.philstar.com/headlines/2021/01/28/2073673/duterte-government-has-no-interest-probing-themselves-global-group

[1102] A/HRC/44/22, para. 5.; Gallagher et al., 'Failing to fulfil the responsibility to protect', p. 7.

[1103] Gallagher et al., 'Failing to fulfil the responsibility to protect', p. 7.

[1104] Ibid.

[1105] Ibid.

The Philippines' 'manifest failing'

Overall, the application of the framework proposed in this chapter to the Duterte's 'war on drugs' suggest that the government of the Philippines is 'manifest failing' to protect its populations. This is evidenced by 1) the continuous presence of grave human rights violations indicative of crimes against humanity; 2) the gravity of these human rights violations; and 3) the imminence of the threat of atrocities in the country, in light of evidence of their continuity and escalation; as well as by 4) Duterte's continuous threats, inciting violence and hatred against drug users and evidence of systematic preparations to commit atrocities; 5) the host state's possession of sufficient territorial control and capacity to protect its populations from the threat posed by non-state actors and state law enforcement agencies' primary responsibility for ongoing grave human rights violations in the country; 6) the lack of measurable steps taken by the government to mitigate the ongoing threat of atrocities; 7) the lack of cooperation with and even hostility towards the international community and its attempts to investigate alleged atrocities taking place in the country; 8) the Philippine government's lack of compliance with the demands made by the international community; and 9) the historical record of mass violence and a climate of impunity in which atrocities currently unfold.

Despite the fact that there is compelling evidence of the Philippines' abject failing to protect its populations, the international response to the threat of atrocities in the country remains weak. In the aftermath of the OHCHR comprehensive report, 62 human rights groups asked the HRC 'to establish an independent body to investigate the killings and other human rights violations' in the country.[1106] This request was followed by a call from UN human rights chief Michelle Bachelet highlighting the 'urgent need to revoke the policies that continue to result in killings and other human rights violations, to bring to justice the perpetrators, and to halt the use of rhetoric inciting violence against people who use or sell drugs'.[1107] Hence, the most recent HRC resolution on the situation of October 2020, which included provisions for supporting the Philippines through

[1106] Eimor Santos, '62 groups ask UN for independent probe on PH killings, human rights abuses', 28 August 2020. [Accessed 24 March 2021]. Available at: https://www.cnnphilippines.com/news/2020/8/28/philippines-drug-war-united-nations-human-rights-council.html

[1107] CNN Philippines Staff, 'UN rights chief: Revoke PH policies, rhetoric that lead to killings and abuses', 16 September 2020. [Accessed 24 March 2020]. Available at: https://www.cnnphilippines.com/news/2020/9/16/united-nations-bachelet-philippines-duterte-drug-war.html

'technical assistance', has fallen short of the expectations of civil society and UN experts for launching an independent investigation to probe human rights violations in the context of the 'war on drugs' and has been labelled as a 'missed opportunity to seek justice for thousands of unlawful killings'.[1108] The lack of decisive action can be attributed to the fact that the stance of the P5 remains unchanged. As Gallagher et al.' analysis has shown, even though the P3 'have expressed public concerns they [continue to] prioritise counter-terrorism and trade over the Responsibility to Protect [, while] China and Russia uphold the view that the war on drugs is a matter of domestic jurisdiction'.[1109] While Philippine political analysts look to the US in expectation that President Joe Biden would take a tougher stance on human rights issues than his predecessor (who even applauded Duterte for the 'unbelievable job on the drug problem'), it is not clear whether the US will take decisive steps to put pressure on the Philippines in relation to the 'war on drugs'.[1110]

Table 5.2. The Philippines' 'Manifest Failing' in the context of Duterte's 'war on drugs'

DETERMINANTS	INDICATORS
1. Relevant human rights violations / acts of atrocity crimes	Killing, murder, torture, extra-judicial executions, arbitrary arrest and detention, rape and sexual violence, abduction, displacement, lack of access to justice and redress.
2. The gravity of the human rights violations taking place	Cumulative death toll of 8, 663 according to low estimates, with other estimates suggesting up to triple that number. 359, 941 internally displaced individuals. 223, 780 'drug personalities' arrested from 14 July 2016 to 31 December 2019 according to government date. Targeting vulnerable groups, including children, the urban poor, the most marginalised and voiceless people with least access to justice and redress.

[1108] Amnesty International, 'Philippines: UN resolution a missed chance for justice but scrutiny continues', 7 October 2020. [Accessed 5 February 2020]. Available at: https://www.amnesty.org/en/latest/news/2020/10/philippines-un-resolution-a-missed-chance-for-justice-but-scrutiny-continues/

[1109] Gallagher et al., 'Failing to fulfil the responsibility to protect', p. 1.

[1110] Reuters Staff, 'Philippines sees "close and friendly" ties with Biden administration' 20 January 2021 [Accessed 24 March 2021], Available at: https://www.reuters.com/article/us-philippines-usa-idUSKBN29P0YN; Gallagher et al., 'Failing to fulfil the responsibility to protect', p. 12.

3. The imminence of the threat of atrocities	There has been a continuous pattern of violence, incitement to violence and mass human rights violations that amount to atrocity crimes in the Philippines since July 2016. Over the last year there has been an escalation in the number of extrajudicial killings and arbitrary arrests as well as a rise in public statements of intent (See discussion of Determinant 4, for examples). The likelihood of prospective serious human rights violations remains high, especially given the absence of mitigating action (see Determinant 6). The anticipated large-scale atrocities are part of a concerted pattern of continuous gross human rights violations.
4. Host state intent	*Public statements of intent*: The President of the Philippines has publicly incited violence against drug suspects for the past four years, encouraging the killing of suspected drug users and dealers. *Systematic preparations*: Thousands of deaths have occurred as part of the Philippine National Police's anti-drug war related operations. Thousands of deaths attributed to unidentified groups have also been linked to the government. The government has offered unplausible justifications for thousands of killings, in an attempt to conceal / deny their illegitimacy.
5. The host state's territorial control and capacity	At present, the Philippine government is not seen as being 'unable' to contain the threat of atrocities posed by non-state actors, owing to its lack of capacity or control over its territory. While non-state actors bear responsibility for the commission of atrocities in some parts of the country, ongoing grave human rights violations are primarily carried out by state law enforcement authorities. Data of the past year suggests that grave human rights violations perpetrated by non-state groups are decreasing, while atrocities perpetrated by the government in the same period have been on the rise.
6. Measures taken by the host state	The Philippines has not taken any of the following measures listed in the framework outlined in this thesis that it could have conceivably taken to mitigate the threat of atrocities: publicly recognised and condemned relevant

	violation of human rights, issued unambiguous and public orders to its police and security forces to respect international humanitarian and human rights law, conducted thorough investigations into reported human rights violations or enforced accountability for relevant violations of human rights by arresting and/or prosecuting the authors of those attacks and it has not formulated an action plan with timelines for implementing specific measures to mitigate the risk of atrocities, in cooperation with relevant stakeholders (including affected populations and prospective victims of atrocities). The lack of action on the part of the Philippine government to address the threat of atrocities has led to the escalation of grave human rights violations in the country.
7. Host state cooperation	The Philippines does not allow/impedes the access of international human rights monitors and investigations into the human rights situation in the country: - the Philippines has withdrawn from the ICC to prevent further investigation of alleged crimes against humanity. - the Philippines has not ensured the safety of human rights workers in the country.
8. The host state's response to international requests to address the threat of atrocities	The Philippines has not complied with demands made by the UN to: - prevent extrajudicial killings and enforced disappearances. - to carry out impartial investigations and to hold perpetrators accountable.
9. The host state's track record of violence and relevant interacts with the international community	The Philippine has: - a track record of mass violence and serious human rights violations associated with atrocity crimes, including extrajudicial killings. - a track record of acting as an agent of mass violence. - a long-standing culture of violence and impunity for serious human rights violations. - a government that has displayed tolerance of and failed to seek accountability for past human rights violations.

CONCLUSION

This chapter offered a framework, which specified a range of determinants, detailed what is required of them, and how they interact in a bid to provide a continuum of steps that could help relevant stakeholders in conducting 'manifest failing' inquiries. Hence, it is important to underscore that the determinants of this framework are not discrete

categories and that some overlap between them is inevitable due to the links and interdependence between them, which were discussed in this chapter and illustrated in the discussion of the human rights situation in the Philippines. While no one determinant is sufficient to ascertain whether a host state is ‘manifestly failing’ to protect its populations or not, taken together the nine determinants that make up the framework proposed in this chapter offer comprehensive normative guidance that encourages and offers a continuum of steps for a thorough investigation into a host state’s ‘manifest failing’, without dictating a one-size-fits-all approach.

CONCLUSION

This thesis argued that the 'manifest failing' threshold, representing the threshold for collective action in response to mass atrocity crimes in Paragraph 139 of the 2005 WSOD, presently lacks the requisite clarity and content to fulfil its function as substantive licence for, and leash against, international action to protect vulnerable populations abroad from the threat of the four crimes. To address these shortfalls, it set out to answer the question of 'What determines a "manifest failing"?' in order to identify a set of substantive determinants and associated indicators that should be applied to appraisals as to whether there is a substantive basis for international action against a national government, on the grounds that it is not adequately fulfilling its primary responsibility to protect its populations from atrocity crimes. There are a number of negative implications of the substantive indeterminacy of the 'manifest failing' test.

On the one hand, the lack of substance means that the 'manifest failing' requirement sets a very low bar for justifying a coercive intervention in another state, thus compromising its function as a substantive leash to unwarranted foreign interference into the domestic affairs of third states. The lack of clarity means that the concept can be stretched to accommodate a wider range of justifications for coercive action that go beyond determining the host state's 'inability' or 'unwillingness' to protect their populations. As Ahmed underscores in relation to the threshold for self-defence against non-state actors, asserting that a host state is 'unable or unwilling' to supress the threat of a non-state actor operating from its territory 'can be a very subjective claim that is open to significant manipulation, particularly because a state's effectiveness… may often not be easily observable to other states and thus provides greater room for conflicting and self-serving interpretations'.[1111] Hence, the concerns expressed by some with regards to the fuzziness of the 'manifest failing' requirement and the potential for powerful states to abuse it are not without merit (see p. 24 of Introduction). This can be especially problematic when there are acute power asymmetries, as there often are, between powerful intervening states and much weaker host states, where atrocities crimes are suspected to be taking place. As seen in the case of the Russian incursion in Georgia, states can take advantage of the indeterminacy of the 'manifest failing' test to frame an intervention in R2P terms,

[1111] Ahmed, 'Defending Weak States', p. 14.

as a threshold requirement devoid of substance does not compel them to conduct a robust evidenced-based inquiry into the host state's 'manifest failing' and give reasons as to why it has been met in a given case. Thus, there are serious potential negative consequences of the 'manifest failing' threshold's limited capacity to serve as a substantive normative constraint on unjustified international action against a state that is allegedly not up to the task of fulfilling the protection responsibilities it has for its own populations and to delegitimise unwarranted action when it occurs.

On the other hand, the nebulous content of the 'manifest failing' threshold requirement renders it impractical for aiding determinations as to whether there is a just cause for collective action to protect foreign populations, effectively hindering its function as a substantive licence for international action. This can contribute to failings to protect on the part of the international community for several reasons. The lack of clarity surrounding the threshold for collective action makes it difficult for state and non-state actors to make a compelling case that a host state has crossed a red line and should be seen as neglecting its obligations to protect their population established under international law. This presents an obstacle to galvanising international attention towards, and subsequent action to tackle, a situation of concern. Consequently, scenarios in which grave human rights violations are taking place can be easily overlooked or ignored until it is too late, despite the calls of human rights groups and international bodies attempting to draw attention to the gravity of the situation and the 'manifest failing' of the host state.

As illustrated in this thesis, in the absence of compelling arguments that a substantive threshold for collective action has been crossed, states unwilling to take steps to prevent or halt an atrocity crisis (on account of their vested interest in obstructing such measures), can easily justify their opposition to responding to the plight of vulnerable populations. They can do so through advocating an alternative narrative, which dissociates the facts of the situation from the matter of atrocity crimes, or by simply rejecting claims that a threshold has been crossed without having to provide robust justification for doing so. In other words, in its present indeterminate state, the 'manifest failing' threshold requirement does not have constraining content that compels states to justify their arguments against collective action with reference to the notion of a substantive threshold for collective action.

In all of the aforementioned scenarios, the substantive indeterminacy of the 'manifest failing' threshold can contribute to what Labonte terms a 'double manifest failure', defined as humanitarian crises in which 'both the host state and the international community have failed to uphold their respective primary and secondary responsibilities to protect under R2P'.[1112] Comprehending the devastating consequences of this twofold failure does not take a leap of the imagination, as both historical and contemporary illustrations of its implications abound, including Sudan, Syria, and Myanmar, to name a few. Against this backdrop, it is clear that the function of the 'manifest failing' threshold as a licence for collective action should not be understated.

To address the negative implications of the substantive indeterminacy of the 'manifest failing' threshold, this thesis brought together insights from three different investigations: 1) existing research on mass atrocity crimes and the 'manifest failing' threshold (Chapter I); 2) the factors of the 'unwilling or unable' test in the analogous context of extraterritorial self-defence against non-state actors (Chapter III); and 3) state's understanding of 'manifest failing' in the case of Libya (Chapter IV). The determinants and indicators of 'manifest failing' identified from these investigations were triangulated, in order to develop a comprehensive normative framework for 'manifest failing' inquiries. In addition to constituting an important part of the main contribution of this thesis, namely, the 'manifest failing' framework outlined in Chapter V, each of these three key chapters (I, III and IV) makes a distinct contribution to the study of the 'manifest failing' threshold in their own right.

The first chapter provided a scoped synthesis of existing scholarly contributions to the study of the 'manifest failing' test, complemented by relevant insights from the literature on mass atrocity crimes and mass violence, in order to justify the inclusion of concrete determinants and indicators in the framework articulated in Chapter V. In so doing, it also contributed to the scholarship on 'manifest failing' and R2P by critically appraising relevant research on the topic to date, outlining gaps in the existing literature and scholarly understanding(s) of the 'manifest failing' threshold that need to be filled. These gaps were address in the thesis through the subsequent exploration of what this requirement should

[1112] Melissa T. Labonte, 'Whose Responsibility to Protect? The Implications of Double Manifest Failure for Civilian Protection', 16 (7) *The International Journal of Human Rights* (2012): 982–1002, p. 982.

entail normatively, particularly from different perspectives, as presented in Chapter III and IV.

Through the investigation of the analogous 'unwilling or unable' test, which applies in the context of self-defence against non-state actors, Chapter III of this thesis offered the first sustained endeavour to explore the relationship between the 'manifest failing' threshold and its predecessor. The analogy drawn between the two thresholds has allowed the thesis to develop a more comprehensive 'manifest failing' framework, by drawing on legal scholarship on the 'unable or unwilling' test, which is far more advanced in specifying the substantive criteria upon which 'unwilling or unable' determinations should be premised. Specifically, this exercise helped to identify a range of analogous determinants of 'manifest failing', which focus on assessing the 'unwillingness' and 'inability' of the host state with reference to its conduct in relation to the threat of atrocities and relevant interactions with the international community. In addition, the study of the 'unwilling or unable' test helped to underscore that the responsibility for the actions of a non-state group is ultimately borne by the host state. In so doing, this thesis drawn attention to the 'manifest failing' of states in scenarios, where an 'unable' state can also be 'unwilling' to deal with the threat of atrocity crimes posed by a non-state actor, which have been largely overlooked in scholarly contributions to the study of the 'manifest failing' threshold. In this way, Chapter III has helped to gain an insight into determining a 'manifest failing' in more complex situations, including a host state colluding with a non-state actor or committing atrocities alongside a non-state perpetrator; a host state that is not taking measures to prevent a non-state group from committing atrocities, even though it has the capacity to do so; and a host state that is unable to stop a non-state perpetration, but is also not seeking help or refusing to accept help from international actors.

In sum, the study of the 'unwilling or unable' test has underscored that excusing an uncooperative or idle state in the face of a threat posed by a non-state actor is not a viable justification to not take decisive collective action against that state. Likewise, the fact that in some situations non-state actors will be committing atrocities alongside the host state, should not detract attention from the 'manifest failing' of the host state, who bears the responsibility for the protection of its population. Thus, Chapter III has expanded thinking on the 'manifest failing' threshold in a way that could not be inspired by the sparse existing literature dedicated to it, nor the broader scholarship on R2P and mass atrocity

crimes, nor by the analysis of narrative case studies. It has also made a contribution to appreciating the role of non-state actors in atrocity crises, which is a key contemporary challenge for the R2P.

The third line of investigation that informed the study of the determinants of a 'manifest failing' in this thesis was the analysis of what is perhaps the most discussed R2P case of the 21st century. Nevertheless, with the exception of Bode's flawed, unsubstantiated and ultimately inadequate attempt to show how states made the determination that the Libyan government was 'manifestly failing' to protect its populations, this topic remained unexplored within the vast literature on Libya. While some may argue that this can be attributed to the fact that the 'manifest failing' of the Libyan government was indeed 'manifest', this is precisely the reason why it is the best case to examine when trying to explore state discourses over the threshold for collective action - in order to identify the determinants and associated indicators that states invoked in debates as to whether or not this threshold has been crossed. In a nutshell, Chapter IV has made novel contributions to the study of the Libyan case by shedding light on the language states used in debates surrounding the threshold for collective action and identifying the determinants and indicators that they offered in their justifications for the adoption of coercive measures in Libya.

Furthermore, triangulating these determinants and indicators with those identified in previous chapters has enhanced the framework for 'manifest failing' inquiries outlined in Chapter V for two reasons. First, it has demonstrated that there is meaningful overlap between states' understanding of 'manifest failing' in Libya and the determinants and indicators derived theoretically in Chapters I and III. Second, the inclusion of the determinants and indicators identified in Chapter IV in the overall framework ensures that it reflects the aspects of the situation that states find important to understanding 'manifest failing' determinations and thus are more likely to take into account (including as reference points for reason-giving) in their evaluation of future cases (as argued in Chapter II). Therefore, the insights derived from the analysis of justificatory discourses on the substantive threshold for collective action in the context of the Libyan intervention have contributed to the articulation of a more compelling 'manifest failing' framework in Chapter V.

The aforementioned three investigations contributed to devising this framework by providing the basis to identify and justify the inclusion of concrete determinants and indicators therein, as well as providing insights into how these determinants should be interpreted and appraised in conjunction with one another in order to judge whether a 'manifest failing' is taking place. Thus, this thesis provided the first comprehensive normative framework for 'manifest failing' inquiries, which addresses the negative implications stemming from the substantive indeterminacy of the 'manifest failing' threshold. In doing so, the thesis makes five important contributions.

First, it fills a lacuna by providing a framework of common reference within which argumentation over collective action can take place. Second, it provides a framework that can guide 'manifest failing' inquiries and assist relevant stakeholders in determining whether a 'manifest failing' is taking place. Specifically, it can aid them in applying the available information about a given case where vulnerable populations are facing a serious threat of atrocity crimes, so as to make a compelling evidence-based case that a national government is 'manifestly failing' to protect its populations. This is important, because part of what over 150 states agreed to at the 2005 World Summit was that they were 'prepared to take collective action' when national governments are 'manifestly failing' to protect their populations from the four atrocity crimes. Third, the framework outlined in this thesis provides normative criteria that require reason-giving in the event of international inaction in the face of a plausible threat of grave atrocity crimes. Fourth, these criteria can serve to 'check' and delegitimise unfounded arguments in favour of coercive action, thus discrediting unwarranted interventions. Fifth, the framework outlined in Chapter IV provides states, non-governmental and international organisations with a tool to raise awareness of a situation in which a national government is not living up to its 'responsibility to protect'. Comprehensive evidence-based assessments pointing to the 'manifest failing' of a national government to protect its populations from atrocity crimes will make it more difficult for reluctant states to avoid addressing arguments that a situation should be viewed from a mass atrocities perspective, thus compelling them to give reasons as to why they are opting not to act, while also limiting the number of plausible justifications they can offer.

As the application of this new 'manifest failing' framework to the atrocity crisis triggered by the Philippines' government 'war on drugs' illustrated, it could be particularly useful in cases where divergent arguments are being made as to whether a situation should be

viewed from a mass atrocity perspective and whether it warrants international attention. And since situations in which collective action merits deliberation are inevitably defined by the competing views and priorities of numerous states involved in decision-making on collective action, most, if not all, situations, where a threat of grave atrocity crimes exists, would benefit from the application of the framework developed in this thesis.

There may well be other determinants and indicators worth adding to the normative framework for 'manifest failing' inquiries outlined in Chapter V and this thesis does not pretend to posit an exhaustive or definitive list. What this thesis constitutes is the first attempt to systematically set out the abstract features of, and provide a robust skeleton for, the 'manifest failing' threshold. As a result, it is not the last step in the process of clarifying the substance of this requirement for collective international action. The determinants of a 'manifest failing' identified in this thesis offer a robust and substantially nuanced starting point from which relevant stakeholders, state or non-state actors, can scrutinise and start to clarify the basis for their 'manifest failing' determinations and give reasons as to why they have deemed a host state to be 'manifestly failing' or not. This practice of reason-giving and conceptual sharpening can help to infuse decisions over collective action with more clarity and transparency. Moreover, it is also important to note that the framework offered here is designed with a view to be applied to more cases beyond the examples and cases looked at in Chapters IV and V. Particular candidates where the framework can offer applied insights would be Myanmar, China and Venezuela. Hence, as a heuristic tool, the framework can be advanced over time to reflect new research into mass atrocity crimes and applied to new cases in which states actively debate the threshold for collective action in response to threats of atrocity crimes.

ANNEX I: DEFINITIONS OF MASS ATROCITY CRIMES

GENOCIDE

The legal definition of the crime of genocide can be found in Article 2 of the Convention on the Prevention and Punishment of the Crime of Genocide.

Box 4. Legal definition of genocide

Convention on the Prevention and Punishment of the Crime of Genocide

Article 2

'In the present Convention, genocide means any of the following acts committed with intent to destroy, in whole or in part, a national, ethnical, racial or religious group, as such:

(a) Killing members of the group;

(b) Causing serious bodily or mental harm to members of the group;

(c) Deliberately inflicting on the group conditions of life calculated to bring about its physical destruction in whole or in part;

(d) Imposing measures intended to prevent births within the group;

(e) Forcibly transferring children of the group to another group.'

CRIMES AGAINST HUMANITY

The legal definition of crimes against humanity can be found in Article 7 of the Rome Statute of the International Criminal Court. [1113]

Box 5. Legal definition of crimes against humanity

Rome Statute of the International Criminal Court

Article 7

'1. For the purpose of this Statute, "crime against humanity" means any of the following acts when committed as part of a widespread or systematic attack directed against any civilian population, with knowledge of the attack:

(a) Murder;

(b) Extermination;

(c) Enslavement;

(d) Deportation or forcible transfer of population;

(e) Imprisonment or other severe deprivation of physical liberty in violation of fundamental rules of international law;

(f) Torture;

(g) Rape, sexual slavery, enforced prostitution, forced pregnancy, enforced sterilization, or any other form of sexual violence of comparable gravity;

(h) Persecution against any identifiable group or collectivity on political, racial, national, ethnic, cultural, religious, gender as defined in paragraph 3, or other grounds that are universally recognized as impermissible under international law, in connection with any act referred to in this paragraph or any crime within the jurisdiction of the Court;

(i) Enforced disappearance of persons;

(j) The crime of apartheid;

(k) Other inhumane acts of a similar character intentionally causing great suffering, or serious injury to body or to mental or physical health.'

[1113] International Criminal Court, 'Rome Statute of the International Criminal Court' (The Hague: International Criminal Court, 2011). [Accessed 20 March 2021]. Available at: https://www.icc-cpi.int/resource-library/documents/rs-eng.pdf

WAR CRIMES

The legal definition of war crimes can be found in Article 8 of the Rome Statute of the International Criminal Court.

Box 6. Legal definition of war crimes

Rome Statute of the International Criminal Court

Article 8

'1. The Court shall have jurisdiction in respect of war crimes in particular when committed as part of a plan or policy or as part of a large-scale commission of such crimes.

2. For the purpose of this Statute, "war crimes" means:

(a) Grave breaches of the Geneva Conventions of 12 August 1949, namely, any of the following acts against persons or property protected under the provisions of the relevant Geneva Convention:

(i) Wilful killing;

(ii) Torture or inhuman treatment, including biological experiments;

(iii) Wilfully causing great suffering, or serious injury to body or health;

(iv) Extensive destruction and appropriation of property, not justified by military necessity and carried out unlawfully and wantonly;

(v) Compelling a prisoner of war or other protected person to serve in the forces of a hostile Power;

(vi) Wilfully depriving a prisoner of war or other protected person of the rights of fair and regular trial;

(vii) Unlawful deportation or transfer or unlawful confinement;

(viii) Taking of hostages.

(b) Other serious violations of the laws and customs applicable in international armed conflict, within the established framework of international law, namely, any of the following acts:

(i) Intentionally directing attacks against the civilian population as such or against individual civilians not taking direct part in hostilities;

(ii) Intentionally directing attacks against civilian objects, that is, objects which are not military objectives;

(iii) Intentionally directing attacks against personnel, installations, material, units or vehicles involved in a humanitarian assistance or peacekeeping mission in accordance with the Charter of the United Nations, as long as they are entitled to the protection given to civilians or civilian objects under the international law of armed conflict;

(iv) Intentionally launching an attack in the knowledge that such attack will cause incidental loss of life or injury to civilians or damage to civilian objects or widespread, long-term and severe damage to the natural environment which would be clearly excessive in relation to the concrete and direct overall military advantage anticipated;

(v) Attacking or bombarding, by whatever means, towns, villages, dwellings or buildings which are undefended and which are not military objectives;

(vi) Killing or wounding a combatant who, having laid down his arms or having no longer means of defence, has surrendered at discretion;

(vii) Making improper use of a flag of truce, of the flag or of the military insignia and uniform of the enemy or of the United Nations, as well as of the distinctive emblems of the Geneva Conventions, resulting in death or serious personal injury;

(viii) The transfer, directly or indirectly, by the Occupying Power of parts of its own civilian population into the territory it occupies, or the deportation or transfer of all or parts of the population of the occupied territory within or outside this territory;

(ix) Intentionally directing attacks against buildings dedicated to religion, education, art, science or charitable purposes, historic monuments, hospitals and places where the sick and wounded are collected, provided they are not military objectives;

(x) Subjecting persons who are in the power of an adverse party to physical mutilation or to medical or scientific experiments of any kind which are neither justified by the medical, dental or hospital treatment of the person concerned nor carried out in his or her interest, and which cause death to or seriously endanger the health of such person or persons;

(xi) Killing or wounding treacherously individuals belonging to the hostile nation or army;

(xii) Declaring that no quarter will be given;

(xiii) Destroying or seizing the enemy's property unless such destruction or seizure be imperatively demanded by the necessities of war;

(xiv) Declaring abolished, suspended or inadmissible in a court of law the rights and actions of the nationals of the hostile party;

(xvi) Pillaging a town or place, even when taken by assault;

(xvii) Employing poison or poisoned weapons;

(xviii) Employing asphyxiating, poisonous or other gases, and all analogous liquids, materials or devices;

(xix) Employing bullets which expand or flatten easily in the human body, such as bullets with a hard envelope which does not entirely cover the core or is pierced with incisions;

(xx) Employing weapons, projectiles and material and methods of warfare which are of a nature to cause superfluous injury or unnecessary suffering or which are inherently indiscriminate in violation of the international law of armed conflict, provided that such weapons, projectiles and material and methods of warfare are the subject of a comprehensive prohibition and are included in an annex to this Statute, by an amendment in accordance with the relevant provisions set forth in articles 121 and 123;

(xxi) Committing outrages upon personal dignity, in particular humiliating and degrading treatment;

(xxii) Committing rape, sexual slavery, enforced prostitution, forced pregnancy, as defined in article 7, paragraph 2 (f), enforced sterilization, or any other form of sexual violence also constituting a grave breach of the Geneva Conventions;

(xxiii) Utilizing the presence of a civilian or other protected person to render certain points, areas or military forces immune from military operations;

(xxiv) Intentionally directing attacks against buildings, material, medical units and transport, and personnel using the distinctive emblems of the Geneva Conventions in conformity with international law;

(xxv) Intentionally using starvation of civilians as a method of warfare by depriving them of objects indispensable to their survival, including wilfully impeding relief supplies as provided for under the Geneva Conventions;

(xxvi) Conscripting or enlisting children under the age of fifteen years into the national armed forces or using them to participate actively in hostilities.

(c) In the case of an armed conflict not of an international character, seriousviolations of article 3 common to the four Geneva Conventions of 12 August 1949, namely, any of the following acts committed against persons taking no active part in the hostilities, including members of armed forces who have laid down their arms and those placed hors de combat by sickness, wounds, detention or any other cause:

(i) Violence to life and person, in particular murder of all kinds, mutilation, cruel treatment and torture;

(ii) Committing outrages upon personal dignity, in particular humiliating and degrading treatment;

(iii) Taking of hostages;

(iv) The passing of sentences and the carrying out of executions without previous judgement pronounced by a regularly constituted court, affording all judicial guarantees which are generally recognized as indispensable.

(d) Paragraph 2 (c) applies to armed conflicts not of an international character and thus does not apply to situations of internal disturbances and tensions, such as riots, isolated and sporadic acts of violence or other acts of a similar nature.

(e) Other serious violations of the laws and customs applicable in armed conflicts not of an international character, within the established framework of international law, namely, any of the following acts:

(i) Intentionally directing attacks against the civilian population as such or against individual civilians not taking direct part in hostilities;

(ii) Intentionally directing attacks against buildings, material, medical units and transport, and personnel using the distinctive emblems of the Geneva Conventions in conformity with international law;

(iii) Intentionally directing attacks against personnel, installations, material, units or vehicles involved in a humanitarian assistance or peacekeeping mission in accordance with the Charter of the United Nations, as long as they are entitled to the protection given to civilians or civilian objects under the international law of armed conflict;

(iv) Intentionally directing attacks against buildings dedicated to religion, education, art, science or charitable purposes, historic monuments, hospitals and places where the sick and wounded are collected, provided they are not military objectives;

(v) Pillaging a town or place, even when taken by assault;

(vi) Committing rape, sexual slavery, enforced prostitution, forced pregnancy, as defined in article 7, paragraph 2 (f), enforced sterilization, and any other form of sexual violence also constituting a serious violation of article 3 common to the four Geneva Conventions;

(vii) Conscripting or enlisting children under the age of fifteen years into armed forces or groups or using them to participate actively in hostilities;

(viii) Ordering the displacement of the civilian population for reasons related to the conflict, unless the security of the civilians involved or imperative military reasons so demand;

(ix) Killing or wounding treacherously a combatant adversary;

(x) Declaring that no quarter will be given;

(xi) Subjecting persons who are in the power of another party to the conflict to physical mutilation or to medical or scientific experiments of any kind which are neither justified by the medical, dental or hospital treatment of the person concerned nor carried out in his or her interest, and which cause death to or seriously endanger the health of such person or persons;

(xii) Destroying or seizing the property of an adversary unless such destruction or seizure be imperatively demanded by the necessities of the conflict;

(f) Paragraph 2 (e) applies to armed conflicts not of an international character and thus does not apply to situations of internal disturbances and tensions, such as riots, isolated and sporadic acts of violence or other acts of a similar nature. It applies to armed conflicts that take place in the territory of a State when there is protracted armed conflict between governmental authorities and organized armed groups or between such groups.'

ETHNIC CLEANSING

Ethnic cleansing has not been recognised as an autonomous international crime, but is commonly understood to contain acts (or elements of acts) that constitute one of the other atrocity crimes, especially genocide. While related to genocide 'ethnic cleansing is focused more closely… on geography and on forced removal of ethnic or related groups from particular areas'.[1114] The utmost overlap between the two crimes 'takes place when forced removal of population leads to a group's destruction'.[1115] However, this relationship between ethnic cleansing and genocide cannot be circumscribed 'in legal terms because there is no international convention that defines ethnic cleansing'.[1116] Consequently, a historical approach is ordinarily employed to provide an operational definition of ethnic cleansing and delineate its key constitutive elements.

The Oxford Book of Genocide Studies identifies the wars in the former Yugoslavia in the 1990s as the moment when 'ethnic cleansing' passed into the vocabulary of terms closely related to genocide.[1117] Accordingly, subsequent definitions of ethnic cleansing (including the one provided in the UN Framework of Analysis for Atrocity Crimes) customarily return to the work of the Commission of Experts appointed by the Secretary-General in 1992, at the request of the UN Security Council, to investigate reported violence against individuals and property, deportation and expulsion in the former Yugoslavia.

The first interim report of the Commission resolved that: '"ethnic cleansing" means rendering an area ethnically homogenous by using force or intimidation to remove persons of given groups from the area. "Ethnic cleansing" is contrary to international law'.[1118] The Commission's final report defined it as 'a purposeful policy designed by one ethnic or religious group to remove by violent and terror-inspiring means the civilian

[1114] Benjamin Lieberman, '"Ethnic Cleansing" versus Genocide', in Donald Bloxham and A. Dirk Moses (eds.) *The Oxford Handbook of Genocide Studies* (New York: Oxford University Press, 2010): 42-60, p.42.

[1115] Ibid.

[1116] Ibid.

[1117] Ibid.

[1118] United Nations Security Council, 'Interim Report of the Commission of Experts Established Pursuant to Security Council Resolution 780 (1992)', S/25274, 26 January 1993, Annex I, para. 55.

population of another ethnic or religious group from certain geographic areas'.[1119] With regards to the acts that constitute ethnic cleansing, both of reports state that:

> 'Based on the many reports describing the policy and practices conducted in the former Yugoslavia, "ethnic cleansing" has been carried out by means of **murder, torture, arbitrary arrest and detention, extra-judicial executions, rape and sexual assaults, confinement of civilian population in ghetto areas, forcible removal, displacement and deportation of civilian population, deliberate military attacks or threats of attacks on civilians and civilian areas, and wanton destruction of property**. Those practices constitute crimes against humanity and can be assimilated to specific war crimes. Furthermore, such acts could also fall within the meaning of the Genocide Convention' (emphasis added).[1120]

[1119] United Nations Security Council, 'Final Report of the Commission of Experts Established Pursuant to Security Council Resolution 780 (1992)', S/1994/674, 27 May 1994, Annex, para. 130.

[1120] S/25274 (1993), Annex I, para.56; S/1994/674 (1994), Annex, para. 129.

BIBLIOGRAPHY

Adams, Simon. 'Libya and the Responsibility to Protect.' *Global Centre for the Responsibility to Protect Occasional Paper Series* No. 3. New York: Global Centre for the Responsibility to Protect, 2012.

Adler-Nissen, Rebecca and Vincent Pouliot. 'Power in Practice: Negotiating the International Intervention in Libya.' *European Journal of International Relations* 20, no. 4 (2014): 889-911.

AFP News Agency. 'Obama Vows to "Take Out" Terror Targets in Pakistan'. 28 September 2008. Accessed 11 January 2021. https://www.youtube.com/watch?v=LqOncFQUZsw.

African Union Peace and Security Council. Communique on Libya, 23 February 2011 Accessed 5 February 2020. http://www.peaceau.org/uploads/psc-communique-on-the-situation-in-libya.pdf.

African Union Peace and Security Council. Communique of the 265th Meeting of the Peace and Security Council, PSC/PR/COMM.2 (CCXLV), 10 March 2011. Accessed 4 April 2020. https://www.gov.za/communiqué-265th-meeting-african-union-peace-and-security-council.

Ahmed, Dawood I. 'Defending Weak States against the Unwilling or Unable Doctrine of Self-Defense.' *Journal of International Law and International Relations* 9 (2013): 1-37.

Al Jazeera. 'Libya protests spread and intensify.' 21 February 2011. Accessed 06 February 2020. https://www.aljazeera.com/news/2011/2/21/libya-protests-spread-and-intensify.

Al Jazeera. 'GCC: Libya regime lost legitimacy.' 11 March 2011. Accessed 16 April 2020. https://www.aljazeera.com/news/middleeast/2011/03/201131021173060618l.html.

Al Jazeera. ‘Why the UN failed to save the Rohingya.’ 28 June 2019. Accessed 26 October 2020. https://www.aljazeera.com/news/2019/6/28/why-the-un-failed-to-save-the-rohingya.

Al Jazeera. ‘UN fails to take action on order against Myanmar on Rohingya.’ 5 February 2020. Accessed 26 October 2020. https://www.aljazeera.com/news/2020/2/5/un-fails-to-take-action-on-order-against-myanmar-on-rohingya.

Al Jazeera. ‘“I will kill you”: Philippines’ Duterte renews drug war threat’. 5 June 2020. Accessed 21 March 2021. https://www.aljazeera.com/news/2020/6/5/i-will-kill-you-philippines-duterte-renews-drug-war-threat.

Al Jazeera. ‘“Appalled”: UN urges probe into killing of Philippine activists.’ 10 March 2021. Accessed 24 March 2021. https://www.aljazeera.com/news/2021/3/10/un-urges-probe-into-killings-of-philippine-activists.

Akande, Dapo and Thomas Liefländer. ‘Clarifying Necessity, Imminence, and Proportionality in the Law of Self-Defense.’ *American Journal of International Law,* 107, no. 3 (2013): 563-570.

Amnesty International. ‘Philippines: UN must intensify pressure to end killings as impunity reigns.’ 25 September 2020. Accessed 21 March 2021. https://www.amnesty.org/en/latest/news/2020/09/philippines-un-pressure-end-killings/.

Amnesty International. ‘Philippines: UN resolution a missed chance for justice but scrutiny continues.’ 7 October 2020. Accessed 5 February 2020. https://www.amnesty.org/en/latest/news/2020/10/philippines-un-resolution-a-missed-chance-for-justice-but-scrutiny-continues/.

Amnesty International UK. ‘More than 7,000 killed in the Philippines in six months, as president encourages murder.’ 18 May 2020. Accessed 21 March 2021. https://www.amnesty.org.uk/philippines-president-duterte-war-on-drugs-thousands-killed.

Araud, Gérard. ‘Libya (UN/Sanctions/Libya).’ Remarks to the press by Gérard Araud, France’s Permanent Representative to the United Nations. 25 February 2011, New

York. Published on 1 March 2011. Accessed 23 September 2019. https://franceintheus.org/spip.php?article2192.

Araud, Gérard. 'Libya/adoption of UNSCR 1970.' Remarks to the press by Gérard Araud, France's Permanent Representative to the United Nations, 26 February 2011, New York. Published on 1 March 2011. Accessed 28 January 2020. https://franceintheus.org/spip.php?article2193.

Asia-Pacific Centre for the Responsibility to Protect. 'Atrocity Crimes Risk Assessment Series: The Philippines. Volume 7.' September 2018. Accessed 24 March 2021. https://r2pasiapacific.org/files/2497/Risk%20Assessment%20The%20Philippines%20FINALwith%20images%20opti%281%29.pdf.

Asia-Pacific Centre for the Responsibility to Protect. 'The Philippines Baseline Assessment of R2P Implementation.' October 2019. Accessed 24 March 2021. https://r2pasiapacific.org/files/4214/Philippines_annex_assessment.pdf.

Asia-Pacific Centre for the Responsibility to Protect. 'November 2019: Risk Assessment Series Volume 9: Myanmar.' November 2019. Accessed 21 October 2020. https://r2pasiapacific.org/files/4243/Risk_Assessment_myanmar_vol9_november20192.pdf.

Asia-Pacific Centre for the Responsibility to Protect. 'The Philippines: Asia Pacific Regional Outlook February 2020.' February 2020. Accessed 24 March 2021. https://r2pasiapacific.org/files/4781/APO%20The%20Philippines%20February%202020%20PDF.pdf.

Austin, Rod. '"War on drugs" makes Philippines fourth most dangerous country – report.' *Guardian*, 9 July 2019. Accessed 24 March 2021. https://www.theguardian.com/global-development/2019/jul/09/war-on-drugs-makes-philippines-fourth-most-dangerous-country-report.

Ban, Ki-moon. 'Implementing the Responsibility to Protect.' Report of the Secretary-General, A/63/677. 12 January 2009.

Ban, Ki-moon. 'Responsibility to Protect: Timely and Decisive Response.' Report of the Secretary-General, A/66/874-S/2012/578, 25 July 2012.

BBC. ‘Libya protests: Defiant Gaddafi refuses to quit.’ 22 February 2011. Accessed 4 February 2020. https://www.bbc.co.uk/news/world-middle-east-12544624.

BBC. ‘UN Secretary-General Ban Ki-moon condemns Libya violence.’ 23 February 2011. Accessed 4 February 2020. https://www.bbc.co.uk/news/av/world-middle-east-12561784/un-secretary-general-ban-ki-moon-condemns-libya-violence.

BBC. ‘China forcing birth control on Uighurs to suppress population, report says.’ 29 June 2020. Accessed 20 March 2021. https://www.bbc.com/news/world-asia-china-53220713.

Bellamy, Alex J. ‘Libya and the Responsibility to Protect: The Exception and the Norm.’ *Ethics & International Affairs* 25, no. 3 (2011): 263-269.

Bellamy, Alex J. *Mass Atrocities and Armed Conflict: Links, Distinctions, and Implications for the Responsibility to Prevent.* Muscatine, IA: The Stanley Foundation, 2011.

Bellamy, Alex J. ‘The Responsibility to Protect and the Problem of Regime Change.’ In *The Responsibility to Protect: Challenges and Opportunities in Light of the Libyan Intervention*, edited by Thomas G. Weiss et al., 20-23. e-International Relations, 2011. Accessed 12 August 2020. https://www.e-ir.info/wp-content/uploads/R2P.pdf.

Bellamy, Alex J. *The First Response: Peaceful Means in the Third Pillar of the Responsibility to Protect*. Muscatine, IA: The Stanley Foundation, 2015.

Bellamy, Alex J. *The Responsibility to Protect: A Defense*. New York: Oxford University Press, 2015.

Bellamy, Alex J. ‘The Three Pillars of the Responsibility to Protect.’ *Pensamiento Propio* 41, no. 20 (2015): 35-64.

Bellamy, Alex J. and Blagovesta Tacheva. ‘R2P and the Emergence of Responsibilities across Borders.’ In *The State and Cosmopolitan Responsibilities*, edited by Richard Beardsworth, Garrett Wallace Brown, and Richard Shapcott, 15-40. Oxford: Oxford University Press, 2019.

Bellamy, Alex J. and Paul D. Williams. ‘The new politics of protection? Côte d’Ivoire, Libya and the responsibility to protect’. *International Affairs* 87, no. 4 (2011): 825-850.

Beresford, Alexander. ‘A responsibility to protect Africa from the West? South Africa and the NATO Intervention in Libya.’ *International Politics* 52, no. 3 (2015): 288-304.

Berlinger, Joshua. ‘Philippines: UN decision to probe drug war “straight from the mouth of the Queen in Alice in Wonderland.” *CNN*, 12 July 2019. Accessed 3 February 2021. https://edition.cnn.com/2019/07/11/asia/philippines-drug-war-un-intl-hnk.

Bethlehem, Daniel. ‘Self-Defense against an Imminent or Actual Armed Attack by Nonstate Actors.’ *The American Journal of International Law* 106, no. 4 (2012): 769-777.

Bode, Ingvild. ‘“Manifestly Failing” and “Unwilling or Unable” as Intervention Formulas: A Critical Assessment.’ *Rethinking Humanitarian Intervention in the 21st Century*, edited by Aiden Warren and Damian Grenfell, 164-191. Edinburgh: Edinburgh University Press, 2017.

Bowcott, Owen. ‘Osama bin Laden death: Pakistan says US may have breached sovereignty.’ *Guardian*, 5 May 2011. Accessed 24 July 2020. https://www.theguardian.com/world/2011/may/05/osama-bin-laden-pakistan-us-sovereignty.

Brown, Gillian and George Yule. *Discourse Analysis*. Cambridge: Cambridge University Press, 1983.

Brunnée, Jutta and Stephen J. Toope. *Legitimacy and Legality in International Law*. Cambridge: Cambridge University Press, 2010.

Burns, William J. *The Back Channel: A Memoir of American Diplomacy and Its Case for Renewal*. New York: Penguin Random House, 2020.

Butcher, Charles, Benjamin E. Goldsmith, Sascha Nanlohy, Arcot Sowmya and David Muchlinski. ‘Introducing the Targeted Mass Killing Data Set for the Study and

Forecasting of Mass Atrocities.' *Journal of Conflict Resolution* 64, no. 7-8 (2020): 1524-1547.

Cameron, David. 'Transcript of the PM's YouTube interview in Oman', 24 February 2011. Published on 26 February 2011. Accessed 29 January 2020. https://www.gov.uk/government/speeches/transcript-of-the-pms-youtube-interview-in-oman.

Cameron, David. 'Prime Minister's statement on Libya', Statement on the situation in Libya to the House of Commons, 28 February 2011. Accessed 30 January 2020. https://www.gov.uk/government/speeches/prime-ministers-statement-on-libya--2.

Cameron, David. 'Speech at a press conference in Brussels.' 11 March 2011, Brussels. Accessed 12 November 2019. https://www.gov.uk/government/speeches/press-conference-in-brussels.

Cameron, David. 'Statement on UN Resolution 1973.' 18 March 2011. Accessed 12 July 2020. https://www.bbc.co.uk/news/uk-politics-12786225.

Cameron, David. 'PM Statement to the House on Libya: Prime Minister David Cameron has secured comprehensive support from MPs for the military intervention in Libya.' 21 March 2011. Accessed 14 January 2021. https://www.gov.uk/government/speeches/pm-statement-to-the-house-on-libya.

Cameron, David and Nicolas Sarkozy. 'Letter from David Cameron and Nicolas Sarkozy to Herman Van Rompuy.' *Guardian*, 10 March 2011. Accessed 14 February 2020. https://www.theguardian.com/world/2011/mar/10/libya-middleeast.

Chalk, Frank and Kurt Jonassohn. *The History and Sociology of Genocide, Analyses and Case Studies*. London: Yale University Press, 1990.

Chayes, Abraham and Antonia Handler Chayes. *The New Sovereignty: Compliance with International regulatory Agreements*. Cambridge, MA: Harvard University Press, 1995.

Chesterman, Simon. 'Legality versus legitimacy: humanitarian intervention, the Security Council, and the rule of law.' *Security Dialogue* 33, no. 3 (2002): 293-307.

Chesterman, Simon. '"Leading from Behind": The Responsibility to Protect, the Obama Doctrine, and Humanitarian Intervention after Libya.' *Ethics & International Affairs* 25, no. 3 (2011): 279-285.

Christakis, Theodore. 'Challenging the "Unwilling or Unable" Test.' In *Self-Defence against Non-State Actors: Impulses from the Max Planck Trialogues on the Law of Peace and War*, edited by Anne Peters and Christian Marxsen, Research Paper No. 2017-07: 17-20. Max Planck Institute for Comparative Public Law & International Law (MPIL), 2017. Accessed 16 July 2020. http://dx.doi.org/10.2139/ssrn.2941640.

CNA. 'Mahathir at UN condemns Myanmar's treatment of Rohingya.' 29 September 2018. Accessed 31 January 2021. https://www.channelnewsasia.com/news/asia/mahathir-at-un-condemns-myanmar-s-treatment-of-rohingya-10770568.

CNN Philippines. 'Duterte on UN rights probe: State your purpose first.' 12 July 2019. Accessed 21 March 2021. https://www.cnnphilippines.com/news/2019/7/12/Rodrigo-Duterte-UN-human-rights-probe.html.

CNN Philippines. 'UN rights chief: Revoke PH policies, rhetoric that lead to killings and abuses.' 16 September 2020. Accessed 24 March 2020. https://www.cnnphilippines.com/news/2020/9/16/united-nations-bachelet-philippines-duterte-drug-war.html.

Cohen, Benjamin J. 'The Way Forward', *New Political Economy* 14, no. 3 (2009): 395-400.

Conde, Carlos. 'Killings in Philippines Up 50 Percent During Pandemic: Drug War' Deaths Rise Dramatically as Country Reels from Covid-19.' *Human Rights Watch*. 8 September 2020. Accessed 21 March 2021. https://www.hrw.org/news/2020/09/08/killings-philippines-50-percent-during-pandemic.

Corten, Olivier. 'Has Practice Led to an "Agreement Between the Parties" Regarding the Interpretation of Article 51 of the UN Charter?' *Self-Defence Against Non-State Actors: Impulses from the Max Planck Trialogues on the Law of Peace and War*,

edited by Anne Peters and Christian Marxsen, Research Paper No. 2017-07: 14-16. Max Planck Institute for Comparative Public Law & International Law (MPIL), 2017. Accessed 16 November 2020. http://dx.doi.org/10.2139/ssrn.2941640.

Corten, Olivier. 'The "Unwilling or Unable" Test: Has it Been, and Could it be, Accepted?' *Leiden Journal of International Law* 29 (2016): 777-799.

Couzigou, Irène. 'The Right to Self-Defence Against Non-State Actors – Criteria of the "Unwilling or Unable" Test.' *Self-Defence Against Non-State Actors: Impulses from the Max Planck Trialogues on the Law of Peace and War*, edited by Anne Peters and Christian Marxsen, Research Paper No. 2017-07: 47-49. Max Planck Institute for Comparative Public Law & International Law (MPIL), 2017. Accessed 16 July 2020. http://dx.doi.org/10.2139/ssrn.2941640.

Crossley, Noëlle. *Evaluating the Responsibility to Protect: Mass Atrocity Prevention as a Consolidating Norm in International Society*. Abingdon and New York: Routledge, 2016.

Danish Institute for International Studies (DIIS). *New Threats and the Use of Force*. Copenhagen: Danish Institute for International Studies, 2005. Accessed 16 July 2020. https://www.files.ethz.ch/isn/16807/new_threats_whole_web.pdf.

Davies, Phil H.J. 'Spies as Informants: Triangulation and the Interpretation of Elite Interview Data in the Study of the Intelligence and Security Services.' *Politics* 21, no. 1 (2001): 73-80.

Davis, Jonathan E. 'From Ideology to Pragmatism: China's Position on Humanitarian Intervention in the Post-Cold War Era.' In *The Legality and Legitimacy of the Use of Force in Northeast Asia*, edited by Brendan Howe and Boris Kondoch, 117-190. Leiden: Brill, 2013.

Deeks, Ashley. '"Unwilling or Unable": Toward a Normative Framework for Extraterritorial Self-Defense.' *Virginia Journal of International Law* 52, no. 3 (2012): 483-550.

de Vera, Ben O. 'Duterte order shuns all loans, grants, aid from 18 countries backing probe of PH killings.' 20 September 2019. Accessed 21 March 2021.

https://business.inquirer.net/279344/duterte-order-shuns-all-loans-grants-aid-from-18-countries-in-favor-of-probe-of-ph-killings.

Denzin, Norman. *The Research Act: A Theoretical Introduction to Sociological Methods.* New York: McGraw-Hill, 1978.

Dinstein, Yoram. *War, Aggression and Self-Defense*. Third Edition. Cambridge: Cambridge University Press, 2003.

Doyle, Michael W. 'The Politics of Global Humanitarianism: The Responsibility to Protect before and after Libya.' *International Politics: Responsibility to Protect and Humanity* 53, no. 1 (2016): 14-31.

Dunford, Robin and Michael Neu. *Just War and the Responsibility to Protect: A Critique*. London: Zed Books, 2019.

Dunford, Robin and Michael Neu. 'The Responsibility to Protect in a world of already existing intervention', *European Journal of International Relations* 25, no. 4 (2019): 1080–1102.

Dunne, Tim and Jess Gifkins, 'Libya and the state of intervention.' *Australian Journal of International Affairs* 65, no. 5 (2011): 515-529.

Dunne, Tim and Katharine Gelber. 'Arguing Matters: The Responsibility to Protect and the Case of Libya.' *Global Responsibility to Protect* 6, no. 3 (2014): 326-349.

Eckstein, Harry. 'Case Study and Theory in Political Science'. In *Case Study Method*, edited by Roger Gomm, Martyn Hammersley, and Peter Foster, 119-164. London: SAGE, 2000.

Ellis-Petersen, Hannah and Emanuel Stoakes. 'UN report condemns its conduct in Myanmar as systemic failure.' *Guardian*, 17 June 2019. Accessed 26 October 2020. https://www.theguardian.com/world/2019/jun/17/un-report-myanmar-rohingya-systemic-failure.

Erzberger, Christian and Gerald Prein. 'Triangulation: Validity and empirically based hypothesis construction. Quality and Quantity.' *Quality and Quantity* 31, no. 2 (1997): 141-154.

European Council, 'Declaration adopted by the extraordinary European Council', EUCO 7/1/11 (Rev 1), 11 March 2011, Brussels. Accessed 16 April 2020. 20 April 2011. https://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ec/119780.pdf

Evans, Gareth. 'Russia, Georgia and the Responsibility To Protect.' *Amsterdam Law Forum* 1, no. 2 (2009): 25-28.

Fairclough, Isabela and Norman Fairclough. *Political Discourse Analysis: A Method for Advanced Students.* Abingdon: Routledge, 2012.

Fairclough, Norman. *Analysing Discourse: Textual Analysis for Social Research.* London and New York: Routledge, 2003.

Farmer, Tracy, Kerry Robinson, Susan J. Elliott, and John Eyles. 'Developing and Implementing a Triangulation Protocol for Qualitative Health Research.' *Qualitative Health Research* 16, no. 3 (2006): 377-394.

Fielding, Nigel G. & Jane L. Fielding. *Linking Data.* Beverley Hills: SAGE, 1986.

Fillon, François. 'Statement by the Government on the intervention of the Armed Forces in Libya for the implementation of UN Security Council Resolution 1973, and the debate on this statement', National Assembly (France), 22 March 2011. Accessed 15 January 2020. http://www.assemblee-nationale.fr/13/cri/2010-2011/20110144.asp#INTER_0.

Flick, Uwe. 'Triangulation Revisited – Strategy of Validation or Alternative.' *Journal for the Theory of Social Behaviour* 22, no. 2 (1992): 175-197.

France 24. 'Gaddafi forces pound key rebel-held town.' 15 March 2011. Accessed 7 February 2020. https://www.france24.com/en/20110315-gaddafi-forces-pound-key-rebel-held-town-ajdabiya-libya-airstrikes.

France 24. 'Russia says NATO strikes on Libya exceed mandate.' 15 April 2011. Accessed 25 March 2020. https://www.france24.com/en/20110415-russia-says-nato-libya-strikes-exceed-un-mandate.

France in the United States website. 'France/UK/US/message to Gaddafi', Joint message addressed to Colonel Gaddafi by France, the United Kingdom and the United States,

with the support of Arab countries. 18 March 2011. Accessed 23 September 2019. https://franceintheus.org/spip.php?article2237.

French Embassy in the UK website. Accessed 29 January 2020. https://uk.ambafrance.org/.

French Ministry of Foreign and European Affairs. 'Situation in Libya.' Communiqué issued by the Ministry of Foreign and European Affairs. 24 February 2011, Paris. Accessed 17 January 2020. https://franceintheus.org/spip.php?article2174.

Gallagher, Adrian. 'A Clash of Responsibilities: Engaging with Realist Critiques of the R2P.' *Global Responsibility to Protect* 4, no. 3 (2012): 334–57.

Gallagher, Adrian. *Genocide and its Threat to Contemporary International Order*. Basingstoke: Palgrave Macmillan, 2013.

Gallagher, Adrian. 'Syria and the indicators of a "manifest failing".' *The International Journal of Human Rights* 18, no.1 (2014): 1-19.

Gallagher, Adrian. 'What constitutes a "Manifest Failing"? Ambiguous and inconsistent terminology and the Responsibility to Protect.' *International Relations* 28, no. 4 (2014): 428-444.

Gallagher, Adrian. 'The promise of pillar II: analysing international assistance under the Responsibility to Protect.' *International Affairs* 91, no. 6 (2015): 1259-1275.

Gallagher, Adrian, Euan Raffle and Zain Maulana. 'Failing to fulfil the responsibility to protect: the war on drugs as crimes against humanity in the Philippines.' *The Pacific Review* 33, no. 2 (2019): 247-277.

Garwood-Gowers, Andrew. 'China and the "responsibility to protect": The implications of the Libyan intervention.' *Asian Journal of International Law* 2, no. 2 (2012): 375-393.

Gavilan, Jodesz. 'CHR still out of the loop in much hyped DOJ drug war panel.' *Rappler*. 8 October 2020. Accessed 24 March 2021. https://www.rappler.com/nation/commission-human-rights-still-looped-out-doj-panel-drug-war-killings.

Gavilan, Jodesz. ‘Despite UN promise, PH gov’t excludes CHR from “partial” drug war probe.’ *Rappler*. 11 January 2021. Accessed 24 March 2021. https://www.rappler.com/nation/chr-excluded-drug-war-review-panel-department-justice.

Gee, James Paul. *How to do Discourse Analysis: A Toolkit*. New York: Routledge, 2011.

Gerring, John. ‘What is a Case Study and What Is It Good for?’ *American Political Science Review* 98, no. 2 (2004): 341-354.

Gerring, John. *Case Study Research: Principles and Practices*. New York: Cambridge University Press, 2007.

Gifkins, Jess. ‘R2P in the UN Security Council: Darfur, Libya and beyond.’ *Cooperation and Conflict* 51, no. 2, (2016): 148-165.

Glanville, Luke. ‘Does R2P matter? Interpreting the impact of a norm.’ *Cooperation and Conflict* 51, no. 2 (2016): 184-199.

Glanville, Luke. ‘Intervention in Libya: From Sovereign Consent to Regional Consent.’ *International Studies Perspectives* 14 (2013): 325-342.

Global Centre for the Responsibility to Protect website. ‘About us’. Accessed 08 August 2020. https://www.globalr2p.org/about/.

Global Centre for the Responsibility to Protect. ‘Myanmar (Burma): Populations at risk.’ 15 September 2020. Accessed 21 October 2020. https://www.globalr2p.org/countries/myanmar-burma.

Global Centre for the Responsibility to Protect. ‘R2P Monitor. Issue 55.’ 15 January 2021. Accessed 31 January 2021. https://www.globalr2p.org/publications/r2p-monitor-issue-55-15-january-2021/.

Global Centre for the Responsibility to Protect. ‘R2P Monitor. Issue 56.’ 15 March 2021. Accessed 20 March 2021. https://www.globalr2p.org/publications/r2p-monitor-issue-56-15-march-2021/

Goldsmith, Jack. ‘Fire When Ready.’ *Foreign Policy*, 20 March 2020. Accessed 24 July 2020. https://foreignpolicy.com/2012/03/20/fire-when-ready/.

Goodenough, Patrick. 'Russia, China Accuse West of Exceeding UN Resolution, Making Libyan Crisis Worse.' *CNS News,* 29 March 2011. Accessed 25 March 2020. https://cnsnews.com/news/article/russia-china-accuse-west-exceeding-un-resolution-making-libyan-crisis-worse.

Government UK website. Accessed 29 January 2020. https://www.gov.uk/.

Grand, Camille. 'The French experience: Sarkozy's War?.' *Precision and Purpose: Airpower in the Libyan Civil War*, edited by Karl P. Mueller et al., 183-202. Santa Monica: RAND Corporation, 2015.

Guardian. 'Libya uprising – Monday 7 March.' 7 March 2011. Accessed 31 March 2020. https://www.theguardian.com/world/blog/2011/mar/07/libya-uprising-live-updates.

Guardian. 'Libya military action - Friday 18 March.' 18 march 2011. Accessed 8 February 2020. https://www.theguardian.com/world/blog/2011/mar/18/libya-military-action-live-updates.

Gutierrez, Jason. 'Philippines Officially Leaves the International Criminal Court.' *New York Times*, 17 March 2019. Accessed 21 March 2021. https://www.nytimes.com/2019/03/17/world/asia/philippines-international-criminal-court.html.

Guzman, Andrew T. and Timothy L. Meyer. 'Customary International Law in the 21st Century.' In *Progress in International Law,* edited by Russell A. Miller and Rebecca M. Bratspies, 197-218. Leiden: Martinus Nijhoff Publishers, 2008.

Guzman, Andrew T. and Timothy L. Meyer. 'International Common Law: The Soft Law of International Tribunals.' *Chicago Journal of International Law* 9, no. 2 (2009): 515-532.

Hansard HC Deb. vol.524 col.649, 7 March 2011. [Online]. Accessed 12 July 2020. https://publications.parliament.uk/pa/cm201011/cmhansrd/cm110307/debtext/110307-0001.htm#11030711000002.

Hansard HC Deb. vol.525 col.611, 18 March 2011. [Online]. Accessed 12 July 2020. https://publications.parliament.uk/pa/cm201011/cmhansrd/cm110318/debtext/110318-0001.htm#11031850000878.

Hansard HC Deb. vol.525 col.705, 21 March 2011. [Online]. Accessed 12 July 2020. https://publications.parliament.uk/pa/cm201011/cmhansrd/cm110321/debtext/110321-0001.htm#1103219000001.

Hardy, Cynthia, Bill Harley and Nelson Phillips. 'Discourse Analysis and Content Analysis: Two Solitudes?.' *Qualitative Methods* 2 (2004): 19-21.

Hastings, Sarah L. 'Triangulation.' In *Encyclopaedia of Research Design. Vol I.*, edited by Neil J. Salkind, 1537-1540. Thousand Oaks, CA: SAGE, 2010.

Hehir, Aidan. 'R2P and international law', *Critical Perspectives on the Responsibility to Protect*, edited by Philip Cunliffe, 84-99. London: Routledge, 2011.

Hehir, Aidan. 'The Permanence of Inconsistency: Libya, the Security Council, and the Responsibility to Protect.' *International Security* 38, no. 1 (2013): 137-159.

Hehir, Aidan. 'The Dog That Didn't Bark? A Response to Dunne and Gelber's Analysis of RtoP's Influence on the Intervention in Libya.' *Global Responsibility to Protect* 7, no. 2 (2015): 211-224.

Hehir, Aidan. 'Assessing the influence of the Responsibility to Protection the UN Security Council during the Arab Spring.' *Cooperation and Conflict* 51, no. 2 (2016): 166-183, 173.

Hehir, Aidan. *Hollow Norms and the Responsibility to Protect*. Hampshire: Palgrave Macmillan, 2019.

Heller, Kevin Jon. 'Ashley Deeks' Problematic Defense of the "Unwilling or Unable" Test'. *Opinio Juris,* 15 December 2011. Accessed 21 July 2020. http://opiniojuris.org/2011/12/15/ashley-deeks-failure-to-defend-the-unwilling-or-unable-test/.

High-level Panel on Threats, Challenges and Change. 'A More Secure World: Our Shared Responsibility.' A/59/565, 2 December 2004.

Hill, Matthew, David Campanale and Joel Gunter. '"Their goal is to destroy everyone": Uighur camp detainees allege systematic rape.' *BBC,* 2 February 2021. Accessed 20 March 2021. https://www.bbc.com/news/world-asia-china-55794071.

Horowitz, Irving. *Taking Lives: Genocide and State Power*. Fifth Edition. New Brunswick, NJ: Transaction Publishers, 2002.

House of Commons. 'United Nations Security Council Resolution 1973.' Order of Business Monday 21 March 2011, Session 2010-11, 19 March 2011 [Online]. Accessed 21 January 2021. https://publications.parliament.uk/pa/cm201011/cmagenda/ob110321.htm.

Hsieh, Hsiu-Fang and Sarah E. Shannon. 'Three Approaches to Content Analysis.' *Qualitative Health Research* 15, no. 9 (2005): 1277-1288; 1280; 1283.

Human Rights Watch. '"They Came To Kill" Escalating Atrocities in the Central African Republic.' 18 December 2013. Accessed 31 January 2021. https://www.hrw.org/report/2013/12/18/they-came-kill/escalating-atrocities-central-african-republic#.

Human Rights Watch. 'Killing Without Consequence: War Crimes, Crimes Against Humanity and the Special Criminal Court in the Central African Republic.' 5 July 2017. Accessed 31 January 2021. https://www.hrw.org/report/2017/07/05/killing-without-consequence/war-crimes-crimes-against-humanity-and-special.

Human Rights Watch. 'ASEAN: Don't Whitewash Atrocities Against Rohingya.' 19 June 2019. Accessed 31 January 2021. https://www.hrw.org/news/2019/06/19/asean-dont-whitewash-atrocities-against-rohingya?search=asean+divided+over++Rohingya.

Human Rights Watch. 'Philippines: New Anti-Terrorism Act Endangers Rights.' 5 June 2020. Accessed 24 March. https://www.hrw.org/news/2020/06/05/philippines-new-anti-terrorism-act-endangers-rights.

Human Rights Watch. 'Philippines: "Drug War" Killings Rise During Pandemic.' 13 January 2021. Accessed 24 March 2021. https://www.hrw.org/news/2021/01/13/philippines-drug-war-killings-rise-during-pandemic

Human Rights Watch. 'Philippines: Events of 2019.' (no date). Accessed 19 March 2021. https://www.hrw.org/world-report/2020/country-chapters/philippines#.

Hurrell, Andrew. 'Legitimacy and the Use of Force: Can the Circle Be Squared?.' *Review of International Studies*: *Force and Legitimacy in World Politics* 31 (2005): 15-32.

Independent. 'Duterte withdraws Philippines from International Criminal Court in stunning U-turn.' 14 March 2018. Accessed 21 March 2021. https://www.independent.co.uk/news/world/asia/duterte-philippines-icc-withdraw-war-drugs-international-criminal-court-extra-judicial-killings-a8256076.html.

International Commission on Intervention and State Sovereignty. 'The Responsibility to Protect: Report of the International Commission on Intervention and State Sovereignty.' Ottawa, ON, Canada: International Development Research Centre, 2001.

International Criminal Court. 'Preliminary examination: Republic of the Philippines.' Accessed 14 January 2021. https://www.icc-cpi.int/philippines.

International Criminal Court – The Office of the Prosecutor. 'Report on Preliminary Examination Activities 2020'. 14 December 2020. Accessed 23 March 2021. https://www.icc-cpi.int/itemsDocuments/2020-PE/2020-pe-report-eng.pdf.

International Law Association. Committee on the Formation of Customary (General) International Law, Report of the Sixty-Ninth Conference, 25-29th July, London, 2000. Accessed 20 October 2020. https://www.law.umich.edu/facultyhome/drwcasebook/Documents/Documents/ILA%20Report%20on%20Formation%20of%20Customary%20International%20Law.pdf.

Iyi, John-Mark. 'Emerging Powers and the Operationalisation of R2P in Africa: The Role of South Africa in the UNSC.' *African Journal of Legal Studies* 7, no. 1 (2014): 149-176.

Jackson, Richard. *Writing the war on terrorism: Language, politics and counter-terrorism*. Manchester: Manchester University Press, 2005.

Jackson, Richard. 'Constructing Enemies: "Islamic Terrorism" in Political and Academic Discourse.' *Government and Opposition* 42, no. 3 (2007): 394-426.

Jackson, Richard. 'Critical Discourse Analysis'. In *Critical Methods in Terrorism Studies*, edited by Priya Dixit and Jacob L. Stump, 77-90. Abingdon: Routledge, 2016.

Jakarta Post. 'Indonesia condemns Rakhine clashes.' 29 August 2017. Accessed 31 January 2021. https://www.thejakartapost.com/news/2017/08/29/indonesia-condemns-rakhine-clashes.html.

Jaworski, Adam and Nikolas Coupland. 'Introduction: Perspectives on discourse analysis.' In *The Discourse Reader*, Second Edition, edited by Adam Jaworski and Nikolas Coupland, 1-37. London: Routledge, 2006.

Jick, Todd D. 'Mixing Qualitative and Quantitative Methods: Triangulation in Action.' *Administrative Science Quarterly. Qualitative Methodology* 24, no. 4 (1979): 602-611.

Johnstone, Ian. 'UN Security Council Deliberations: The Power of the Better Argument.' *European Journal of International Law* 14, no. 3 (2003): 437-480.

Jones, Adam. *Genocide: A Comprehensive Introduction*, Second edition. New York: Routledge, 2011.

Jones, Adam. *Genocide: A Comprehensive Introduction*. New York: Routledge, 2006.

Jorgensen, Marianne and Louise J. Phillips. *Discourse Analysis as Theory and Method.* London: SAGE, 2002.

Juppé, Alain. 'Libya', Reply given by Alain Juppé to a question in the National Assembly, 1 March 2011. Accessed 17 January 2020. http://www.assemblee-nationale.fr/13/cri/2010-2011/20110128.asp.

Juppé, Alain. 'Situation in Libya.' Reply given by Alain Juppé to a question in the National Assembly. French National Assembly, 8 March 2011. Accessed 15 January 2020. http://www.assemblee-nationale.fr/13/cri/2010-2011/20110134.asp#P302_53875.

Juppé, Alain. 'Arab democracy movements – Migration/Union for the Mediterranean/investment/Middle East peace process – Libya.' Hearing of Alain Juppé before the National Assembly Foreign Affairs Committee (excerpts), 15 March

2011. Published on 23 March 2011. Accessed 23 September 2019. https://franceintheus.org/spip.php?article2244

Juppé, Alain. 'Press conference given by Alain Juppé, Ministre d'Etat, Minister of Foreign and European Affairs.' 17 March 2011, published on 21 March 2011. Accessed 23 September 2019. https://franceintheus.org/spip.php?article2235.

Juppé, Alain. 'Libya/Côte d'Ivoire.' Interview on Libya to France 2 television, 19 March 2011, Paris. Published on 21 March 2011. Accessed 28 January 2020. https://franceintheus.org/spip.php?article2238.

Juppé, Alain. 'London Conference on Libya/Consensus.' Press conference given by Alain Juppé (excerpts). 29 March 2011. Accessed 6 November 2019. https://uk.ambafrance.org/Alain-Juppe-s-press-conference,18821.

International Criminal Court, 'Rome Statute of the International Criminal Court'. The Hague: International Criminal Court, 2011. Accessed 20 March 2021. Available at: https://www.icc-cpi.int/resource-library/documents/rs-eng.pdf

Keating, Tom. 'The UN Security Council on Libya: Legitimation or Dissimulation.' In *Libya: The Responsibility to Protect and the Future of Humanitarian Intervention*, edited by Aidan Hehir and Robert Murray, 162-190. Basignstoke: Palgrave, 2013.

Koena, Jean Fernand. 'Central African Republic declares state of emergency.' *Washington Post,* 21 January 2021. Accessed 21 January 2021. https://www.washingtonpost.com/world/africa/central-african-republic-declares-state-of-emergency/2021/01/22/852fc84e-5c90-11eb-a849-6f9423a75ffd_story.html.

Koh, Harold Hongju. 'The Lawfulness of the U.S. Operation Against Osama bin Laden.' *Opinio Juris,* 19 May 2011. Accessed 26 July 2020. http://opiniojuris.org/2011/05/19/the-lawfulness-of-the-us-operation-against-osama-bin-laden/.

Kuperman, Alan J. 'Rethinking the Responsibility to Protect', *The Whitehead Journal of Diplomacy and International Relations* 10, no. 1 (2009): 33-43.

Kuperman, Alan J. ‘Obama's Libya debacle: how a well-meaning intervention ended in failure.’ *Foreign Affairs* 94, no. 2 (2015): 66-77.

Kurtz, Gerrit and Philipp Rotmann. ‘The Evolution of Norms of Protection: Major Powers Debate the Responsibility to Protect.’ *Global Society* 30, no. 1 (2016): 3-20.

Labonte, Melissa T. ‘Whose Responsibility to Protect? The Implications of Double Manifest Failure for Civilian Protection.’ *The International Journal of Human Rights* 16, no. 7 (2012): 982–1002.

Leaning, Jennifer. ‘Introduction.’ In *Humanitarian Crises: The Medical and Public Health Response*, edited by Jennifer Leaning, Susan M. Briggs and Lincoln C. Chen, 1-15. Cambridge, MA and London: Harvard University Press, 1999.

Leaning, Jennifer. ‘Human Security and Conflict.’ In *Risking Human Security: Attachment and Public Life*, edited by Marci Green, 125-150. Abingdon and New York: Routledge, 2008.

Leaning, Jennifer. ‘Enforced Displacement of Civilian Populations in War: A Potential New Element in Crimes against Humanity.’ *International Criminal Law Review* 11, no. 3 (2011): 445-462.

Leaning, Jennifer and Sam Arie. ‘Human Security: A Framework for Assessment in Conflict and Transition.’ *Harvard Center for Population and Development Studies, Working Paper Series* 11, no. 8 (2011): 1-70.

Lewis, Michael. ‘Obama’s Way.’ *Vanity Fair*, 11 September 2012. Accessed 31 May 2020. https://www.vanityfair.com/news/2012/10/michael-lewis-profile-barack-obama.

Lieberman, Benjamin. ‘“Ethnic Cleansing” versus Genocide.’ In *The Oxford Handbook of Genocide Studies*, edited by Donald Bloxham and A. Dirk Moses, 42-60. New York: Oxford University Press, 2010.

Lischin, Luke. ‘Surviving or thriving? COVID-19 and violent non-state actors in the Southern Philippines.’ *New mandala*, 9 September 2020. Accessed 5 February 2021. https://www.newmandala.org/surviving-or-thriving-covid-19-and-violent-non-state-actors-in-the-southern-philippines/.

Macdonald, Roderick A. ‘Three Metaphors of Norm Migration in International Context.’ *Brooklyn Journal of International Law* 34, no. 3 (2009): 603-653.

Makela, Finn. ‘The Drug Testing Virus.’ *Revue Juridique Themis* 43, no. 3 (2009): 651-703.

Marks, Stephen P. and Nicholas Cooper. ‘The Responsibility to Protect: Watershed or Old Wine in a New Bottle?’ *Jindal Global Law Review* 2, no. 1 (2010): 86-130.

McGreal, Chris. ‘Libyan rebels in retreat as Gaddafi attacks by air, land and sea.’ *Guardian*, 10 March 2011. Accessed 7 February 2020. https://www.theguardian.com/world/2011/mar/10/libya-gaddafi-takes-ras-lanuf.

McLeod, Ian. *Legal Method.* Second Edition. Basingstoke: Macmillan, 1996.

Miles, Matthew B. and A. Michael Huberman. *Qualitative Data Analysis*. Second Edition. Thousand Oaks, CA: SAGE, 1994.

Ministry of Foreign and European Affairs. ‘France/Libya/UNSCR 1973.’ Statements issued by the Ministry of Foreign and European Affairs Spokesperson (excerpts), 18 March 2011, Paris. Published on 21 March 2011. Accessed 29 January 2020. https://franceintheus.org/spip.php?article2236.

Mitzen, Jennifer. ‘Reading Habermas in anarchy: multilateral diplomacy and global public spheres.’ *American Political Science Review* 99, no. 3 (2005): 401-417.

Morris, Justin. ‘Libya and Syria: R2P and the spectre of the swinging pendulum.’ *International Affairs* 89, no. 5 (2013): 1265-1283.

Murray, Robert W. ‘Humanitarianism, Responsibility or Rationality? Evaluating Intervention as State Strategy.’ In *Libya: The Responsibility to Protect and the Future of Humanitarian Intervention*, edited by Aidan Hehir and Robert Murray, 15-33. Basignstoke: Palgrave, 2013.

National Assembly (France). 2nd session of 1 March. 1 March 2011. Accessed 17 January 2020. http://www.assemblee-nationale.fr/13/cri/2010-2011/20110128.asp.

National Assembly (France). 1st session of 2 March. 2 March 2011. Accessed 17 January 2020. http://www.assemblee-nationale.fr/13/cri/2010-2011/20110130.asp.

National Assembly (France). 1st session of 8 March. 8 March 2011. Accessed 23 August 2020. http://www.assemblee-nationale.fr/13/cri/2010-2011/20110134.asp.

National Assembly (France). 1st session of 22 March, 22 March 2011. Accessed 15 January 2020. http://www.assemblee-nationale.fr/13/cri/2010-2011/20110144.asp#INTER_0.

National Assembly (France). Liste des comptes rendus intégraux 2011. Accessed 29 January 2020. http://www.assemblee-nationale.fr/13/cri/2010-2011/.

Newman, Edward. 'Review Article – Humanitarian Intervention Legality and Legitimacy.' *The International Journal of Human Rights* 6, no. 4 (2002): 102-120.

O'Connell, Mary Ellen. 'Evidence of terror.' *Journal of Conflict and Security Law* 7, no. 1 (2002): 19-36.

O'Connell, Mary Ellen. 'Responsibility to Peace.' In *Critical Perspectives on the Responsibility to Protect*, edited by Philip Cunliffe, 71-83. London: Routledge, 2011.

O'Connell, Mary Ellen. 'Self-Defence, Pernicious Doctrines, Peremptory Norms.' In *Self-Defence against Non-State Actors*, Max Planck Trialogues, edited by Mary Ellen O'Connell, Christian J. Tams, and Dire Tladi, 174–257. Cambridge: Cambridge University Press, 2019.

Obama, Barack. 'Executive Order 13566–Libya.' The White House President Barack Obama. 25 February 2011. Accessed 23 September 2019. https://obamawhitehouse.archives.gov/the-press-office/2011/02/25/executive-order-13566-libya.

Obama, Barack. 'President Obama on Libya: "These Sanctions Therefore Target the Qaddafi Government, While Protecting the Assets that Belong to the People of Libya".*' The White House President Barack Obama*. 25 February 2011. Accessed 23 September 2019. https://obamawhitehouse.archives.gov/blog/2011/02/25/president-obama-libya-these-sanctions-therefore-target-qaddafi-government-while-prot.

Obama, Barack. 'Remarks by the President on Libya.' *The White House President Barack Obama*. 23 February 2011. Accessed 23 September 2019.

https://obamawhitehouse.archives.gov/the-press-office/2011/02/23/remarks-president-libya.

Obama, Barack. 'Remarks by President Obama and President Calderón of Mexico at Joint Press Conference.' *The White House President Barack Obama*. 3 March 2011. Accessed 14 February 2020. https://obamawhitehouse.archives.gov/the-press-office/2011/03/03/remarks-president-obama-and-president-calder-n-mexico-joint-press-confer.

Obama, Barack. 'Remarks by the President on the Situation in Libya.' *The White House President Barack Obama*. 18 March 2011. Accessed 25 January 2020. https://obamawhitehouse.archives.gov/the-press-office/2011/03/18/remarks-president-situation-libya.

Obama, Barack. 'Remarks by the President on Libya.' *The White House President Barack Obama*. 19 March 2011. Accessed 23 September 2019. https://obamawhitehouse.archives.gov/the-press-office/2011/03/19/remarks-president-libya.

Obama, Barack. 'Letter from the President regarding the commencement of operations in Libya.' *The White House President Barack Obama*. 21 March 2011. Accessed 23 September 2019. https://obamawhitehouse.archives.gov/the-press-office/2011/03/21/letter-president-regarding-commencement-operations-libya.

Obama, Barack. 'Remarks by the President in Address to the Nation on Libya.' *The White House President Barack Obama*. 28 March 2011. Accessed 21 January 2021. https://obamawhitehouse.archives.gov/the-press-office/2011/03/28/remarks-president-address-nation-libya.

Obama, Barack. 'Weekly Presidential Address: The Military Mission in Libya.' *The White House President Barack Obama*. 26 March 2011. Accessed 25 January 2020. https://obamawhitehouse.archives.gov/the-press-office/2011/03/26/weekly-address-president-obama-says-mission-libya-succeeding.

Obama, Barack. *A Promised Land.* New York: Penguin Random House, 2020.

Organization of the Islamic Conference. 'Final Communique Issued by the Emergency Meeting of the Committee of Permanent Representatives to the Organization of the

Islamic Conference on the Alarming Developments in Libyan Jamahiriya.' 8 March 2011. Accessed 31 March 2020. https://www.oic-oci.org/topic/?t_id=5022&ref=2110&lan=en.

Pak, Hui-Chol and Hye-Ryon Son, 'Military Intervention in Syria and the "Unwilling or Unable" Test: Lawful or Unlawful?.' Russian *Law Journal* 7, no. 4 (2019): 73-98.

Pape, Robert A. 'When Duty Calls: A Pragmatic Standard for Humanitarian Intervention.' *International Security* 37, no. 1 (2012): 41-80.

Patag, Kristine Joy. 'Duterte government has no interest in probing themselves – global group.' *PhilStar*. 28 January 2021. Accessed 21 March 2021. https://www.philstar.com/headlines/2021/01/28/2073673/duterte-government-has-no-interest-probing-themselves-global-group.

Pattison, James. 'Perilous Noninterventions? The Counterfactual Assessment of Libya and the Need to be a Responsible Power.' *Global Responsibility to Protect* 9, no. 2 (2017): 219-228.

Pavlova, Ludmila. 'The Concept "Responsibility to Protect": Analysis and Legal Evaluation.' *Journal of International Law and International Relations* 4, no. 67 (2013): 3-8.

Pinto, Rogerio. 'Mixed Methods Design.' In *Encyclopaedia of Research Design. Vol I.*, edited by Neil J. Salkind, 812-818. Thousand Oaks, CA: SAGE, 2010.

Pollentine, Marc. *Constructing the Responsibility to Protect*, Ph.D. thesis. Cardiff University, 2012.

Power, Samantha. *The Education of an Idealist*. London: HarperCollins, 2019.

Powers, Maggie. 'The Responsibility to Protect after Libya – dead, dying or thriving?' *Open Democracy*, 24 June 2014. Accessed 27 February 2020. https://www.opendemocracy.net/en/openglobalrights-openpage-blog/responsibility-to-protect-after-libya-dead-dying-or-thriving/.

Powers, Maggie. 'Responsibility to protect: dead, dying, or thriving?' *The International Journal of Human Rights* 19, no. 8 (2015): 1257-1278.

Puri, Hardeep Singh. *Perilous Interventions: The Security Council and the Politics of Chaos*. London: HarperCollins, 2016.

Ratner, Steven R. and Jason S Abrams. *Accountability for Human Rights Atrocities: Beyond the Nuremberg* Legacy. Second Edition. Oxford: Oxford University Press, 2001.

Reike, Ruben. 'Libya and the Responsibility to Protect: Lessons for "Pillar Three"'. In *Operationalizing the Responsibility to Protect: A Contribution to the Third Pillar Approach*, edited by Daniel Fiott, Robert Zuber and Joachim Koops, 73-79. Brussels: The Madariaga – College of Europe Foundation, Global Action to Prevent War, the Global Governance Institute and the International Coalition for the Responsibility to Protect, 2012.

Reinold, Theresa. 'State weakness, irregular warfare, and the right to self-defence post-9/11.' *American Journal of International Law* 105, no. 2 (2011): 244-286.

Reuters. 'Gaddafi launches counter-offensive on Libyan rebels.' 6 March 2011. Accessed 2 July 2020. https://www.reuters.com/article/us-libya-protests/gaddafi-launches-counter-offensive-on-libyan-rebels-idUSTRE71G0A620110306.

Reuters. 'Malaysia's dissent on Myanmar statement reveals cracks in ASEAN façade.' 25 September 2017. Accessed 31 January 2021. https://www.reuters.com/article/us-myanmar-rohingya-malaysia/malaysias-dissent-on-myanmar-statement-reveals-cracks-in-asean-facade-idUSKCN1C0124.

Reuters. 'Philippines sees "close and friendly" ties with Biden administration.' 20 January 2021. Accessed 24 March 2021. https://www.reuters.com/article/us-philippines-usa-idUSKBN29P0YN.

Reuters. 'U.N. "deeply worried" over Philippine killings, violent rhetoric.' 9 March 2021. Accessed 24 March 2021. https://www.reuters.com/article/us-philippines-rights-un/u-n-deeply-worried-over-philippine-killings-violent-rhetoric-idUSKBN2B11JG?il=0.

Rice, Susan. *Tough Love: My Story of the Things Worth Fighting For*. New York: Simon and Schuster, 2019.

Robertson, Phil. 'Another Spike in Philippines' "Drug War" Deaths.' *Human Rights Watch*, 28 September 2020. Accessed 21 March 2021. https://www.hrw.org/news/2020/09/28/another-spike-philippines-drug-war-deaths.

Rosenberg, Sheri and Ekkehard Strauss, 'A Common Standard for Applying R2P', APCR2P Brief, 2 (2), 2012.

Rosenberg, Sheri and Ekkehard Strauss. 'A Common Approach to the Application of the Responsibility to Protect.' In *Operationalizing the Responsibility to Protect: A Contribution to the Third Pillar Approach*, edited by Daniel Fiott, Robert Zuber and Joachim Koops, 55-72. Brussels: The Madariaga – College of Europe Foundation, Global Action to Prevent War, the Global Governance Institute and the International Coalition for the Responsibility to Protect, 2012.

Rosenberg, Sheri and Ekkehard Strauss. 'A Common Standard for Applying R2P.' *Cardozo Law School, Holocaust, Genocide, and Human Rights Program* (2013): 1-148.

Rosenberg, Sheri, Tibi Galis, and Alex Zucker. 'Introduction.' In *Reconstructing Atrocity Prevention*, edited by Sheri Rosenberg, Tibi Galis, and Alex Zucker, 1-16. New York: Cambridge University Press, 2016.

Rossman, Gretchen B. and Bruce L. Wilson. 'Numbers and Words: Combining Quantitative and Qualitative Methods in a Single Large-Scale Evaluation Study.' *Evaluation Review* 9, no. 5 (1985): 627-643.

Ruzzene, Attilia. 'Drawing Lessons from Case Studies by Enhancing Comparability.' *Philosophy of the Social Sciences* 42, no. 1 (2012): 99-120.

Santos, Eimor. '62 groups ask UN for independent probe on PH killings, human rights abuses.' *CNN Philippines*, 28 August 2020. Accessed 24 March 2021. https://www.cnnphilippines.com/news/2020/8/28/philippines-drug-war-united-nations-human-rights-council.html.

Sarkozy, Nicolas. 'Situation in Libya.' Communiqué issued by the Presidency of the Republic. 21 February 2011, Paris. Accessed 17 January 2020. https://franceintheus.org/spip.php?article2161.

Sarkozy, Nicolas. Statement at the Council of Ministers on the crisis in Libya. 23 February 2011, Paris. Accessed 17 January 2020. https://franceintheus.org/spip.php?article2176.

Sarkozy, Nicolas. 'Situation in Libya and the Mediterranean – Extraordinary European Council – Press conference given by Nicolas Sarkozy, President of the Republic (excerpts).' 11 March 2011, Brussels. Accessed 12 November 2019. https://franceintheus.org/spip.php?article2223.

Sarkozy, Nicolas. 'Paris Summit for the Support of the Libyan People – Statement by Nicolas Sarkozy, President of the Republic.' 19 March 2011. Accessed 23 September 2019. https://franceintheus.org/spip.php?article2241

Scharf, Michael P. 'How the War Against ISIS Changed International Law.' *Case Western Reserve Journal of International Law* 48 (2016): 1-54.

Schmitt, Eric. 'U.S. "Gravely Concerned" Over Violence in Libya.' *New York Times,* 20 February 2011. Accessed 29 March 2020. https://www.nytimes.com/2011/02/21/world/middleeast/21diplomacy.html.

Schmitt, Michael N. '"Change Direction" 2006: Israeli Operations in Lebanon and the International Law of Self-Defense.' *Michigan Journal of International Law* 29, no. 2 (2009): 127-164.

Security Council Report. 'Update Report No. 3: Libya.' 25 February 2011. Accessed 4 February 2020. https://www.securitycouncilreport.org/update-report/lookup_c_glkwlemtisg_b_6586331.php.

Security Council Report. 'Update report No. 1: Libya.' 14 March 2011. Accessed 7 February 2020. https://www.securitycouncilreport.org/update-report/lookup_c_glKWLeMTIsG_b_6621881.php.

Security Council Report. 'What's in Blue: Insights on Libya.' 14 March 2011. Accessed 7 February 2020. https://www.whatsinblue.org/2011/03/insights-on-libya-3.php.

Security Council Report. 'Special Research Report: UN Sanctions, No. 3.' 25 November 2013. Accessed 21 January 2021.

https://www.securitycouncilreport.org/atf/cf/%7B65BFCF9B-6D27-4E9C-8CD3-CF6E4FF96FF9%7D/special_research_report_sanctions_2013.pdf.

Security Council Report. 'Chronology of Events: Libya.' 2 August 2019 (revised). Accessed 8 February 2020. https://www.securitycouncilreport.org/chronology/libya.php.

Shaw, Martin. *What is Genocide?*. Cambridge: Polity Press, 2007.

Sill, Rudra and Peter Katzenstein. 'Analytic Eclecticism in the Study of World Politics: Reconfiguring Problems and Mechanisms across Research Traditions', *Perspectives on Politics* 8, no. 2 (2010): 411-431.

Sjöstedt, Britta. 'Applying the Unable/Unwilling State Doctrine – Can a State Be Unable to Take Action?' In *Self-Defence Against Non-State Actors: Self-Defence Against Non-State Actors: Impulses from the Max Planck Trialogues on the Law of Peace and War. Research Paper No. 2017-07*, edited by Anne Peters and Christian Marxsen, 35-38. Max Planck Institute for Comparative Public Law & International Law (MPIL), 2017. Accessed 16 July 2020. http://dx.doi.org/10.2139/ssrn.2941640.

Skinner, Quentin. 'Some Problems in the Analysis of Political Thought and Action.' *Political Theory* 2, no. 3 (1974): 277-303.

Spandler, Kilian. 'Lessons from ASEAN's Rakhine response.' *East Asia Forum*, 5 February 2020. Accessed 26 October 2020. https://www.eastasiaforum.org/2020/02/05/lessons-from-aseans-rakhine-response/.

Stake. Robert E. *The Art of Case Study Research*. Thousand Oaks, CA: SAGE, 1995.

Talabong, Rambo and Jodesz Gavilan. 'PNP has a human rights office, but what has it done?' *Rappler*, 26 November 2020. Accessed 24 March 2021. https://www.rappler.com/newsbreak/in-depth/philippine-national-police-human-rights-affairs-office-what-has-it-done.

Taylor, Andrew. *Failed States*. Basingstoke: Palgrave Macmillan, 2013.

Teitt, Sarah. 'The Responsibility to Protect and China's Peacekeeping Policy.' *International Peacekeeping* 18, no. 3 (2011): 298-312.

Thakur, Ramesh. ‘R2P after Libya and Syria: Engaging emerging powers.’ *The Washington Quarterly* 36, no. 2 (2013): 61-76.

True-Frost, Cora. ‘The Security Council and Norm Consumption.’ *International Law and Politics* 40 (2007): 115-217.

Tupas, Emmanuel. ‘13 killed daily during quarantine.’ *PhilStar*. 13 September 2020. Accessed 24 March 2020. https://www.philstar.com/nation/2020/09/13/2042020/13-killed-daily-during-quarantine.

Turner, Scott F., Laura B. Cardinal and Richard M. Burton. ‘Research Design for Mixed Methods: A Triangulation-based Framework and Roadmap.’ *Organizational Research Methods* 20, no. 2 (2017): 243-267.

UK Parliament website. Accessed 29 January 2021. https://www.parliament.uk/.

Ulfstein, Geir and Hege Føsund Christiansen. ‘The legality of the NATO bombing in Libya.’ *International & Comparative Law Quarterly* 62, no. 1 (2013): 159-171.

UN News. ‘UN chief urges restraint by Arab leaders as protests continue.’ 21 February 2011. Accessed 4 February 2020. https://news.un.org/en/story/2011/02/367112-un-chief-urges-restraint-arab-leaders-protests-continue.

UN News. ‘Ban strongly condemns Qadhafi’s actions against protesters, calls for punishment.’ 23 February 2011. Accessed 4 February 2020. https://news.un.org/en/story/2011/02/367342-ban-strongly-condemns-qadhafis-actions-against-protesters-calls-punishment.

United Nations. Charter of the United Nations, 24 October 1945. Accessed 21 January 2021. https://www.un.org/en/charter-united-nations/.

United Nations. ‘UN Secretary-General Special Adviser on the Prevention of Genocide, Francis Deng, and Special Adviser on the Responsibility to Protect, Edward Luck, on the Situation in Libya.’ New York, 22 February 2011. 17 January 2021. http://www.un.org/en/genocideprevention/documents/media/statements/2011/English/2011-02-22-OSAPG,%20Special%20Advisers%20Statement%20on%20Libya,%2022%20February%202011.pdf.

United Nations. Index to proceedings of the Security Council: 66th year, 2011. Accessed 05 November 2019. http://research.un.org/en/docs/sc/quick/meetings/2011.

United Nations. 'Framework of Analysis for Atrocity Crimes: A tool for prevention.' 2014. Accessed 27 July 2020. https://www.refworld.org/docid/548afd5f4.html.

United Nations General Assembly. Convention on the Prevention and Punishment of the Crime of Genocide, A/RES/260 (III), 9 December 1948.

United Nations General Assembly. 2005 World Summit Outcome, A/RES/60/1, 24 October 2005.

United Nations General Assembly. A/RES/65/265. 1 March 2011.

United Nations General Assembly. 'Promotion and protection of human rights in the Philippines.' Resolution adopted by the Human Rights Council on 11 July 2019, A/HRC/RES/41/2, 17 July 2019.

United Nations General Assembly. A/HRC/RES/45/33, 13 October 2020.

United Nations Human Rights Council, 'Report of the United Nations High Commissioner for Human Rights on the situation of human rights in the Philippines.' Advance Edited Version, A/HRC/44/22, 04 June 2020.

United Nations Human Rights Office of the High Commissioner press release. 'Pillay calls for international inquiry into Libyan violence and justice for victims.' 22 February 2011. Accessed 4 February 2020. https://newsarchive.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=10743&LangID=E.

United Nations Human Rights Office of the High Commissioner press release. 'UN Human Rights Council recommends suspension of Libya.' 25 February 2011. Accessed 4 February 2020. https://www.ohchr.org/EN/NewsEvents/Pages/HRCSpecialSessionLibya.aspx.

United Nations Human Rights Office of the High Commissioner. 'Situation of Human Rights in the Libyan Arab Jamahiriya: Statement by Navi Pillay, UN High Commissioner for Human Rights', Human Rights Council - 15th Special Session.

Geneva, 25 February 2011. Accessed 17 January 2021. https://newsarchive.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=10760&LangID=E.

United Nations Security Council, Provisional Rules of Procedure of the Security Council, S/96/Rev.7, 1 January 1983.

United Nations Security Council, 'Interim Report of the Commission of Experts Established Pursuant to Security Council Resolution 780 (1992)', S/25274. 26 January 1993, Annex I.

United Nations Security Council, 'Final Report of the Commission of Experts Established Pursuant to Security Council Resolution 780 (1992)', S/1994/674, 27 May 1994, Annex.

United Nations Security Council. S/PV.6486, 22 February 2011.

United Nations Security Council. S/PV.6490, 25 February 2011.

United Nations Security Council. S/PV.6491, 26 February 2011.

United Nations Security Council. S/RES/1970, 26 February 2011.

United Nations Security Council. 'Council of the League of Arab States, Resolution No. 7360.' Extraordinary session, 12 March 2011. S/2011/137. 15 March 2011.

United Nations Security Council. S/2011/142, 17 March 2011.

United Nations Security Council. S/PV.6498, 17 March 2011.

United Nations Security Council. S/RES/1973, 17 March 2011.

United Nations Security Council. S/PV.6528, 4 May 2011.

United Nations Security Council. S/PV.6531 (Resumption 1), 10 May 2011.

United Nations Security Council. S/PV.6531, 10 May 2011.

United Nations Security Council. S/PV.6620, 16 September 2011.

United Nations Security Council. S/RES/2009, 16 September 2011.

United Nations Security Council. S/2011/612, 4 October 2011.

United Nations Security Council. S/PV.6627, 4 October 2011.

United Nations Security Council. S/RES/2016, 27 October 2011.

United Nations Security Council. S/RES/2017, 31 October 2011.

United Nations Security Council, S/PV.6650 (Resumption 1), 9 November 2011.

United Nations Security Council. S/PV.6650, 9 November 2011.

United Nations Security Council. S/RES/2022, 2 December 2011.

United Nations Security Council. S/PV.6711, 4 February 2012.

United Nations Security Council. S/2012/77, 4 February 2012.

United Nations Security Council. S/PV.6810, 19 July 2012.

United Nations Security Council. S/2012/538, 19 July 2012.

United Nations Security Council. S/RES/2127, 5 December 2013.

United Nations Security Council. S/RES/2134, 28 January 2014.

United Nations Security Council. 'Letter dated 23 September 2014 from the Permanent Representative of the United States of America to the United Nations addressed to the Secretary-General', S/2014/695, 23 September 2014.

United Nations Security Council. S/RES/2249, 20 November 2015.

United Nations Security Council. S/RES/2399, 30 January 2018.

United Nations Security Council, 'UN Security Council Press Statement on Libya, SC/10180-AFR/2120.' 22 February 2011. Accessed 4 February 2020. https://www.un.org/press/en/2011/sc10180.doc.htm.

United Nations Security Council. 'Actions with Respect to Threats to the Peace, Breaches of the Peace, and Acts of Aggression.' Accessed 12 May 2020. https://www.un.org/securitycouncil/content/repertoire/actions.

Valente, Catherine. ‘Duterte orders military to ‘shoot and kill’ armed communist rebels.’ *Manila Times*, 6 March 2021. Accessed 24 March 2021. https://www.manilatimes.net/2021/03/06/news/duterte-orders-military-to-shoot-and-kill-armed-communist-rebels/847906/.

Valentino, Benjamin A. *Final Solutions, Mass Killing and Genocide in the 20th Century*. New York: Cornell University Press, 2004.

van Dijk, Teun. ‘What is Political Discourse Analysis?’ In *Linguistics*, edited by Jan Blommaert and Chris Bulcaen, 12-52. Amsterdam: Benjamins, 1987.

Villamor, Felipe. ‘Philippines Plans to Withdraw From International Criminal Court.’ *New York Times*, 14 March 2018. Accessed 21 March 2021. https://www.nytimes.com/2018/03/14/world/asia/rodrigo-duterte-philippines-icc.html.

Voeten, Erik. ‘The Political Origins of the UN Security Council’s Ability to Legitimize the Use of Force.’ *International Organization* 59 (2005): 527-557.

Walzer, Michael. *Just and Unjust Wars: A Moral Argument with Historical Illustrations.* Fourth Edition. New York: Basic Books, 2006.

Watt, Nicholas and Patrick Wintour. ‘Libya no-fly zone call by France fails to get David Cameron’s backing.’ *Guardian*, 23 February 2011. Accessed 29 January 2020. https://www.theguardian.com/world/2011/feb/23/libya-nofly-zone-david-cameron.

Watts, Sir Arthur. ‘The Importance of International Law.’ In *The Role of Law in International Politics: Essays in International Relations and International Law*, edited by Michael Byers, 5-16. New York: Oxford University Press, 2001.

Weaver, Catherine, ‘IPE’s Split Brain’, *New Political Economy* 12, no. 3 (2009): 337-346.

Weiner, Allen S. ‘The Use of Force and Contemporary Security Threats: Old Medicine for New Ills?’ *Stanford Law Review* 59, no. 2 (2006): 415-504.

Welsh, Jennifer. ‘Civilian Protection in Libya: Putting Coercion and Controversy Back into RtoP.’ *Ethics and International Affairs* 25, no. 3 (2011): 255-262.

Welsh, Jennifer M. 'Norm Contestation and the Responsibility to Protect.' *Global Responsibility to Protect* 5, no. 4 (2013): 365-396.

Welsh, Jennifer. 'R2P's Next Ten Years: Deepening and Extending the Consensus.' In *The Oxford Handbook of the Responsibility to Protect*, edited by Alex J. Bellamy and Time Dunne, 984-1000. Oxford: Oxford University Press, 2016.

Wheeler, Nicholas. *Saving Strangers: Humanitarian Intervention in International Society*. New York: Oxford University Press, 2000.

White House Barack Obama website. Accessed 29 January 2020. https://obamawhitehouse.archives.gov/.

White House Office of the Press Secretary. 'Readout of President Obama's Call with Chancellor Angela Merkel of Germany.' 26 February 2011. Accessed 14 February 2020. https://obamawhitehouse.archives.gov/the-press-office/2011/02/26/readout-president-obamas-call-chancellor-angela-merkel-germany.

White House Office of the Press Secretary. 'Readout of President Obama's meeting with his national security team on Libya.' 15 March 2011. Accessed 7 January 2020. https://obamawhitehouse.archives.gov/the-press-office/2011/03/15/readout-president-obamas-meeting-his-national-security-team-libya.

Widmaier, Wesley and Luke Glanville. 'The benefits of norm ambiguity: constructing the responsibility to protect across Rwanda, Iraq and Libya.' *Contemporary Politics* 21, no. 4 (2015): 367-383.

Williams, Paul D. 'Briefing: The Road to Humanitarian War in Libya.' *Global Responsibility to Protect* 3, no. 2 (2011): 248-259.

Wilmshurst, Elizabeth. 'The Chatham House Principles of International Law on the Use of Force by States in Self –Defence.' *The International and Comparative Law Quarterly* 55, no. 4 (2006): 963-972.

Yin, Robert K. *Case Study Research and Applications: Design and Methods*. Sixth Edition. Thousand Oaks, CA: SAGE, 2018.

Zwingel, Suzanne. ‘How Do Norms Travel? Theorizing International Women’s Rights in Transnational Perspective.’ *International Studies Quarterly* 56, no. 1 (2012): 115–129.