

2019

Challenging and Refining the "Unwilling or Unable" Doctrine

Craig Martin

Follow this and additional works at: <https://scholarship.law.vanderbilt.edu/vjtl>



Part of the [International Law Commons](#), and the [Military, War, and Peace Commons](#)

Recommended Citation

Craig Martin, Challenging and Refining the "Unwilling or Unable" Doctrine, 52 *Vanderbilt Law Review* 387 (2021)

Available at: <https://scholarship.law.vanderbilt.edu/vjtl/vol52/iss2/3>

This Article is brought to you for free and open access by Scholarship@Vanderbilt Law. It has been accepted for inclusion in Vanderbilt Journal of Transnational Law by an authorized editor of Scholarship@Vanderbilt Law. For more information, please contact mark.j.williams@vanderbilt.edu.

Challenging and Refining the “Unwilling or Unable” Doctrine

Craig Martin*

ABSTRACT

This Article challenges and proposes refinements to the “unwilling or unable” doctrine. Governments after 9/11 have invoked the doctrine to justify the use of force in self-defense against non-state actors (NSAs) operating within the territory of nonconsenting states. Responding to criticism that it lacked substance and a legal foundation, Daniel Bethlehem famously developed more detailed principles to embed the policy firmly in law, strike a balance between the interests of target states and territorial states, and bridge the gap between scholars and policy makers. His principles were embraced by governments as reflecting custom. The effort was laudable, but the principles fell short of their objective, and they create a risk of destabilizing the jus ad bellum regime.

This Article notes that the principles do not reflect custom, and it examines some of the ways in which they are inconsistent with the established understanding of the jus ad bellum regime. Specifically, they: lower the threshold for what constitutes an armed attack; eviscerate the temporal component from the concept of imminence, thereby destabilizing the core principle of necessity; improperly import the law of state responsibility into the jus ad bellum analysis; and undermine the independence of the international humanitarian law (IHL) and the jus ad bellum regimes. Finally, the principles do not provide sufficient guidance

* Professor of Law and Co-Director of the International and Comparative Law Center, Washburn University School of Law. This Article was presented during its development at the Law and Society Association Annual Conference in Mexico City, the ASIL Mid-Year Conference in St. Louis, and the AALS Annual Conference in San Diego, as well as smaller workshops along the way, and I am grateful for the many comments and suggestions made by fellow-panel members and conference participants, as well as many colleagues who have provided comments and input at various stages of the project's development. I would like to thank in particular, Scott Anderson, Bill Burke-White, Cindy Buys, Lois Chiang, Ed Cottrill, Benson Cowan, Cosette Creamer, Ashley Deeks, Jacques Delisle, Matt Erie, Brian Farrell, Eric Feldman, Craig Forcese, Stuart Ford, Terry Gill, Markus Gunneflo, Adil Haque, Kevin Jon Heller, Vladyslav Lanovoy, Asaf Lubin, Tom McDonnell, Mark Nevitt, Ved Nanda, Laurent Sacharoff, Leila Sadat, Yahli Shereshevsky, Freddy Sourgens, Milena Sterio, Jeremy Telman, and Tim Webster. I am of course responsible for any errors.

on how or by whom a range of key determinations are to be made, particularly regarding the “ability” or “unwillingness” of the territorial state. The principles lump all these determinations together and suggest that they may all be made unilaterally by the target state, governed only by a single, low reasonableness standard. All of this weakens the constraints of the jus ad bellum regime more generally, thus raising the risk of inter-state war.

The Article takes seriously the operational imperatives in dealing with the threat posed by terrorist organizations but proposes refinements to the principles to address each of these problems, so as to achieve greater consistency with established principles of the jus ad bellum regime. It develops new ideas on imminence, and drawing upon theories of self-judgment in international law, it disaggregates the decisions that have to be made and proposes differentiated standards to govern their execution and later assessment.

TABLE OF CONTENTS

I.	INTRODUCTION.....	389
II.	THE LAWS OF WAR AND ORIGINS OF THE DOCTRINE	394
	A. Jus ad Bellum Regime Assumptions	395
	B. Jus ad Bellum and IHL	399
	C. Background and Origins of the Unwilling or Unable Doctrine	401
	D. The Bethlehem Principles	406
III.	CRITIQUE OF THE DOCTRINE	412
	A. The Myth of Custom	412
	B. The Perversion of Imminence.....	415
	C. Incoherence on Armed Attack	423
	D. The Conflation of Jus ad Bellum with other Regimes.....	427
	E. Self-Judging Armed Attack, Necessity, and Unwillingness.....	435
IV.	REFINING THE DOCTRINE.....	442
	A. Restoring the Integrity of Jus ad Bellum	442
	B. Differentiated Standards for Distinct Target State Decisions	446
	C. Minimum Good Faith Standard for All Self- Judgment.....	452
	D. Transparency and Accountability	457
V.	CONCLUSION	459

I. INTRODUCTION

In the aftermath of the terrorist attacks on the United States on September 11, 2001 (the 9/11 attacks), the United States and other Western states developed legal justifications for the use of force against transnational terrorist organizations operating from within the territory of weak or sympathetic states. The primary justification articulated by the United States for using lethal force against members of the terrorist organizations themselves, as distinct from the states within which they were operating, has been that the United States is in a “transnational non-international armed conflict” with al-Qaeda, the Taliban, and associated forces.¹ This formulation soon encompassed such disparate groups as al-Qaeda in the Arabian Peninsula (AQAP), al-Shabab, Boko Haram, and Daesh (also known as the Islamic State in Syria or ISIS), to name a few.² But this justification, even if accepted as entirely valid, only provided authority under international humanitarian law (IHL) for the use of lethal force

1. Initially the United States argued that the conflict with al-Qaeda and associated forces constituted a war to which the law of armed conflict did not apply, but the Supreme Court rejected such arguments and held that it was a non-international armed conflict to which Common Article 3 of the Geneva Conventions applied. See *Hamdan v. Rumsfeld*, 548 U.S. 557, 628–31 (2006). It was subsequently called a “transnational non-international armed conflict,” or an “internationalized non-international armed conflict,” in an attempt to distinguish it from the traditional understanding of non-international armed conflict as being internal to a state. See generally David E. Graham, *Defining Non-International Armed Conflict: A Historically Difficult Task*, 88 NON-INTERNATIONAL ARMED CONFLICT IN THE TWENTY-FIRST CENTURY 43 (2012). This distinction has been roundly rejected, however. See generally KUBO MACAK, *INTERNATIONALIZED ARMED CONFLICTS IN INTERNATIONAL LAW* (2018); Anne Quintin, *Symposium: Reflections on Conflict Classification*, OPINIO JURIS (Jan. 16, 2019), <http://opiniojuris.org/2019/01/16/symposium-reflections-on-conflict-classification/> [<https://perma.cc/YXC9-H6NP>] (archived Feb. 5, 2019).

2. THE WHITE HOUSE, *REPORT ON THE LEGAL AND POLICY FRAMEWORKS GUIDING THE UNITED STATES’ USE OF MILITARY FORCE AND RELATED NATIONAL SECURITY OPERATIONS* 3–7 (2016), https://www.justsecurity.org/wp-content/uploads/2016/12/framework.Report_Final.pdf [<https://perma.cc/63VE-2MXC>] (archived Feb. 5, 2019) [hereinafter US LEGAL AND POLICY FRAMEWORKS]. For analysis of the US LEGAL AND POLICY FRAMEWORKS Report, see Benjamin Wittes, *The White House Releases a “Report on the Legal Policy Frameworks” on American Use of Military Force*, LAWFARE (Dec. 5, 2016), <https://www.lawfareblog.com/white-house-releases-report-legal-and-policy-frameworks-american-uses-military-force> [<https://perma.cc/QK6V-VLXB>] (archived Feb. 5, 2019); see also Rebecca Ingber, *Co-Belligerency*, 42 YALE J. INT’L L. 67, 68 (2017); Robert Chesney, *A Daisy Chain of Associated Forces? On the Potential Use of Force in Niger Against al-Mourabitoun*, LAWFARE (Oct. 26, 2017), <https://www.lawfareblog.com/daisy-chain-associated-forces-potential-use-force-niger-against-al-mourabitoun> [<https://perma.cc/3TZS-57NW>] (archived Feb. 5, 2019).

against the members of these armed groups in certain circumstances. It did not answer growing objections that firing missiles at people within the territory of other sovereign states, without the consent of those states, violated the international law rights of those states.³ It was argued that such strikes constituted a use of force against the nonconsenting state, and unless it came within one of the two permissible exceptions, was in violation of the prohibition on the use of force in Article 2(4) of the United Nations (UN) Charter.⁴ A justification grounded in the *jus ad bellum* regime, which governs the use of force against states, was thus needed to accompany the IHL rationales in defense of the lethal operations against non-state armed groups (NSAs).⁵

The United States, soon followed by other states, turned to an old doctrine from the pre-UN system of international law, now famously known as the “unwilling or unable” doctrine, to justify this use of force against NSAs operating within nonconsenting states.⁶ While the doctrine has its origin in much older neutrality law,⁷ it was reformulated to apply to the circumstances of states that are the victims of armed attacks (referred to here as target states) mounted by NSAs from within the territory of some other state (territorial states).⁸ The updated doctrine suggested that the target states have the right, under the doctrine of self-defense in the *jus ad bellum* regime, to use force against the NSAs within that territorial state in response to

3. There is a massive literature on this issue. See, e.g., Philip Alston (Special Rapporteur on Extrajudicial, Summary, or Arbitrary Executions), *Study on Targeted Killings*, UN Doc. A/HRC/14/24/Add.6, 12–15 (May 28, 2010) [hereinafter *Alston Study*]; NOAM LUBELL, EXTRATERRITORIAL USE OF FORCE AGAINST NON-STATE ACTORS 25–42 (2005); NILS MELZER, TARGETED KILLING IN INTERNATIONAL LAW 51–55 (2008); Craig Martin, *Going Medieval: Targeted Killing, Self-Defence, and the Jus ad Bellum Regime*, in TARGETED KILLING: LAW AND MORALITY IN AN ASYMMETRICAL WORLD (Claire Finkelstein & Jens Ohlin eds., 2012) [hereinafter *Martin Going Medieval*]; Mary Ellen O’Connell, *Unlawful Killing with Combat Drones: A Case Study of Pakistan, 2004–2009*, in SHOOTING TO KILL: THE LAW GOVERNING LETHAL FORCE IN CONTEXT (Simon Bronitt ed., 2012); TOM RUYTS, ‘ARMED ATTACK’ AND ARTICLE 51 OF THE UN CHARTER: EVOLUTIONS IN CUSTOMARY LAW AND PRACTICE 368–510 (2010); Christof Heyns et al., *The International Law Framework Regulating the Use of Drones*, 65 INT’L & COMP. L.Q. 791, 800–05 (2016); David Kretzmer, *Targeted Killing of Suspected Terrorists: Extra-Judicial Executions or Legitimate Means of Defence*, 16 EUR. J. INT’L L. 171, 186–88 (2005).

4. Martin, *Going Medieval*, *supra* note 3, at 224.

5. See *infra* Part II.C & D, for details of this development.

6. Elena Chachko & Ashley Deeks, *Who is on Board with “Unwilling or Unable”?*, LAWFARE (Oct. 10, 2016), <https://www.lawfareblog.com/who-board-unwilling-or-unable> [<https://perma.cc/QTE4-BLLG>] (archived Feb. 27, 2019).

7. Ashley S. Deeks, “Unwilling or Unable”: *Toward a Normative Framework for Extraterritorial Self-Defence*, 52 VA. J. INT’L L. 483, 505 (2012) [hereinafter *Deeks Unwilling*].

8. See *infra* Part II.C.

imminent or actual armed attacks, so long as the territorial state is unwilling or unable to prevent the NSA attacks, and is unwilling to consent to the target state using force to do so itself.⁹

The killing of Osama Bin Laden in Pakistan may be the most famous instance of such a use of force against members of an NSA located in a nonconsenting state (though there remains controversy over whether Pakistan had secretly consented to the raid).¹⁰ But the targeted killing program of the United States has involved missile strikes against members of organized armed groups and terrorist organizations in the tribal regions of Pakistan, Yemen, Somalia, Sudan, Syria, Libya, and Mali, to name just the best known theatres.¹¹ These were countries in which the United States was not yet involved as a belligerent in an ongoing armed conflict (in contrast to its role in Afghanistan and post-occupation Iraq), and in many instances the governments of these territorial states objected to the strikes.¹² The United States thus required some *jus ad bellum* justification for the use of force, and it increasingly invoked the unwilling or unable doctrine, as part of a broader claim of self-defense, for this purpose.¹³

9. See *infra* Part II.D.

10. Cora Currier et al., *What the Snowden Files Say About the Osama Bin Laden Raid*, THE INTERCEPT (May 18, 2015), <https://theintercept.com/2015/05/18/snowden-osama-bin-laden-raid/> [<https://perma.cc/5PJN-4E8N>] (archived Feb. 5, 2019); Husain Haqqani, *What Pakistan Knew About the Bin Laden Raid*, FOREIGN POLICY (May 31, 2015), <https://foreignpolicy.com/2015/05/13/what-pakistan-knew-about-the-bin-laden-raid-seymour-hersh/> [<https://perma.cc/ALG9-VCAS>] (archived Feb. 5, 2019); Seymour M. Hersh, *The Killing of Osama Bin Laden*, LONDON REV. OF BOOKS (May 21, 2015), <https://www.lrb.co.uk/v37/n10/seymour-m-hersh/the-killing-of-osama-bin-laden> [<https://perma.cc/F3NK-BFEV>] (archived Feb. 5, 2019); Eric Schmitt et al., *U.S. Was Braced for Fight With Pakistanis in Bin Laden Raid*, N.Y. TIMES (May 10, 2011), <https://www.nytimes.com/2011/05/10/world/asia/10intel.html> [<https://perma.cc/2KL5-XDKB>] (archived Feb. 5, 2019).

11. ROSA BROOKS, HOW EVERYTHING BECAME WAR AND THE MILITARY BECAME EVERYTHING: TALES FROM THE PENTAGON 12, 112 (2016); see also GEN. JOHN P. ABIZAID & ROSA BROOKS, RECOMMENDATIONS AND REPORT OF THE TASK FORCE ON U.S. DRONE POLICY 19 (2d ed. 2015).

12. See O'Connell, *supra* note 3; Mark Mazzetti, *U.S. Intensifying a Secret Campaign of Yemen Airstrikes*, N.Y. TIMES (June 8, 2011), <https://www.nytimes.com/2011/06/09/world/middleeast/09intel.html> [<https://perma.cc/5DLT-M2RB>] (archived Feb. 6, 2019); David Sanger & Eric Schmitt, *As Rift Deepens, Kerry Has a Warning for Pakistan*, N.Y. TIMES (May 15, 2011), <https://www.nytimes.com/2011/05/15/us/politics/15dipl.html> [<https://perma.cc/HEL7-9NYR>] (archived Feb. 6, 2019).

13. See, e.g., Eric Holder, Att'y Gen., Address at Northwestern University School of Law (Mar. 5, 2012), <https://www.justice.gov/opa/speech/attorney-general-eric-holder-speaks-northwestern-university-school-law> [<https://perma.cc/EAU3-U3QN>] (archived Feb. 6, 2019); Harold Koh, U.S. Dep't of State Legal Adviser, The Obama Administration

But as the United States (among others) increasingly relied upon the doctrine, the doctrine was also more frequently criticized as having little grounding in law, and as having none of the actual substance necessary to shape the decisions or in any way constrain the actions of powerful target states.¹⁴

In defense of the doctrine, efforts were made to provide it with a more solid grounding in the *jus ad bellum* regime, and to infuse it with substantive principles that would stipulate the conditions under which target states could justifiably use force against NSAs in nonconsenting territorial states.¹⁵ This effort was led most famously by a former British government official named Daniel Bethlehem, who in 2012 published a set of sixteen principles, now commonly referred to as the “Bethlehem Principles,” with the explicit goal of providing a more sound legal foundation for the doctrine.¹⁶ Many governments and policy makers, and indeed some scholars, quickly embraced the doctrine so defined as representing the current state of customary international law.¹⁷

I applaud the spirit and purpose of Bethlehem’s efforts, and I acknowledge his admonition that scholars must better understand the reality within which states are now operating and take seriously the threats that they must address.¹⁸ But in this Article I argue that some aspects of his principles remain dangerously inconsistent with parts of the *jus ad bellum* regime and its relationship with other legal regimes, and excessively privilege the interests of powerful target states at the expense of the rights of weak territorial states. At the same time, in the spirit of his project, I propose ways of refining the principles. Thus, the Article explores how the doctrine, as articulated in the Bethlehem

and International Law: Address to the Annual Meeting of the American Society of International Law (Mar. 25, 2010), <https://2009-2017.state.gov/s//releases/remarks/139119.htm> [<https://perma.cc/E7PN-SNV9>] (archived Feb. 6, 2019).

14. See generally, e.g., Steven R. Ratner, *Self-Defense Against Terrorists: The Meaning of Armed Attack*, in THE LEIDEN POLICY RECOMMENDATIONS ON COUNTER-TERRORISM AND INTERNATIONAL LAW (N. Schrijver & L. Van Den Herk eds., 2012); Dawood I. Ahmed, *Defending Weak States Against the “Unwilling or Unable” Doctrine of Self-Defense*, 9 J. INT’L L. & INT’L REL. 1 (2013); Deeks *Unwilling*, *supra* note 7; Christian J. Tams, *The Use of Force Against Terrorists*, 20 EUR. J. INT’L L. 359 (2009); Iona Ebben, *The Use of Force Against a Non-State Actor in the Territory of Another State: Applying the Self-Defense Framework to Al-Qaeda* (2011) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1960508 [<https://perma.cc/JZY9-UWEJ>] (archived Feb. 15, 2019).

15. See *infra* Part II.C for details of these developments.

16. See generally Daniel Bethlehem, *Self-Defense Against Imminent or Actual Armed Attack by Nonstate Actors*, 106 AM. J. INT’L L. 770 (2012) [hereinafter Bethlehem *Self-Defense Against NSAs*]; Daniel Bethlehem, *Principles of Self-Defense—A Brief Response*, 107 AM. J. INT’L L. 579 (2013) [hereinafter Bethlehem *A Brief Response*].

17. See *infra* Part III.A.

18. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 773.

Principles, might be refined in ways that would pragmatically address the threats posed by NSAs operating from within unwilling states, while at the same time bringing the principles of the doctrine into greater compliance with the *jus ad bellum* regime, and preserving the integrity of the regime's relationships with IHL and the law of state responsibility.

The unwilling or unable doctrine, and particularly one specific policy maker's formulation of it, may at first glance seem a rather esoteric topic of limited significance. But this is misleading. The legitimacy and legality of this doctrine, which has come to be embodied and articulated by the Bethlehem Principles, is of great importance to the integrity of the *jus ad bellum* regime, and by extension, to the future of international peace and security. For while the doctrine cannot yet be understood as being part of customary international law,¹⁹ and aspects of it are inconsistent with long-established principles of *jus ad bellum*,²⁰ if international law does evolve to embrace the doctrine, the threshold for using force in self-defense would be significantly lowered for all purposes and in all circumstances, and the *jus ad bellum* regime would be destabilized and weakened in a number of other ways, which will be explained below. Thus, while the doctrine has been developed to more effectively deal with the narrow and specific threat posed by transnational terrorist organizations, it threatens to weaken the broader *jus ad bellum* regime in ways that would increase the much more serious risk of interstate armed conflict. And not only is that the primary risk that the *jus ad bellum* regime was designed to address, but the risk of armed conflict between major powers is already far graver today than it was a scant decade ago. This is not to trivialize the threat posed by transnational terrorism, but it poses a far lesser risk to states than the prospect of interstate war. Therefore, this doctrine designed to address terrorism cannot be allowed to undermine the legal regime developed to constrain interstate use of force. As between the two, the integrity of the *jus ad bellum* regime is far more important. And thus understanding how the doctrine threatens to weaken the *jus ad bellum* regime, and considering how to bring the doctrine back into compliance with that regime, is important, particularly since refining the doctrine can be done without significantly undermining its effectiveness.

Part II of the Article begins by identifying the positions it takes on several contentious issues in *jus ad bellum*, in order to explain some of the assumptions and premises of its argument. It also provides some background on the history of the unwilling or unable doctrine, and then

19. See *infra* Part III.A.

20. See *infra* Part III.B-D.

reviews the Bethlehem Principles themselves. In Part III, the Article moves to critically examine those aspects of the Bethlehem Principles that it challenges as being problematic. Specifically, it critiques how the principles: (i) have been accepted as custom; (ii) distort and weaken the concept of imminence; (iii) develop and rely upon a novel definition of the concept of “armed attack”; (iv) conflate principles of *jus ad bellum* with those of IHL, and import considerations of the law of state responsibility into the doctrine of self-defense; and (v) purport to confer authority on the target state to unilaterally make a number of important determinations—including the determination that the territorial state is “unwilling” to either take action or consent to such action—without unpacking and examining the differing nature of each of those determinations, or providing differentiated standards to govern how and by whom they are to be made.

In Part IV, the Article turns to the problem of trying to refine the doctrine so as to better comply with the *jus ad bellum* regime, while still retaining a framework that target states will not summarily dismiss as being impractical or insufficient. This begins with a rejection or reformulation of those aspects of the Bethlehem Principles that are most egregiously inconsistent with established international law—and which are not, in any event, central to the operation of the doctrine—in order to restore the integrity of the *jus ad bellum* regime. Then, drawing upon theories about self-judgment and evidentiary standards in international law, the Article unpacks and identifies the nature of each of the decisions that the target state is purportedly permitted to make unilaterally under the doctrine, and explores the kinds of standards that should govern how these decisions are made and then later assessed. All of this aims to bring the doctrine into greater compliance with the *jus ad bellum* regime, and redress the serious asymmetry in how the doctrine balances the rights and interests of target states and territorial states.

II. THE LAWS OF WAR AND ORIGINS OF THE DOCTRINE

The unwilling or unable doctrine is claimed to be part of the doctrine of self-defense in the *jus ad bellum* regime, though the use of force against NSAs in nonconsenting states also implicates IHL. While most readers will be familiar with the basic principles of the *jus ad bellum* regime, there are important aspects of the doctrine of self-defense that remain contested and controversial.²¹ In this Part,

21. There is a massive literature on the topic, but some of the leading treatises include YORAM DINSTEIN, *WAR AGGRESSION AND SELF-DEFENCE* (6th ed. 2017) [hereinafter *DINSTEIN AGGRESSION*]; THOMAS M. FRANCK, *RECOURSE TO FORCE: STATE*

therefore, I identify what position I am taking on the contested issues for the purposes of my arguments on the unwilling or unable doctrine. In addition, I explain the salient aspects of the traditional understanding of the relationship between *jus ad bellum* and IHL,²² before going on to review the development of the unwilling or unable doctrine and reprise the Bethlehem Principles. Those readers well acquainted with these areas of the law may want to merely skim these subparts.

A. *Jus ad Bellum Regime Assumptions*

The modern *jus ad bellum* regime was established with the UN system after World War II.²³ The UN Charter (the Charter) prohibits the threat or use of force against the political independence or territorial integrity of other states, or in any other way inconsistent with the principles enshrined in the Charter.²⁴ The Charter provides for two general exceptions to this prohibition, being the Article 51 right of individual and collective self-defense in response to an armed attack, and the use of force by members of the United Nations as authorized by the UN Security Council under Article 42 in order to restore or maintain international peace and security.²⁵ The unwilling or unable doctrine is argued to be part of the doctrine of self-defense reflected in the Article 51 right.²⁶ To emphasize the obvious, the modern *jus ad bellum* regime, developed in the wake of two catastrophic world wars, reflected an effort to create a stronger system of constraints on the use of force in order to reduce the incidence of armed conflict among states.

ACTION AGAINST THREATS AND ARMED ATTACKS (2002); JUDITH GARDAM, NECESSITY, PROPORTIONALITY, AND THE USE OF FORCE BY STATES (2004); CHRISTINE GRAY, INTERNATIONAL LAW AND THE USE OF FORCE (4th ed. 2018); LUBELL, *supra* note 3; LINDSEY MOIR, REAPPRAISING THE RESORT TO FORCE: INTERNATIONAL LAW, *JUS AD BELLUM* AND THE WAR ON TERROR (2010); RUYS, *supra* note 3.

22. The brief overview provides what I take to be a relatively mainstream perspective on the relevant principles, and I flag those aspects that are contested without delving too deeply into the debates.

23. For the history of these developments, see generally STEPHEN NEFF, WAR AND THE LAW OF NATIONS: A GENERAL HISTORY (2005). For an argument that the Kellogg-Briand Pact of 1928 (formally The General Treaty for Renunciation of War as an Instrument of National Policy, 94 League of Nations Treaty Series No. 2137 (1928)), is the real inflection point for the development of the modern *jus ad bellum* regime, see generally OONA A. HATHEWAY & SCOTT J. SHAPIRO, THE INTERNATIONALISTS: HOW A RADICAL PLAN TO OUTLAW WAR REMADE THE WORLD (2017).

24. U.N. Charter art. 2(4).

25. *Id.* arts. 51, 39–43, respectively. For analysis of the regime, see *supra* sources cited note 21.

26. See *infra* Part II.C & D.

It is at the very center of the UN system, which is primarily for the purpose of maintaining international peace and security.²⁷ This broad purpose should inform how we think about proposals to modify the regime in order to deal with much narrower threats.

While the *jus ad bellum* regime in general is grounded in treaty, the International Court of Justice (ICJ) has held that the principles of the regime also exist independently in almost identical form in customary international law.²⁸ While there is some debate as to whether the scope of the right of self-defense is broader in custom than it is under the Charter,²⁹ here I will adhere to the well-established view that the scope remains the same for both.³⁰ The language of Article 51 of the Charter limits the right of self-defense to the use of force in response to an "armed attack."³¹ This has given rise to two distinct issues, both of which continue to be the subject of considerable debate, and even some movement in terms of state practice, and both of which are relevant to our inquiry: first, can states act in anticipation of an armed attack, and if so, based on what standard; and second, what use of force rises to the level of constituting an armed attack?

With respect to the first issue, some of the most careful and persuasive scholarly studies of the issue suggest that the current state of the law is that self-defense is only permitted in response to an actual armed attack, and thus anticipatory self-defense is not permitted.³² There is also, however, a strong body of scholarship, and some movement developing in state practice, that reflects the contrary view.³³ Since the unwilling or unable doctrine both assumes the

27. U.N. Charter, *supra* note 24, Preamble, art. 1.

28. *Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.)*, Judgment, 1986 I.C.J. Rep. 14, ¶¶ 172–82, 187–201 (June 27) [hereinafter *Nicaragua*].

29. For discussion of and refutation of this argument, see IAN BROWNLIE, *THE USE OF FORCE BY STATES*, 269–78 (1963); Ian Brownlie, *Legal Regulation of the Use of Force*, 8 INT'L & COMP. L.Q., 717–20 (1959). For more recent analysis, see RUYSS, *supra* note 3, at 53–125.

30. *Nicaragua*, 1986 I.C.J. Rep. at ¶¶ 187–202; RUYSS, *supra* note 3, at 19 ("In sum, the idea that the adoption of the UN Charter gave rise to parallel existence of two substantially divergent track for the international law on the use of force . . . is artificial and theoretically unsustainable.")

31. U.N. Charter, *supra* note 24, art. 51.

32. DINSTEN AGGRESSION, *supra* note 21, at 221–25; GRAY, *supra* note 21, at 169–74; RUYSS, *supra* note 3, at 255–304.

33. See, e.g., John A. Cohan, *The Bush Doctrine and the Emerging Norm of Anticipatory Self-Defense in Customary International Law*, 15 PACE INT'L L. REV. 283, 315 (2003); The United Nations Secretary-General's High-Level Panel on Threats, Challenges and Security, *A More Secure World: Our Shared Responsibility*, UN Doc. A/59/565 (2004), ¶¶ 188–93 [hereinafter UN High-Level Panel on Threats]. See generally W. Michael Reisman & Andrea Armstrong, *Centennial Essay: The Past and Future of the Claim of Preemptive Self-Defense*, 100 AM. J. INT'L L. 525 (2006). On state practice, see GRAY, *supra* note 21, at 169–75; RUYSS, *supra* note 3, at 29–52. On the difference between

legitimacy of anticipatory self-defense, and is most frequently invoked in circumstances of anticipatory strikes, this Article assumes that a narrow conception of anticipatory self-defense is valid. According to this conception of anticipatory self-defense, states may use force in anticipation of an “imminent armed attack.”³⁴ What constitutes an “imminent armed attack” is itself an issue of considerable debate, but let us begin with the understanding that at its narrowest, it is an attack that is already irrevocably in motion or on the verge of being launched.³⁵ As will be discussed below, this does not extend to more temporally distant threats—self-defense cannot be claimed to justify the use of force to prevent the development of potential future threats, or to punish past attacks.³⁶ This is consistent with the understanding that the exercise of self-defense is governed by the principle of necessity, meaning that the use of force is a last resort and the only means of preventing or terminating an armed attack. The principle of necessity implies some degree of immediacy, for if the attack is not currently in motion or immediately pending, there is likely time to explore other alternatives to the use of force to prevent it.³⁷ But there is debate over the meaning and scope of the concept of imminence, to which I will return below.

With respect to the second issue, namely what use of force rises to the level of an “armed attack” for purposes of Article 51, the threshold is quite high. The ICJ has held that the use of force constituting an “armed attack” sufficient to trigger the right of self-defense is substantially greater than the use of force that is itself subject to the general prohibition.³⁸ This of course means that not every unlawful use of force rises to the level of armed attack triggering the right of self-defense—which implies that the law contemplates circumstances in which states are limited to non-forceful ways of responding to low

anticipatory and preemptive/preventative self-defense, see *infra* text accompanying note 142.

34. FRANCK, *supra* note 21, at 97; RUYS, *supra* note 3, at 250–54.

35. DINSTEIN AGGRESSION, *supra* note 21, at 232–33, 251–52. Dinstein also argues that an attack irrevocably in motion may constitute an actual armed attack, even if no shot has yet been fired, and also proposes an even more “immediate” category of “interceptive self-defense.” *Id.* at 228–32.

36. See *infra* text accompanying note 120 (discussing preventative self-defense).

37. GRAY, *supra* note 21, at 157–58; RUYS, *supra* note 3, at 250–54.

38. The mining of a naval vessel, and the firing of a sea-sparrow missile at an ocean-going oil-tanker, for instance, were held not to constitute armed attacks for the purposes of triggering the right of self-defense. See *Case Concerning Oil Platforms (Iran v. U.S.)*, 2003 I.C.J. Rep. 160, ¶¶ 51, 64 (Nov. 6). See also *Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.)*, Judgment, 1986 I.C.J. Rep. 14, ¶¶ 191, 210–11, 230–32 (June 27).

levels of unlawful uses of force. The United States does not share this view, but it is among a small minority that takes this position.³⁹ Here I am assuming the established view that the threshold for armed attack is high, but as we will explore below, just how high is a key issue in the debate over the doctrine.⁴⁰

There are two other issues of some controversy regarding the availability and operation of the doctrine of self-defense in response to attacks launched by NSAs. The first relates to whether an attack launched by an NSA qualifies as an “armed attack” triggering the right of self-defense at all. There have been arguments that the *jus ad bellum* regime only governs states, and thus does not contemplate an attack by a non-state entity.⁴¹ This issue has largely been settled after 9/11, when the UN Security Council quite clearly recognized that the 9/11 attacks constituted an “armed attack” justifying the right of self-defense,⁴² and it is assumed here that this is the current state of the law. The second issue, however, relates to the lawful target of the use of force undertaken in self-defense. The traditional view was that Article 51 authorized the use of force against states to which the armed attacks of NSAs could be attributed, but not against NSAs *as such*—meaning not against the NSAs independent of the state in which they are operating. In other words, it is the state to which the operations of the NSA can be imputed or attributed that is the sole legal object of the target state’s use of force for purposes of the *jus ad bellum* regime. Conversely, it is not lawful to use force against an NSA within the territory of a state which is in no way responsible for the actions of the NSA.⁴³ While this has been the traditional view, it is at the very heart of the debate over the unwilling or unable doctrine, and is deeply contested, as we will come to below. In short, some argue that the use of force against an NSA within the territory of another state does not constitute a use of force against the state, and thus does not implicate

39. RUYS, *supra* note 3, at 260; see also Michael Schmitt & Ryan Goodman, *Best Advice for Policymakers on “Bloody Nose” Strike Against North Korea: It’s Illegal*, JUST SECURITY (Jan. 23, 2018), <https://www.justsecurity.org/51320/advice-policymakers-bloody-nose-strike-north-korea-illegal/> [<https://perma.cc/E6Y7-WFPG>] (archived Feb. 6, 2019).

40. See *infra* Part II.C.

41. For analysis of both the issue and the debate, see DINSTEIN AGGRESSION, *supra* note 21, at 240–49; LUBELL, *supra* note 3, at 29–42; RUYS, *supra* note 3, 419–33.

42. S.C. Res. 1368 (Sept. 12, 2001); S.C. Res. 1373 (Sept. 28, 2001).

43. *Nicaragua*, 1986 I.C.J. Rep. at ¶¶ 227–30. The ICJ has also suggested that the U.N. Security Council resolutions referred to, see *supra* note 42, were not inconsistent with nor changed the proposition that self-defense is only available when the acts of an NSA can be attributed to a territorial state. See *Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory*, Advisory Opinion, 2004 I.C.J. Rep. 136, ¶ 139 (July 9).

jus ad bellum at all.⁴⁴ This view is not widely accepted. But it also ignores the fact that even if not a violation of the prohibition against the use of force, the strike against the NSA is at minimum an unlawful intervention and violation of the sovereignty of the territorial state, and that self-defense is the only circumstance precluding wrongfulness under the law of state responsibility that would justify these separate violations.⁴⁵

B. *Jus ad Bellum and IHL*

While the *jus ad bellum* regime dictates the conditions for the use of force or commencement of armed conflict between states, the *jus in bello* regime, or IHL, governs the specific conduct of armed forces within armed conflict.⁴⁶ IHL is based on two core ideas that coexist in constant tension—namely, on the one hand that there must be legal constraints placed on how armed forces fight, particularly on who and what may be targeted, and on the other hand, the notion that there must be legal authority for the use of deadly force by the legitimate armed forces of a state in the pursuit of valid military objectives in war.⁴⁷ We need not delve into even a cursory review of the main principles, but it is important for the later analysis of the Bethlehem Principles to note that IHL has very specific rules on targeting, which are in turn based on the principle of distinction—one of the four core

44. See *infra* Part III.D. The more extreme contrary views suggest that neither attribution of the acts of the NSA to the territorial state, nor the territorial state's consent are required for a use of force against the NSA within the territorial state. See, e.g., Jordan J. Paust, *Self-Defense Targetings of Non-State Actors and Permissibility of U.S. Use of Drones in Pakistan*, 19 J. TRANSNAT'L L. & POL'Y 237, 249–50 (2010).

45. See generally Federica I. Paddeu, *Self-Defence as a Circumstance Precluding Wrongfulness*, 30 LEIDEN J. INT'L L. 93 (2017).

46. Jasmine Moussa, *Can Jus ad Bellum Override Jus in Bello? Reaffirming the Separation of the Two Bodies of Law*, 90 INT'L REV. RED CROSS 872, 965–66 (2008).

47. On the IHL regime generally, see, for example, ANTHONY CULLEN, *THE CONCEPT OF NON-INTERNATIONAL ARMED CONFLICT IN INTERNATIONAL HUMANITARIAN LAW* (2010); CUSTOMARY INTERNATIONAL HUMANITARIAN LAW, VOLS. I & II (Jean-Marie Henckaerts & Louise Doswald-Beck eds., 2005) [hereinafter HENCKAERTS, CUSTOMARY IHL]; YORAM DINSTEIN, *THE CONDUCT OF HOSTILITIES UNDER THE LAW OF INTERNATIONAL ARMED CONFLICT* (2004) [hereinafter DINSTEIN CONDUCT]; *THE HANDBOOK OF HUMANITARIAN INTERNATIONAL LAW* (Dieter Fleck ed., 3d ed. 2013); INTERNATIONAL INSTITUTE OF HUMANITARIAN LAW, *THE MANUAL ON THE LAW OF NON-INTERNATIONAL ARMED CONFLICT, WITH COMMENTARY* (2006) [hereinafter SANREMO NIAC MANUAL]; LINDSAY MOIR, *THE LAW OF INTERNAL ARMED CONFLICT* (2002); GARY D. SOLIS, *THE LAW OF ARMED CONFLICT: INTERNATIONAL HUMANITARIAN LAW IN WAR* (2010).

principles of the IHL regime.⁴⁸ Also, it is important to note that the IHL regime only operates in the context of armed conflict, and the IHL regime itself provides the criteria for determining the existence of an armed conflict, whether international or non-international in character.⁴⁹

What is key for our purposes is the relationship between the *jus ad bellum* and IHL regimes. While related, governing different aspects of war, they are in many important respects independent and distinct. Specifically, this separation is crucial to the principle of equality in IHL, which dictates that the rights and obligations under IHL apply equally to the armed forces of all belligerents regardless of which side ultimately has legal authority to use force under the rules of *jus ad bellum*. This is considered essential to maximizing adherence to the rules of IHL, and thus achieving the ultimate objective of reducing the amount of human suffering in armed conflict.⁵⁰ While the two regimes are largely independent in this sense, there continues to be an important connection between them.⁵¹ In particular, when a state uses armed force against or within the territory of another state, in the sense captured by Article 2(4) of the UN Charter, such that the rules of *jus ad bellum* apply, then that use of force (or the armed attack it is

48. See Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of International Armed Conflict, arts. 44(3), 48 and 51(3), 8 June 1977, 1125 U.N.T.S. 3 [hereinafter Additional Protocol I]; Protocol Additional to the Geneva Conventions of 12 August 1949, and Relating to the Protection of Victims of Non-International Armed Conflicts (Protocol II), art. 13(1), 8 June 1977, 1125 U.N.T.S. 609, [hereinafter Additional Protocol II]; DINSTEIN CONDUCT, *supra* note 47, at 27–28, 82–87; HENCKAERTS, CUSTOMARY IHL, *supra* note 47, at 3–36; SOLIS, *supra* note 47, at 251–53.

49. Geneva Convention Relative to the Treatment of Prisoners of War arts. 2–3, Aug. 12, 1949 U.S.T. 3316, 75 U.N.T.S. 135; Additional Protocol I, *supra* note 48, arts. 1(3)–(4); Additional Protocol II, *supra* note 48, art. 1(1); see also Prosecutor v. Tadić, Case No. IT-94-1-A, Judgment, ¶ 84 (Int'l Crim. Trib. for the Former Yugoslavia July 15, 1999). Considerable controversy surrounds these criteria, particularly in relation to US claims about a transnational non-international armed conflict against certain NSAs, but it does not bear directly on the unwilling or unable doctrine. See, e.g., YORAM DINSTEIN, NON-INTERNATIONAL ARMED CONFLICTS IN INTERNATIONAL LAW 1–19 (2014) [hereinafter DINSTEIN NIAC]. See generally Geoffrey S. Corn & Eric T. Jensen, *Transnational Armed Conflict: A 'Principled' Approach to the Regulation of Counter-Terror Combat Operations*, 42 ISR. L. REV. 46 (2012).

50. DINSTEIN AGGRESSION, *supra* note 21, at 156–68; NEFF, *supra* note 23, at 340–46, 366–69. See generally KEHCHIRO OKIMOTO, THE DISTINCTION AND RELATIONSHIP BETWEEN *JUS AD BELLUM* AND *JUS IN BELLO* (2011); Eyal Benvenisti, *Rethinking the Divide between Jus ad Bellum and Jus in Bello in Warfare Against Non-State Actors*, 34 YALE J. INT'L L. 541 (2009); Alexander Orakhelashvili, *Overlap and Convergence: Jus ad Bellum and Jus in Bello*, 12 J. CONFLICT & SECURITY L. 157 (2007); Robert D. Sloane, *The Cost of Conflation: Preserving the Dualism of Jus ad Bellum and Jus in Bello in the Contemporary Law of War*, 34 YALE J. INT'L L. 47 (2009).

51. But see Sloane, *supra* note 50, at 67–69 (suggesting that the independence of the regimes in the Charter era is actually exaggerated).

responding to) constitutes the initiation of an international armed conflict to which the rules of IHL will apply.⁵² The use of force in *jus ad bellum* terms triggers the operation of the *lex specialis* of IHL, which then operates to govern the conduct of armed forces in the ensuing hostilities—and it is thus IHL and not *jus ad bellum* that provides the rules dictating the conditions under which persons and objects can be lawfully targeted.⁵³ Once the armed conflict is initiated, the *jus ad bellum* principles of necessity and proportionality do continue to operate, in that the defending state is limited to using force in self-defense that continues to be necessary to prevent continued attacks, and that is proportionate to the risk of harm posed by the armed attacks to which the state is responding.⁵⁴ But when a target state uses force against an NSA within a territorial state, this does constitute the use of force by one state against another.⁵⁵ What is more, the operation of IHL is also triggered, and thus both *jus ad bellum* and IHL will have to be considered for the purposes of determining the legality of quite different aspects of this action.⁵⁶ As I will return to below, it is important not to conflate or confuse the distinct aspects that are governed by each regime, or to apply principles from one regime to aspects of armed conflict that are governed by the other regime.

C. Background and Origins of the Unwilling or Unable Doctrine

As explained at the outset, the Bethlehem Principles have come to shape the debate on the unwilling or unable doctrine. Moreover, they appear to have heavily influenced government positions, as evidenced by the explicit reference to and endorsement of them in a number of

52. DINSTEN AGGRESSION, *supra* note 21, at 156–62 (also tracing how the evolution of *jus ad bellum* led to arguments in favor of re-integration of the regimes); DINSTEN CONDUCT, *supra* note 47, at 14–16; MELZER, *supra* note 3, at 247–51, 394–95; Kretzmer, *supra* note 3, at 188.

53. See *infra* Part II.D.

54. Terry D. Gill, *When Does Self-Defence End?*, in THE OXFORD HANDBOOK OF THE USE OF FORCE IN INTERNATIONAL LAW 737 (Marc Weller ed., 2014); GRAY, *supra* note 21, at 157–65.

55. See sources cited *supra* note 43. Some, of course, argue that a use of force against NSAs within a territorial state does not constitute a use of force against the state, or trigger an international armed conflict. See, e.g., Paust, *supra* note 44, at 249–58.

56. Advisory Opinion on the Legality of the Threat or Use of Nuclear Weapons, 1996 I.C.J. 226, ¶¶ 75–87, 95 [hereinafter *Threat of Nuclear Weapons*]; GARDAM, *supra* note 21, at xiii.

important government statements.⁵⁷ They tend to be seen as representing the current state of the law by a number of significant members of the international community.⁵⁸ The focus of my analysis, therefore, will be on critiquing and trying to refine the doctrine as articulated in the Bethlehem Principles. But before turning to the detail of those principles, it may be helpful to explore some history of the unwilling or unable doctrine in order to provide some context. For while the doctrine has become far more prominent and controversial in the wake of the 9/11 attacks, it does have an older pedigree.

Ashley Deeks explored the theoretical and historical foundations for the doctrine in an important article that pre-dated the Bethlehem Principles.⁵⁹ There is a very long history to the more basic concept that a state may lawfully take action against enemy forces operating from within the territory of another non-belligerent state, if that state is unwilling or unable to prevent those enemy forces from using its territory in this way. The concept can be traced back to Emer de Vattel, and was originally part of the law of neutrality, as later codified in the 1907 Hague Convention.⁶⁰ Under the law of neutrality, neutral states had an obligation to not permit belligerent forces to operate from within their territory.⁶¹ In the event that neutral states failed to prevent such operations, other belligerents were permitted to use force against those enemy forces within the neutral state's territory.⁶² In the nineteenth and early twentieth centuries, the principle was primarily applied to states, in that the "enemy forces" that neutral states were obliged to prevent from operating within their territory were typically the armed forces of another belligerent state.⁶³

While the concept has its theoretical origins in neutrality law and thus primarily implicates state action, the application of the principle to non-state actors is not exclusively a late twentieth century

57. See, e.g., Right Honourable Jeremy Wright, QC MP and Attorney General of the United Kingdom, Address to the International Institute for Strategic Studies, The Modern Law of Self-Defence (Jan. 11, 2017), <https://www.gov.uk/government/speeches/attorney-generals-speech-at-the-international-institute-for-strategic-studies> [<https://perma.cc/837J-44HC>] (archived Feb. 6, 2019) [hereinafter Wright *Self-defense*]; George Brandis, QC Attorney General of Australia, The Right of Self-Defense Against Imminent Armed Attack in International Law (Apr. 11, 2017), <https://www.ejiltalk.org/the-right-of-self-defence-against-imminent-armed-attack-in-international-law/#more-15255> [<https://perma.cc/X8ER-L3VC>] (archived Feb. 6, 2019).

58. These include, most prominently, the United States, the United Kingdom, Australia, and Israel.

59. See generally Deeks *Unwilling*, *supra* note 7.

60. *Id.* at 499.

61. *Id.* at 497–503.

62. *Id.* at 499–501.

63. Paul Seger, *The Law of Neutrality*, in THE OXFORD HANDBOOK OF INTERNATIONAL LAW IN ARMED CONFLICT 248–272 (Andrew Clapham & Paola Gaeta eds., 2014).

phenomenon. As many have noted, the famous *Caroline* incident of 1837, which is typically cited as defining the test for anticipatory self-defense, was an early example of the application of the unwilling or unable doctrine to justify the use of force against non-state actors within the territory of a nonconsenting state.⁶⁴ It will be recalled that the *Caroline* incident involved the use of force by Britain against non-state actors, Canadian rebels and their American supporters, operating within the territory of the United States. The insurgents had seized an island within British-Canadian waters, from which they were shelling Canada and threatening to launch an invasion of Canadian territory. The British attacked, set fire to, and destroyed a seventy-one-foot vessel that was being used by the Canadian insurgents to transport weapons and personnel from within the United States to the island.⁶⁵

This resulted in a diplomatic dispute between Britain and the United States that spanned a number of years. It is this dispute that produced the now famous statement—that the use of force had to be justified by a “necessity of self-defense, instant, overwhelming, leaving no choice of means, and no moment for deliberation”—that has come to inform the test for self-defense. But the statement was made in an exchange of diplomatic notes between U.S. Secretary of State Daniel Webster and Lord Ashburton, the representative of the United Kingdom in Washington for purposes of treaty negotiation, some five years after the incident itself.⁶⁶ A less well-known part of those notes was the British assertion that the use of force was a justifiable exercise of self-defense because the United States had been unwilling or unable to prevent the Canadian rebels from conducting attacks against British Canada.⁶⁷ Indeed, Ashburton began his argument by suggesting that Webster’s “ingenious” formulation of the principle was actually beside the point and inapplicable to the case.⁶⁸ Ashburton suggested that the

64. Deeks *Unwilling*, *supra* note 7, at 501–03; *see also* Abraham D. Sofaer, *On the Necessity of Pre-Emption*, 14 EUR. J. INT’L L. 209, 216–17 (2009).

65. Dinstein points out that the *Caroline* Incident was not about self-defense in the modern sense, since there was no prohibition on the use of force at the time—it was really about constraints on measures short of war so as to avoid escalation. DINSTEIN *AGGRESSION*, *supra* note 21, at 225; GRAY, *supra* note 21, at 158. For a recent but sure to be the seminal work on the factual and intellectual history of the *Caroline* incident, *see generally* CRAIG FORCESE, *DESTROYING THE CAROLINE* (2018).

66. Hunter Miller, *British-American Diplomacy The Caroline Case*, in 4 TREATIES AND OTHER INTERNATIONAL ACTS OF THE UNITED STATES OF AMERICA 80–121 (1934). For extensive analysis of the exchange and their legal significance, *see* FORCESE, *supra* note 65, Parts III–V.

67. Miller, *supra* note 66.

68. *Id.*

case should be settled by consideration of the following question: “if cannon are moving and setting up in a battery which can reach you and are actually destroying life and property by their fire, if you have remonstrated for some time without effect and see no prospect of relief, when begins your right to defend yourself . . . ?”⁶⁹ It is worth noting that while Lord Ashburton was clearly suggesting that a state has a right to use force in self-defense against individuals who are launching attacks from within the territory of another sovereign, he also was acknowledging that the right only arises after there has been remonstrance “for some time without effect and no prospect of relief.”⁷⁰

While the doctrine thus has origins that long predate the 9/11 attacks, it gained much greater significance and prominence in the militarized response to those attacks, and to transnational terrorism more generally in the years that followed. The United States ramped up its use of force against transnational terrorist organizations within the territory of nonconsenting states, particularly in the form of targeted killing with drones, and it explicitly articulated the unwilling or unable doctrine as a justification for doing so.⁷¹ This made the doctrine far more prominent, and extended and applied the doctrine to specifically justify action against NSAs.⁷² As the legality of the targeted killing program in particular became hotly debated, the validity of the unwilling or unable doctrine also came under increasing scrutiny.⁷³ For while the doctrine may have been recognized within the context of the law of neutrality, and there had been some instances of states invoking it to justify strikes against NSAs within nonconsenting states in the past, it was (and arguably continues to be) hotly disputed whether the doctrine as applied to NSAs constitutes a principle of customary international law that forms part of the broader doctrine of self-defense.⁷⁴

The doctrine did receive further support in important fora even prior to Bethlehem’s article. For instance, in 2006 Chatham House published a set of principles on the use of force, based on a survey of noted academics and policy makers in the field, that asserted that states may use force in self-defense against non-state actors operating within the territory of nonconsenting states, if those states were

69. *Id.*

70. *Id.*

71. *See, e.g.*, sources cited *supra* note 13.

72. *See id.*

73. *Legality of Targeted Killing Program under International Law*, LAWFARE, <https://www.lawfareblog.com/legality-targeted-killing-program-under-international-law> (last visited Feb. 27, 2019) [<https://perma.cc/ZQ4W-ZFEV>] (archived Feb. 27, 2019). *See generally* sources cited *supra* note 14.

74. *See, e.g.*, Deeks *Unwilling*, *supra* note 7; sources cited *supra* note 14. *See infra* Part III.A.

unwilling or unable to prevent the NSAs from mounting armed attacks against the defending states.⁷⁵ In articulating this position, it explicitly distinguished the ICJ decision in the *Case Concerning Armed Activities on the Territory of the Congo*, which had held that the unwillingness or inability of a state to deal with a threat did *not* justify a use of force against that state in self-defense, absent some evidence that the territorial state supported and was involved in the activity of the NSAs.⁷⁶ Similarly, in 2010 the principles emanating from a multi-year consultative process among academics and policy makers in Europe were published, and submitted to both the foreign minister and to the parliament of the Netherlands, under the title *The Leiden Policy Recommendations on Counter-Terrorism and International Law*.⁷⁷ This set of principles also acknowledged that states could, in exceptional circumstances, use force in self-defense against terrorist organizations within the territory of another state without its consent, in the event that the territorial state was “unwilling or unable to deal with the terrorist attacks.”⁷⁸ And there has been an ever-growing body of scholarship and articulation of policy positions that similarly adopt and embrace the unwilling or unable doctrine.⁷⁹

While the publication of such principles helped to bolster arguments that the doctrine was (or was emerging as) part of the customary international law doctrine of self-defense, they by no means settled the debate. What is more, even those who supported the validity of the doctrine acknowledged that it had never been developed in form and substance beyond a basic principle that operated at a very high level of generality and abstraction.⁸⁰ Neither the Chatham House nor the Leiden principles, for instance, provided any detailed criteria

75. See generally Elizabeth Wilmhurst, *Chatham House Principles of International Law on the Use of Force by States in Self-Defence*, 55 INT'L & COMP. L.Q. 963 (2006).

76. *Case Concerning Armed Activities on the Territory of the Congo (Dem. Rep. Congo v. Uganda)*, Judgment, 2005 I.C.J. Rep. 1, ¶¶ 146–47 (Dec. 19).

77. See generally Nico Schrijver & Larissa van den Herik, *Leiden Policy Recommendations on Counter-Terrorism and International Law*, 57 NETH. INT'L L. REV. 531 (2010).

78. *Id.* at 540, ¶ 32.

79. See, e.g., DINSTEIN AGGRESSION, *supra* note 21; LUBELL, *supra* note 3; Greg Travalio & John Altenburg, *Terrorism, State Responsibility, and the Use of Military Force*, 4 U. CHI. J. INT'L L. 97 (2003). [Even to some extent the special rapporteur on targeted killing.] See Alston Study, *supra* note 3; UN High-Level Panel on Threats, *supra* note 33, at ¶¶ 188–95.

80. Deeks *Unwilling*, *supra* note 7, at 503–06.

governing the application of the doctrine.⁸¹ On the basis of what criteria is it determined that the NSA poses an imminent threat of armed attack against the target state? Does the target state get to make that determination unilaterally, or does the territorial state have a say? On what criteria is it determined that the territorial state is unwilling or unable to deal with the threat posed by the NSA? Must the target state “remonstrate” with the territorial state, as suggested by Lord Ashburton, before a determination of unwillingness is made? And again, who gets to decide? What is more, there are separate questions about the relationship of the doctrine with IHL—would the use of force against the NSAs be part of, or give rise to, an international armed conflict, triggering the operation of IHL? As Deeks explained in her detailed exploration of the doctrine in 2012, up until that time the doctrine provided answers to none of these questions.⁸² And so into this void stepped Daniel Bethlehem, trying to put some more flesh on the bones.

D. *The Bethlehem Principles*

Daniel Bethlehem explicitly wrote his article with a view to addressing the perceived gap between the academic and the operational perspectives on the use of force against NSAs.⁸³ Having served as the principal legal advisor of the United Kingdom’s Foreign and Commonwealth Office, he came from the policy world.⁸⁴ In the article, he wrote that the substance of the principles he was outlining was drawn from discussions he was privy to within the foreign ministry, the defense ministry, and among military legal advisors from other countries.⁸⁵ He argued that while the scholarly discourse on the issue was important, the debate within and among governments about what the appropriate principles are and should be was more significant.⁸⁶ It is these debates, he argued, that were “material both to the crystallization and development of customary international law and to the interpretation of treaties.”⁸⁷ What is more, the academic

81. See generally, Wilmhurst, *supra* note 75; Schrijver & van den Herik, *supra* note 77.

82. Deeks *Unwilling*, *supra* note 7, at 503–06.

83. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 770. For initial commentary on the article, see also Ashley Deeks, *Readings: Daniel Bethlehem on Principles Governing Self-Defense Against Non-State Actors*, LAWFARE (Jan. 10, 2015), <https://www.lawfareblog.com/readings-daniel-bethlehem-principles-governing-self-defense-against-non-state-actors> [<https://perma.cc/XAW4-HRHW>] (archived June 14, 2019) [hereinafter Deeks *Readings*].

84. Deeks *Readings*, *supra* note 83.

85. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 770.

86. *Id.*

87. *Id.*

debate, and particularly the arguments of scholars advocating for a more restrictive approach to the law of self-defense, faced “significant challenges” in shaping operational thinking, in part because the doctrinal debate had “yet to produce a clear set of principles that effectively address the specific operational circumstances faced by states.”⁸⁸

Bethlehem also noted, however, that while the government-level debates were more important than the scholarly discourse, they were largely invisible—elements of them only became apparent through infrequent statements or remarks by government officials, or as evidenced by state practice.⁸⁹ Moreover, while important, the government-level debates were no more successful than the scholarly debates in developing clear, detailed principles to flesh out and govern the operation of the doctrine.⁹⁰ He thus saw a need to “formulate principles, capable of attracting a broad measure of agreement, that apply, or ought to apply, to the use of force in self-defense against an imminent or actual armed attack by nonstate actors.”⁹¹ His article proceeded to lay out sixteen such principles. Before getting to the principles themselves, it is important to emphasize that Bethlehem himself explicitly recognized that several of these principles did not then represent either customary international law or treaty law. Indeed, he acknowledged that several of them were likely to be controversial, and he stated as his objective that they be merely the basis for further debate, and potentially be the locus of future consensus.⁹² It is ironic that despite his cautionary qualifiers, several governments have latched onto his principles as reflecting established law. But it is also important to recognize that the Bethlehem Principles have come to represent and embody the substance of the unwilling or unable doctrine, for which reason any analysis of the doctrine necessarily requires a focus on the Bethlehem Principles.

Turning to the principles themselves, it is worth noting at the outset that some, such as the treatment of imminence and the nature of armed attack, could be construed as being separate from the unwilling or unable doctrine itself, which is explicitly addressed more narrowly in only the last few of the Bethlehem Principles. I take the view, however, that the Bethlehem Principles were intended to constitute a single coherent formulation of the doctrine, and so should

88. *Id.* at 773.

89. *Id.* at 770–71.

90. *Id.* at 770–74

91. *Id.* at 773.

92. *Id.* at 773–74; *see also* Bethlehem *A Brief Response*, *supra* note 16.

be understood as such.⁹³ Also worth noting is that some of the principles are not at all controversial, or at least may be taken as given for the purposes of discussing the specifics of the doctrine. Thus, Principles One through Three merely stipulate that states have a right of self-defense against imminent or actual armed attack by NSAs; that the use of force in self-defense should be a last resort; and that the use of force in self-defense must be limited to what is necessary to address the attack, and proportionate to the threat faced.⁹⁴ Principle Four, however, is more novel, providing in part that: “the term ‘armed attack’ includes both discrete attacks and a series of attacks that indicate a concerted pattern of continuing armed activity.”⁹⁵ This contemplates precisely the kinds of sporadic and relatively low-intensity attacks that are frequently launched by transnational terrorist organizations. Principle Five further supports the formulation in Principle Four, by providing that a series of attacks may be considered as constituting a concerted pattern of continuing armed activity where there is a “reasonable and objective basis” for concluding that those perpetrating the attacks (actual or imminent) are acting in concert.⁹⁶ In other words, so long as the groups are determined to be working together, the occasional attacks of al-Qaeda, along with the infrequent attacks of AQAP, and those of Boko Haram, may be considered together as a series of attacks constituting a concerted pattern of continuing armed activity, which in the aggregate comprise an armed attack justifying an exercise of the right of self-defense. Depending on how it is interpreted, however, this definition of “armed attack” is not consistent with the high threshold established by the International Court of Justice in such cases as *Military and Paramilitary Activities in and Against Nicaragua (Nicaragua v. United States)*⁹⁷ and *Case Concerning Oil Platforms (Iran v. United States)*,⁹⁸ about which I will have more to say below.⁹⁹

Principles Six and Seven continue this focus on the idea of the threat posed by individuals acting in concert but do so in a manner that raises issues about the relationship between *jus ad bellum* and IHL. Principle Six provides that those acting in concert include “those planning, threatening, and perpetrating armed attacks and those

93. My thanks to Ashley Deeks for a helpful discussion on the different possible interpretations of the Principles.

94. *Bethlehem Self-Defense Against NSAs*, *supra* note 16, at 775.

95. *Id.*

96. *Id.*

97. *Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.)*, Judgment, 1986 I.C.J. Rep. 14, ¶¶ 191–95 (June 27).

98. *Case Concerning Oil Platforms (Iran v. U.S.)*, 2003 I.C.J. Rep. 160, ¶¶ 51, 64 (Nov. 6); *see also* RUYSS, *supra* note 3, at 7–11.

99. *See* sources cited *infra* notes 183–185, 284–285 and accompanying text.

providing material support essential to those attacks, such that they can be said to be taking part in those attacks.”¹⁰⁰ Bethlehem explained in a footnote that this is analogous to, but quite distinct from, the concept of “taking direct part in hostilities” in IHL.¹⁰¹ Nevertheless, Principle Seven provides that states may use force in self-defense against those “actively planning, threatening, or perpetrating armed attacks,” as well as against “those in respect of whom there is a strong, reasonable, and objective basis for concluding that they are taking a direct part in those attacks through the provision of material support essential to the attacks.”¹⁰² As I will return to below, the combination of these two rules would seem to suggest that states may use force against individuals determined to be providing material support for armed attacks, independent and perhaps regardless of whether they could be considered targetable under IHL.¹⁰³

Principle Eight relates to the highly controversial issue of what constitutes an “imminent” armed attack. It provides five factors to assist in this determination, which ultimately must be assessed by reference to “all relevant circumstances.”¹⁰⁴ The factors are (i) the nature and immediacy of the threat; (ii) the probability of an attack; (iii) whether the anticipated attack is part of a concerted pattern of continuing armed activity; (iv) the likely scale of the attack and the likely harm it will cause; and (v) the likelihood that there will be other opportunities to undertake effective action in self-defense that will cause less harm (particularly in the form of collateral damage).¹⁰⁵ The principle concludes with the assertion that “the absence of specific evidence of where an attack will take place or of the precise nature of an attack does not preclude a conclusion that an armed attack is imminent . . . provided that there is a reasonable and objective basis for concluding that an armed attack is imminent.”¹⁰⁶ The concept of imminence has become highly controversial since 9/11, and as will be discussed below, Bethlehem’s test in Principle Eight does not begin to resolve the problems.¹⁰⁷

Principles Nine and Ten lay the foundation for the core of the doctrine. Principle Nine provides that states have an obligation to take all reasonable steps to ensure that their territory is not used by NSAs

100. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 775.

101. *Id.* at 775 n.c.

102. *Id.* at 775.

103. See sources cited *infra* notes 195–202—and accompanying text.

104. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 775.

105. *Id.*

106. *Id.* at 776.

107. See *infra* Part III.B.

for purposes of mounting armed attacks on other states. Principle Ten establishes a baseline presumption that states may not use force against NSAs operating in the territory of another state without the consent of that state, subject to certain exceptions and qualifications.¹⁰⁸ It is the exceptions to this presumption that constitute the core of the doctrine. Thus, Principle Eleven provides that there is no requirement for consent in the event that there is a “reasonable and objective basis” that the territorial state is colluding with the NSA (thus referred to as a colluding state), or where it is “otherwise unwilling to effectively restrain the armed activities” of the NSA (thus referred to as a harboring state), leaving the victim state with no other reasonably effective means of countering the imminent or actual armed attack.¹⁰⁹ So Principle Eleven covers the “unwilling” aspect of the test, which relates to both “colluding” and merely “harboring” states.¹¹⁰ It should be noted that the “reasonable and objective basis” test contemplated by Bethlehem is a higher standard than the typical “reasonableness” test—he provided in a footnote that it “requires that the conclusion is capable of being reliably supported with a high degree of confidence on the basis of credible and all reasonably available information.”¹¹¹

Principle Twelve provides that consent is not required in the event that there is a “reasonable and objective basis” for concluding that the territorial state is “unable to effectively restrain the armed activities of the nonstate actor,” again leaving the target state with “no other reasonably available effective means to address the imminent or actual armed attack.”¹¹² Principle Twelve goes on to provide that in these circumstances, the victim state is only exempted from the obligation to obtain consent if there is also a “strong, reasonable, and objective basis for concluding” that efforts to seek consent would undermine the effectiveness of the exercise of self-defense, or would increase the risk of armed attack. It also states that in the process of seeking consent, the target state must provide the territorial state with a reasonable plan of action, which if it refuses to take would then constitute “unwillingness” and bring us back to Principle Eleven.¹¹³ This last element, the refusal of consent, is important, because it is the criteria for determining when an “unable” state is really an “unwilling” state, and as I will discuss below, unwillingness should be the sole focus of the doctrine—but of course, as discussed above, the target state can

108. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 776.

109. *Id.*

110. *Id.* at 775.

111. *Id.* at 775 n.a.

112. *Id.* at 776.

113. *Id.*

also infer this unwillingness.¹¹⁴ The addition of the word “strong” to the “reasonable and objective basis” here (and in Principle Seven, for using force against those merely supporting an armed attack) is meant to raise the standard because it is “the basis for taking action against persons other than those planning, threatening or perpetrating an armed attack.”¹¹⁵

Principle Thirteen closes the loop by providing that the consent given by the territorial state may be “strategic or operational, generic or ad hoc, express or implied.”¹¹⁶ The governing consideration is whether it is reasonable to regard the “representation or conduct” as authoritative consent.¹¹⁷ While there is a presumption against the inference of consent from historical acquiescence, the presumption can be rebutted, and thus even past acquiescence may be the basis for implied consent, particularly where such acquiescence has operated in circumstances in which objection would have been reasonably expected.¹¹⁸ Little is said about who makes each of these determinations, but it is implicit that the target state has the discretion to make these decisions for itself. All of this of course lowers the bar considerably on the determination of what constitutes consent, and opens the door to controversy over whether consent has actually been given in any particular situation—controversy that has surrounded many US drone strikes in Pakistan and Yemen.¹¹⁹ Finally, Principles Fourteen, Fifteen, and Sixteen merely provide that the doctrine does not operate to the prejudice of the operation of the UN Charter, the UN Security Council, or of customary international law relating to the use of force or self-defense, or principles relating to the law of state responsibility.¹²⁰

114. On why the focus should be “unwillingness” and not “inability,” see *infra* text accompanying notes 224–226.

115. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 775 n.d.

116. *Id.* at 777.

117. Bethlehem *A Brief Response*, *supra* note 16, at 585.

118. *Id.*

119. See, e.g., Mark Mazzetti, *U.S. Intensifying a Secret Campaign of Yemen Airstrikes*, N.Y. TIMES (June 8, 2011), <https://www.nytimes.com/2011/06/09/world/middleeast/09intel.html> [<https://perma.cc/E56G-RPUW>] (archived Feb. 6, 2019); Sanger & Schmitt, *supra* note 12.

120. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 777.

III. CRITIQUE OF THE DOCTRINE

Before launching into a critique of the unwilling or unable doctrine, it should be reiterated that there is much to be applauded in Bethlehem's effort. He wrote that one of his primary objectives was to move the policy

away from the rhetoric of a global war on terrorism, with its lack of geographic and temporal and other limitations, hinged on status-based targeting and driven by operational decision-making, and back to a legal framework that turns on an appreciation of imminent threats, of sovereignty, of limitations rooted in necessity and proportionality.¹²¹

Such efforts to ensure that counterterrorism policy is properly informed by international law in general and the *jus ad bellum* regime in particular, and to thereby make the policy more likely compliant with the law, is certainly to be welcomed. The problem is that some aspects of the principles are not consistent with the established *jus ad bellum* regime, or with the relationship between the *jus ad bellum* regime and either IHL or the law of state responsibility. Moreover, they do not fully achieve the other stated objective, which is to elaborate a clear set of criteria for determining when any particular use of force would be lawful.¹²²

A. *The Myth of Custom*

The first point that should be addressed is that there tends to be a growing but arguably mistaken belief, reflected in both scholarship and in government statements, that the unwilling or unable doctrine is now established customary international law.¹²³ Or, to put it in the most charitable terms, there has been a failure to distinguish between the normative and the descriptive in arguments and statements regarding the doctrine. This criticism does not extend to Bethlehem, who was quite clear in both of his articles that he understood that many of the principles would be controversial, and that he was making normative arguments for their adoption rather than claiming they had already crystallized as custom.¹²⁴ But the same cannot be said for

121. Bethlehem *A Brief Response*, *supra* note 16, at 580.

122. Ashley Deeks makes a similar point. See *generally* Deeks *Readings*, *supra* note 83.

123. *But cf.* Olivier Corten, *The 'Unwilling or Unable' Test: Has it Been, and Could it Be, Accepted?*, 29 LEIDEN J. INT'L L. 779 (2016) (arguing that while certain states, such as the U.S. and the U.K., seem to believe that the unwilling or unable doctrine is now customary international law, this is not truly the case).

124. Bethlehem *A Brief Response*, *supra* note 16, at 580–81; Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 773.

many others writing in the field, and perhaps more importantly, government officials making formal statements on their governments' official policies.¹²⁵ For instance, Jeremy Wright, attorney general of the United Kingdom, in a speech in early 2017, suggested that the law had changed as a result of the 9/11 attacks, and went on to say that

a number of states have also confirmed their view that self-defence is available as a legal basis where the state from whose territory the actual or imminent attack emanates is unable or unwilling to prevent the attack or is not in an effective control of the relevant part of its territory.¹²⁶

While parts of his speech reflected an acknowledgment that not all states agreed with the British position, and that parts of his claims on imminence were normative, the implication seemed to be that the unwilling or unable doctrine was fast becoming custom if it had not already crystallized as such.¹²⁷ Most recently, representatives of the United Kingdom and Australia, speaking at the American Society of International Law Annual Meeting in Washington, DC, suggested that the Bethlehem Principles were custom.¹²⁸

Turning to the actual evidence, according to one review of state positions on the issue, which Jeremy Wright himself referenced, only ten countries have explicitly endorsed the unwilling or unable doctrine as a justification for their own use of force in the UN era.¹²⁹ Several of these examples, such as that of Canada, Australia, the Netherlands, and Germany, were primarily drawn from letters to the UN Security Council in accordance with Article 51 of the Charter to explain their support for operations against ISIS in Syria and Iraq as being part of collective self-defense actions.¹³⁰ Another twelve states were listed as

125. For a compilation of government positions, see Chachko & Deeks, *supra* note 6. For examples of full statements by government representatives, see, for example, Brandis, *supra* note 57; Wright *Self-defense*, *supra* note 57.

126. Wright *Self-defense*, *supra* note 57.

127. *Id.*

128. The ASIL Proceedings for 2018 are not yet available, but for an analysis of the panel, see Alex Moorehead, *Brazil's Spirited Defense of the Legal Prohibition on the Use of Force and Self Defense*, JUST SECURITY (Apr. 20, 2018), <https://www.justsecurity.org/55126/brazils-robust-defense-legal-prohibition-force/> [<https://perma.cc/8G7T-NHWJ>] (archived Feb. 6, 2019); asil1906, *Use of Force against Non-State Actors*, YOUTUBE (Apr. 5, 2018), <https://www.youtube.com/watch?v=f166g4JgOpA> [<https://perma.cc/2W5M-9KXY>] (archived Feb. 15, 2019).

129. Deeks *Unwilling* *supra* note 7, at 549–50; see also Chachko & Deeks, *supra* note 6.

130. Chachko & Deeks, *supra* note 6. For further discussion of state practice in relation to Syria, see Jutta Brunnée & Stephen J. Toope, *Self-Defence against Non-State Actors: Are Powerful States Willing but Unable to Change International Law?*, 67 INT'L

either having provided implicit endorsement (three countries) or being ambiguous cases (nine states), while four were found to have explicitly objected.¹³¹ Ashley Deeks' earlier and more exhaustive study includes an appendix of all cases of the use of force against either an NSA or the armed forces of a third country, within the territory of a nonconsenting state—so a sample that goes beyond only uses of force against NSAs—between the years 1817 and 2011 (so prior to the more recent rash of cases involving air strikes against ISIS in Syria and Iraq).¹³² She found only thirty-nine cases. Of these, a full twenty-two were undertaken by the United States, the United Kingdom, Russia, or Israel.¹³³ Thus, those who assert these claims of custom tend to overly privilege and weight the practice of a handful of Western First-World states, and to either ignore or discount the inconsistent practice and explicit objections emanating from the Global South. Alonso Gurmendi has recently detailed the many ways in which the countries of Latin America have repudiated the concept, in both practice and official government statements.¹³⁴ In sum, the evidence does not reflect the kind of widespread state practice and *opinio juris* that is typically required to corroborate claims regarding the establishment of new principles of customary international law.¹³⁵

In addition to this dearth of evidence to support the claims for custom, the positions taken also arguably give far too little significance to the decisions of the ICJ, which actually stand as evidence against the proposition that this doctrine is custom. As will be discussed below, the ICJ has held that the use of force against NSAs in nonconsenting

& COMP. L.Q. 263, 276–77 (2018); Monica Hakimi, *Defensive Force Against Non-State Actors: The State of Play*, 91 INT'L L. STUD. 1, 20–30 (2015).

131. Chachko & Deeks, *supra* note 6.

132. See generally Deeks *Unwilling supra* note 7.

133. *Id.* at 501–03, 549–50.

134. Alonso Gurmendi, *Leticia & Pancho: The Alleged Historic Precedents for Unwilling or Unable in Latin America, Explored (Part I, Leticia)*, OPINIO JURIS (Nov. 7, 2018), <http://opiniojuris.org/2018/11/08/leticia-pancho-the-alleged-historic-precedents-for-unwilling-or-unable-in-latin-america-explored-part-ii-pancho-villa/> [https://perma.cc/8PUV-ZFVQ] (archived Feb. 6, 2019); Alonso Gurmendi, *State Practice regarding Self-Defence against Non-State Actors: An Incomplete Picture*, OPINIO JURIS (Oct. 22, 2018), <http://opiniojuris.org/2018/10/17/state-practice-regarding-self-defence-against-non-state-actors-an-incomplete-picture/> [https://perma.cc/AQ6F-STVN] (archived Feb. 6, 2019); Alonso Gurmendi, *The Latin American View of Jus ad Bellum*, JUST SECURITY (May 16, 2018), <https://www.justsecurity.org/56316/latin-american-view-jus-ad-bellum/> [https://perma.cc/7NEE-49B5] (archived Feb. 6, 2019).

135. See, e.g., RUYS, *supra* note 3, at 236; MALCOLM N. SHAW, INTERNATIONAL LAW 60–62 (8th ed. 2017); KAROL WOLFKE, CUSTOM IN PRESENT INTERNATIONAL LAW 162–67 (1993); Brunnee & Toope, *supra* note 130, at 276–77. Nor can the argument be made that the custom is formed by “specially affected states”—such arguments ignore both the negative manner in which that principle operates, and the neutral fashion in which one must define “specially affected.” See generally, e.g., Kevin Jon Heller, *Specially-Affected States and the Formation of Custom*, 112 AM. J. INT'L L. 191 (2018).

states in response to armed attacks is only justified if the acts of the NSA can be attributed to the territorial state.¹³⁶ Daniel Bethlehem mentions several of the decisions that are inconsistent with aspects of his principles, but he then casts doubt on their authority,¹³⁷ as do the Chatham House Principles.¹³⁸ But the judgments of the ICJ cannot be so easily discounted, particularly when making arguments about what the current state of the law is, and in the absence of convincing contrary evidence. And even when making normative arguments, it is dangerous to the international rule of law to cavalierly dismiss and disregard the judgments of the ICJ.

B. *The Perversion of Imminence*

The second problem posed by the doctrine relates to the concept of imminence. There has been considerable controversy over the concept of imminence as it relates to the doctrine of self-defense more broadly.¹³⁹ As noted above, “anticipatory self-defense” is the use of force in response to an “imminent armed attack,” which many consider to be a lawful exercise of the right of self-defense.¹⁴⁰ And as noted above, imminent has been traditionally understood to mean “about to occur,” or an attack “irrevocably in motion,”¹⁴¹ and is a concept integral to the principle of necessity. In the run up to the invasion of Iraq in 2003, there was great debate over whether “preventative self-defense” also legitimately came within the scope of the doctrine of self-defense¹⁴² (the debate is all the more confused because different people use the term “pre-emptive” self-defense as synonymous with “preventative,” while other people use “pre-emptive” to mean

136. See sources cited *infra* note 219 and accompanying text.

137. Bethlehem *A Brief Response*, *supra* note 16, at 581–82.

138. Wilmhurst, *supra* note 75, at 969–70, nn.19 & 22.

139. For a good overview, see Noam Lubell, *The Problem of Imminence in an Uncertain World*, in *THE OXFORD HANDBOOK OF THE USE OF FORCE IN INTERNATIONAL LAW* (Marc Weller ed., 2015); RUY, *supra* note 3, at 250–54. See also generally Noura S. Erakat, *New Imminence in the Time of Obama: The Impact of Targeted Killings on the Law of Self-Defense*, 56 ARIZ. L. REV. 195 (2014); Alan L. Schuller, *Inimical Inceptions of Imminence*, 18 UCLA J. INT'L L. & FOR. AFF. 161 (2014).

140. See sources cited *supra* notes 32–33 and accompanying text.

141. See sources cited *supra* note 35 and accompanying text.

142. The US assertion of the right to a preventative self-defense was made in 2002. *THE WHITE HOUSE, THE NATIONAL SECURITY STRATEGY OF THE UNITED STATES OF AMERICA* 13–15 (2002). For analysis of the debate, see, for example, DINSTEN AGGRESSION, *supra* note 21, at 221–28; RUY, *supra* note 3, at 250–54; Reisman & Armstrong, *supra* note 33. See generally David A. Sadoff, *A Question of Determinacy: The Legal Status of Anticipatory Self-Defense*, 40 GEO. J. INT'L L. 523 (2009).

“anticipatory”).¹⁴³ Preventative self-defense, also famously referred to as the “Bush Doctrine,” is the use of force in response to a growing potential for an armed attack in the future, or even a developing but still distant threat that cannot yet be classified as a threat of armed attack.¹⁴⁴ This required a reformulation of imminence as an essential aspect of self-defense, with arguments that the magnitude of the harm that will result if the threat materializes allows for an elongation of the concept of imminence, notwithstanding the lack of immediacy.¹⁴⁵ As Condoleezza Rice pithily captured the idea, and George W. Bush famously repeated, states cannot wait for a smoking gun in the form of a mushroom cloud before acting in self-defense against the threat posed by weapons of mass destruction.¹⁴⁶ But the concept of preventative self-defense enshrined in the Bush Doctrine was widely rejected after the Iraq War of 2003.¹⁴⁷

Notwithstanding the condemnation of the concept of preventative self-defense, however, the ideas about imminence that were at its core have bled into the thinking about the rationales for targeted killing and the unwilling or unable doctrine.¹⁴⁸ For instance, we see the same kinds of arguments for the elongation, or even complete elimination, of the temporal relationship between the preparation for an armed attack and its execution. This thinking was reflected most clearly in the American government arguments developed to justify the killing of Anwar al-Awlaki, the American imam who was a propagandist for AQAP. These arguments were laid out in a Department of Justice white paper,¹⁴⁹ which was based on a later-disclosed Office of Legal

143. See, e.g., CHRISTOPHER GREENWOOD, *ESSAYS ON WAR IN INTERNATIONAL LAW* 668 (2006); Lord Goldsmith, Att’y Gen., Transcript of the Iraq Inquiry Hearings: Morning Session (Jan. 27, 2010), <https://webarchive.nationalarchives.gov.uk/20171123124424/http://www.iraqinquiry.org.uk/the-evidence/transcripts-videos-of-hearings/hearings-2010-01-27/> [<https://perma.cc/4847-UQ48>] (archived Feb. 27, 2019).

144. See sources cited *supra* note 142.

145. See Cohan, *supra* note 33, at 337–45; Reisman & Armstrong, *supra* note 33; Sadoff, *supra* note 142. See generally Michael Schmitt, *Preemptive Strategies in International Law*, 24 MICH. J. INT’L L. 513 (2003).

146. *Bush: Don’t wait for mushroom cloud*, CNN (Oct. 7, 2002), <http://edition.cnn.com/2002/ALLPOLITICS/10/07/bush.transcript/> [<https://perma.cc/XWD8-2BLJ>] (archived Feb. 6, 2019).

147. DINSTEN AGGRESSION, *supra* note 21, at 221–28; RUY, *supra* note 3, at 322–42 (providing detailed primary sources as evidence of governments rejecting the Bush Doctrine).

148. See, e.g., DAVID J. BARRON, U.S. DEPT OF JUSTICE, OFFICE OF LEGAL COUNSEL, MEMORANDUM FOR THE ATTORNEY GENERAL RE: APPLICABILITY OF FEDERAL CRIMINAL LAWS AND THE CONSTITUTION TO CONTEMPLATED LETHAL OPERATIONS AGAINST SHAYKH ANWAR AL-AULAQI (2010), <https://fas.org/irp/agency/doj/olc/aulaqi.pdf> [<https://perma.cc/S28K-LNPN>] (archived Mar. 13, 2019) [hereinafter OLC AL-AULAQI MEMO].

149. The Department of Justice White Paper, *Lawfulness of a Lethal Operation Directed Against a U.S. Citizen Who is a Senior Operational Leader of Al-Qa’ida or An*

Counsel (OLC) memo.¹⁵⁰ The killing was justified (in part) on the grounds that it was a legitimate use of force in self-defense against an imminent armed attack,¹⁵¹ but the white paper articulated a theory of imminence that virtually eliminated all temporal meaning from the concept:

[T]he condition that an operational leader present an “imminent” threat of violent attack against the United States does not require the United States to have clear evidence that a specific attack on U.S. persons and interests will take place in the immediate future . . . this definition of imminence, which would require the United States to refrain from action until preparations for an attack are concluded, would not allow the United States sufficient time to defend itself . . . The United States is likely to have only a limited window of opportunity within which to defend Americans in a manner that has both a high likelihood of success and sufficiently reduces the probabilities of civilian casualties . . . Delaying action against individuals continually planning to kill Americans until some theoretical end state of the planning for a particular plot would create an unacceptably high risk that the action would fail and that American casualties would result.¹⁵²

Given that imminence is at root a temporal concept, this definition arguably hollowed out its essence and rendered it meaningless. The white paper further argued, “the threat posed by al-Qa’ida and its associated forces demands a broader concept of imminence in judging when a person continually planning terror attacks presents an imminent threat, making use of force appropriate.”¹⁵³ The argument was that once a group had plotted attacks against the United States in the past, it could be assumed that they were continuing to do so in the present, and the United States was entitled to use any window of opportunity to respond to the likelihood of such attacks in the future.¹⁵⁴ As it happens, Anwar al-Awlaki was killed more than a year after he was placed on a kill list, long after an application for an order to stop the killing brought by his father was dismissed by a federal court (in part on grounds that his father lacked standing, because al-Awlaki himself could have himself brought the claim, and because there had not yet been a violation of the law of nations),¹⁵⁵ and after several

Associated Force, <https://casebook.icrc.org/case-study/us-lethal-operations-against-al-qaida-leaders#toc-document-a-department-of-justice-white-paper-on-lethal-operations-against-al-qa-ida-leaders> (last visited Feb. 15, 2019) [<https://perma.cc/4RNJ-HHLA>] (archived Feb. 15, 2019) [hereinafter DoJ White Paper].

150. OLC AL-AULAQI MEMO, *supra* note 148.

151. DoJ White Paper, *supra* note 149, at 1.

152. *Id.* at 7.

153. *Id.*

154. *Id.* at 7–8; see also OLC Memo on Al-Aulaki, *supra* note 148.

155. Al-Aulaqi v. Obama, 727 F. Supp. 2d 1, 33, 37 (D.D.C. 2010).

missed attempts.¹⁵⁶ The government never suggested in its explanations of the strike that he had been in the midst of launching any specific attack when he was killed.¹⁵⁷

The Bethlehem Principles would appear to embrace this idea that imminence can somehow mean something other than “about to occur,” “close at hand,” or “impending”—in other words, can have a meaning that is divorced from the temporal component that is the very essence of the normal understanding of the concept of imminence.¹⁵⁸ I say “appear to” because it is not entirely clear how the five factors in Principle Eight would operate together. The first factor, it will be recalled, is “the nature and immediacy of the threat,” which would suggest an understanding of imminence in traditional temporal terms.¹⁵⁹ But the other elements militate against this. Before examining these in turn, it is helpful to note that the final concluding assertion in Principle Eight echoes the Department of Justice white paper in asserting that the lack of specific evidence as to where an attack will take place, or the precise nature of the attack, does not preclude the conclusion that an attack is imminent.¹⁶⁰ It is difficult to imagine how it could be possible to establish that an attack is virtually certain, and either in motion or immediately pending, and that the use of force is therefore the only available option to prevent it, without knowing the nature or location of the attack. The third element of Principle Eight raises similar issues—it asks whether the anticipated attack is part of a concerted pattern of continuing armed activity.¹⁶¹

156. Mark Mazzetti et al., *C.I.A. Strike Kills U.S.-Born Militant in a Car in Yemen*, N.Y. TIMES (Sept. 30, 2011), <https://www.nytimes.com/2011/10/01/world/middleeast/anwar-al-awlaki-is-killed-in-yemen.html> [<https://perma.cc/VLR7-8CTE>] (archived Feb. 6, 2019). Also killed were Samir Khan, an American American-Pakistani dual national who published an English Language on-line magazine called “Inspire” for al-Qaeda. For more detailed examination of the campaign to kill Al-Awlaki, see MARK MAZZETTI, *THE WAY OF THE KNIFE: THE CIA, A SECRET ARMY, AND A WAR AT THE ENDS OF THE EARTH* (2013) [hereinafter MAZZETTI KNIFE]; JEREMY SCAHILL, *DIRTY WARS: THE WAR IS A BATTLEFIELD* (2013).

157. SCAHILL, *supra* note 156; Conor Friedersdorf, *How Team Obama Justifies the Killing of a 16-Year-Old American*, THE ATLANTIC (Oct. 24, 2012), <https://www.theatlantic.com/politics/archive/2012/10/how-team-obama-justifies-the-killing-of-a-16-year-old-american/264028/> [<https://perma.cc/JT64-YUJ8>] (archived Feb. 6, 2019).

158. It is instructive to consider how American courts have interpreted the concept of imminence as part of the doctrine of self-defense in domestic criminal law. *See, e.g.*, *State v. Norman*, 378 S.E.2d 8, 13 (Sup. Ct. N.C. 1989) (“[T]he term ‘imminent,’ . . . has been defined as ‘immediate danger, such as must be instantly met, such as cannot be guarded against by calling the assistance of others or the protection of the law.’ Or cases have sometimes used the phrase ‘about to suffer’ interchangeably with ‘imminent’ to describe the immediacy of threat that is required to justify killing in self-defense.”).

159. *See supra* Part II.D. *See supra* text accompanying note 105.

160. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 775–76.

161. *Id.*

Given that the principles define “armed attack” as including an accumulation of small strikes, this element of Principle Eight could be taken to mean that when the next sporadic terrorist strike in a series or pattern of such strikes is deemed to breach the threshold in making all of them together an armed attack for purposes of justifying self-defense (an issue to be analyzed below), then this next strike need not be already in motion or immediately pending in order for it to be “imminent.”¹⁶² In other words, such a potential future strike may be classified as “imminent” based on the nontemporal consideration of whether it is determined to be part of a pattern of strikes that cumulatively constitute an armed attack, and force may be used in self-defense in response to that future strike, no matter how distant in time it may be.

The second and fourth of the five elements of Principle Eight most clearly reflect the essential logic of the Bush Doctrine, in that they require an assessment of the probability of an attack, as well as the likely scale of the attack and the likely harm it will cause.¹⁶³ It is quite true that a standard formula for measuring the risk associated with an event is to assess the probability of that event occurring, multiplied by the magnitude of the harm it would cause.¹⁶⁴ This formula is not, however, a proper measure of imminence. Risk and imminence are entirely different concepts. The product of probability and harm has nothing to do with the timing of a risk materializing, or the time available to find ways to avoid the harm.¹⁶⁵ Risk as measured in this way is no doubt a vitally important policy consideration in trying to develop responses to the threats of both weapons of mass destruction and transnational terrorist attacks. But it is improper to smuggle the notion of risk into the concept of imminence for the purposes of

162. See *infra* Parts III.C, IV.A.

163. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 775.

164. See, e.g., MARVIN RAUSAND, RISK ASSESSMENT: THEORY, METHODS, AND APPLICATIONS 1–28 (2013). Tort aficionados will also note that this is inherent in the Learned Hand test, which is essentially a question of whether the burden of avoiding the risk is greater than the risk itself. *United States v. Carroll Towing Co.*, 159 F.2d 169, 173 (2d Cir. 1947).

165. Akande and Lieflander, in an article critical of the Bethlehem Principles, do however accept that these factors do serve as two of the essential elements of imminence, and even that where a threat is sufficiently probable and severe, “the mere fact that it is still temporally remote should provide no injunction against action where that action is necessary and proportionate.” But the last clause of course begs the question, since imminence is an element of necessity; and in any event, even they do not eliminate the temporal element, but develop something closer to the “last clear chance” concept that I consider below. Dapo Akande & Thomas Lieflander, *Clarifying Necessity, Imminence, and Proportionality in the Law of Self-Defense*, 107 AM. J. INT’L L. 563, 565 (2013).

lowering the legal threshold for the use of force in self-defense. The occurrence of some event can be both quite probable and very likely to cause massive harm, and yet still be far in the future, thus providing lots of time to pursue various alternatives for avoiding or eliminating the risk.¹⁶⁶ Crucially, for the purposes of the doctrine of self-defense, this formulation of probability and magnitude of harm does not make a use of force in response necessary. The very fact that we have time before the risk is likely to materialize means that we at least have an obligation to search for alternatives to the use of force as a means of preventing the attack, including ways to change the conditions and circumstances that make the event probable in the first place. This search for alternatives brings us to the last factor, and perhaps the most interesting and complex aspect of the problem of imminence.

The fifth factor in Principle Eight is 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.”¹⁶⁷ There are two problems with this that need to be untangled and addressed separately. The first is that by including considerations of collateral damage, it not only introduces elements that have nothing to do with the immediacy of the attack, but once again conflates concepts from IHL into a *jus ad bellum* analysis. These considerations of collateral harm are of course absolutely necessary for the analysis of whether the available options for a use of force in self-defense would be lawful in IHL terms in any given context, but they have no bearing on the imminence of the attack being responded to. If the IHL analysis indicated that waiting to respond might increase the probability of collateral damage, even to the point of making the planned operations unlawful, that would not make the attack any more imminent so as to justify an earlier response. The second problem is raised by the “other opportunities” notion.¹⁶⁸ On this view, imminence is separated from its temporal roots in a slightly different manner. Rather than using imminence as meaning that the attack is about to occur, as the common notion of the concept would suggest, it is used to mean that the window of opportunity for preventing the attack is about to close. The attorney general of Australia in a recent speech clearly articulated this “last feasible window of opportunity” conception of imminence, and several commentators responding to his remarks

166. The ICJ itself has, in the context of necessity in the law of state responsibility, itself addressed the relationship between risk and imminence. See *Gabcikovo-Nagymaros Project (Hung. v. Slov.)*, Judgment, 1997 I.C.J. Rep. 7, 42 (Sept. 25). (“[I]mminence’ is synonymous with ‘immediacy’ or ‘proximity’”, and while the word peril “certainly evokes the idea of ‘risk’,” the peril “must have been a threat to the interest at the actual time.”).

167. *Bethlehem Self-Defense Against NSAs*, *supra* note 16, at 775–76.

168. *Id.*

noted that this actually blurs or indeed collapses the distinction between “anticipatory” and “preventative” self-defense, effectively embracing the Bush Doctrine.¹⁶⁹

This criticism of the “last window of opportunity” test is partly right, but it does not quite capture the entire problem, nor does it fully address the valid intuition that underlies this attempt to reformulate the concept of imminence. This reflects what I think is a broader problem with the imminence debate, which is that both sides have tended to lose sight of the fact that imminence is a crucial element of the principle of necessity—and that its meaning and significance always have to be assessed and understood in the context of the role it plays in establishing the necessity of a use of force. Necessity is the core principle of the doctrine of self-defense, meaning that a use of force in response to an armed attack must be a last resort, the only alternative, to preventing the attack—and imminence only matters for purposes of establishing that necessity. To refer back to the *Caroline* incident, the “no moment for deliberation” was inextricably tied to the “no choice of means” as factors explaining necessity.¹⁷⁰ Imminence and immediacy are simply factors for establishing that there are no alternatives to the use of force available as a means of responding to and preventing the anticipated or ongoing armed attack. If the anticipated attack is not imminent, in the traditional sense of being in the immediate future, then typically—though not always—there is time to explore other options that will be still available, and the use of force would not truly be a “last resort.” When, for instance, former Prime Minister Menachem Begin conceded, in a speech at Israel’s National Defense College in 1982, that on the eve of Israel’s preemptive strikes that commenced the 1967 war, Israel had “had choices,” he was essentially admitting that the Israeli use of force had not been a necessary last resort to prevent an imminent Egyptian attack.¹⁷¹

The valid intuition of the “last window of opportunity” formulation, however, is that there may be situations in which there

169. Brandis, *supra* note 57; see, e.g., Adil Haque, *Imminence and Self-Defense Against Non-State Actors: Australia Weighs In*, JUST SECURITY (May 30, 2017), <https://www.justsecurity.org/41500/imminence-self-defense-non-state-actors-australia-weighs/> [<https://perma.cc/J8UD-L779>] (archived Feb. 6, 2019).

170. See *supra* text accompanying notes 65–66.

171. RICHARD HAAS, WAR OF NECESSITY WAR OF CHOICE 9–10, 299 (2009). *But see* DINSTEN AGGRESSION, *supra* note 21, at 211–12 (“That [an Egyptian armed attack was at an early stage], at least, was the widely shared perception . . . in June 1967, based on sound judgment of events. Hindsight knowledge, suggesting that . . . the situation may have been less desperate than it appeared, is immaterial.”).

are actually no alternatives remaining, and the necessity of response is thus immediate, even if the attack itself is not imminent in the traditional sense.¹⁷² A thought experiment may help illustrate this: suppose we learn somehow that an invading force of a more advanced alien life form is on its way to Earth, and that it will arrive in ten years. Ten years is not immediate or imminent in the traditional sense. But what if we also learn that it will take us close to ten years to develop an adequate defense to the invading force? Then, all of a sudden, the necessity of response takes on the character of “instant, overwhelming, leaving no choice of means and no moment of deliberation.”¹⁷³ To hesitate is to be doomed. It is possible to characterize a threat such as this as imminent, not in the sense that the harm itself will materialize in the immediate future, but in the sense that a particular response is immediately necessary, in that there are no alternative options, and if the response is not immediate it will be too late and thus ineffective.¹⁷⁴ Climate change is a real example of this form of imminent threat, in that the materialization of the risk of harm is not immediately at hand (though in 2019 that assertion is increasingly questionable), but if an immediate response is not undertaken it will be impossible to prevent the materialization of catastrophic harm in the future.¹⁷⁵

It is important to note, however, that in the context of the doctrine of self-defense, even this conception of imminence—let us call it the true “last clear chance” formulation—continues to be an essentially temporal concept and continues to operate primarily as a factor for establishing necessity. In essence, it continues to be about the time remaining within which alternatives to the use of force can be explored before a use of force is the only remaining means of preventing or responding to an attack. It is this form of imminence that the “last window of opportunity” test is imperfectly grasping for.

I say that the unwilling or unable doctrine formulation is imperfect, and different from this “last clear chance” formulation, for two reasons. First, as a practical matter, in the context of action against NSAs in nonconsenting states, the target state is virtually never going to have sufficiently compelling evidence that: (i) an armed attack in the future is certain (a conclusion that requires compelling evidence of both-capability and a decision reflecting specific intent); (ii) the current window of opportunity to use force is truly the last and only

172. This is explored as well by Akande & Lieflander, *supra* note 165, at 565.

173. Miller, *supra* note 66.

174. This thought experiment comes from Eliezer Yudkowsky, who used it in a podcast to explain how and why the threat posed by artificial general intelligence can be viewed as being imminent. *AI: Racing Toward the Brink: A Conversation with Eliezer Yudkowsky*, SAM HARRIS (Feb. 6, 2018), <https://samharris.org/podcasts/116-ai-racing-toward-brink/> [<https://perma.cc/3RMX-Y3HU>] (archived Feb. 27, 2019).

175. See generally, e.g., INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE, GLOBAL WARMING OF 1.5°C (2018), https://report.ipcc.ch/sr15/pdf/sr15_spm_final.pdf [<https://perma.cc/3R8P-QWBL>] (archived Feb. 6, 2019).

chance to prevent the threat from materializing; and, therefore, (iii) there are no other alternatives remaining to prevent the attack, and thus the immediate use of force is necessary. And all three of these elements must arguably be satisfied to establish the true “last clear chance” element of imminence and thus necessity. But in the context of the unwilling or unable doctrine, it is far more likely that the target state will have established that there is “a chance” to target members of the NSA, and that it will have little way of knowing that it is “the last chance” to use force to prevent the future attack, or that there will be no other alternatives to the use of force in the time available to explore them.¹⁷⁶

This is compounded by the second reason that the doctrine’s window of opportunity element is different from the “last clear chance” formulation, which is that it is coupled with elements that effectively strip “imminence” of its essential temporal meaning. Because, while formulated as something like a “last clear chance” conception of imminence, it really just confuses an attractive current or imminent window of opportunity to strike the NSA with the idea that an *immediate response* is the *only way* to prevent a certain specific armed attack. What is missing is any requirement to establish that there is a true necessity to act immediately, that there are no alternatives available—including, importantly, through consultation and possible collaboration with the territorial state. This conception of imminence tends to smuggle in aspects of the Bush Doctrine,¹⁷⁷ and if it is legitimated and entrenched in the approach to the use of force against nonconsenting states harboring NSAs, it may in turn influence thinking about imminence in other *jus ad bellum* contexts.

C. Incoherence on Armed Attack

Closely related to the problems with the conception of imminence is an apparent incoherence in how the principles characterize and rely upon the concept of “armed attack.” The problem is twofold. First, the manner in which the principles define armed attack is a significant

176. See generally, e.g., Philip Alston, *The CIA and Targeted Killings Beyond Borders*, 2 HARV. NAT. SEC. J. 283 (2011) [hereinafter Alston *CIA*] (in the context of targeted killing); Haque, *supra* note 169 (on these evidentiary difficulties). In this sense, the “last clear chance” test shares some features of the “ticking time bomb” hypothetical in the torture debate, which depends on premises that require a level of knowledge that in practical terms is virtually impossible to establish or verify—and thus, while providing a veneer of principled theoretical justification, the hypothetical is highly misleading as any kind of guide to policymaking.

177. See Haque, *supra* note 169; *supra* text accompanying note 147–154.

departure from the established understanding in international law.¹⁷⁸ Second, even if we were to accept the definition embraced by the Bethlehem Principles, that definition cannot be coherently reconciled with other parts of the argument that target states make, and which the Bethlehem Principles implicitly support, to justify uses of force against nonconsenting states.

Beginning with the problem of characterization, it will be recalled that Principle Four of the Bethlehem Principles defines “armed attack” as including “both discrete attacks and a series of attacks that indicate a concerted pattern of continuing armed activity.”¹⁷⁹ It notes that the distinction between discrete attacks and a series of attacks may be relevant to considerations of the necessity to respond in self-defense. Similarly, the third element of the five elements for assessing imminence in Principle Eight calls for consideration of whether the anticipated attack is part of a concerted pattern of continuing armed activity.¹⁸⁰ All of this suggests that the principles embrace the view that a series of strikes may in the aggregate constitute an “armed attack” even though no one of the strikes on its own would satisfy the test. The United States, of course, does not accept that there is a material gap between a use of force that would violate Article 2(4) of the Charter, and an armed attack for purposes of Article 51 of the Charter.¹⁸¹ But the US view is an outlier.¹⁸² The ICJ has made clear that there is a sizeable gap between the two.¹⁸³ Moreover, in the *Oil Platforms* case, it clearly established that the threshold for establishing an armed attack is quite high, with neither a mine strike on a naval vessel nor a missile strike on an oil tanker, both of which were part of a pattern of attacks, satisfying the test.¹⁸⁴ A careful analysis of state practice and *opinio juris* suggests that the vast majority of the international community subscribes to the more traditional ICJ view.¹⁸⁵ And while there is some academic support for the idea that “an accumulation of events” may in the aggregate constitute an armed attack, it is arguably a minority view that is not

178. The seminal work on armed attack is RUY, *supra* note 3.

179. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 775.

180. *Id.*

181. Schmitt & Goodman, *supra* note 39; see also William H. Taft, *Self-Defense and the Oil Platforms Decision*, 29 YALE J. INT'L L. 295, 300–02 (2004) (views of the then-legal counsel to the State Department).

182. GRAY, *supra* note 21, at 153–57; RUY, *supra* note 3, at 143–57. *But see* DINSTEN AGGRESSION, *supra* note 21, at 209–11.

183. Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), Judgment, 1986 I.C.J. Rep. 14, ¶¶ 189–91 (June 27).

184. Case Concerning Oil Platforms (Iran v. U.S.), 2003 I.C.J. Rep. 160, ¶¶ 51, 64, 161–219 (Nov. 6). For a critique of this decision, see Taft, *supra* note 181.

185. RUY, *supra* note 3, 143–74.

yet reflected in established law.¹⁸⁶ Thus, the Bethlehem Principles incorporate a conception of armed attack that is inconsistent with established law. Moreover, they do not provide any guidance whatsoever as to the standards that might apply in determining when a series of attacks by disparate but purportedly associated groups satisfies this test of “concerted pattern of continuing armed activity” sufficient to constitute an armed attack. It is worth noting that similar issues arise in trying to determine whether conduct that is alleged to constitute crimes against humanity is sufficiently “widespread” or “systemic” to satisfy the definition, and a growing body of jurisprudence and scholarship has grappled with developing the contours of the concept.¹⁸⁷

Target states and the Bethlehem Principles are advancing a lower threshold for armed attack precisely because the risk posed by transnational terrorist organizations is characterized by exactly this pattern of sporadic, repeated, lower-intensity attacks. These are attacks against civilian targets, often months apart, in disparate locations, and often on the scale of car bombs, efforts to bring down a single plane, or vehicular attacks against pedestrians.¹⁸⁸ Attacks are launched by different groups that have ill-defined affiliations with one another. Attacks on the scale of 9/11, which certainly cross the threshold for an armed attack, are more the exception rather than the rule. I do not want to be misunderstood as suggesting that NSAs cannot be responsible for armed attacks as a matter of law,¹⁸⁹ but rather that the low-intensity and sporadic attacks often mounted by terrorist organizations do not, individually or in the aggregate, rise to the level of an armed attack as that term has been traditionally understood in *jus ad bellum*.¹⁹⁰ What is more, even if we were to accept

186. For some support for the “accumulation” argument, see, for example, DINSTEIN AGGRESSION, *supra* note 21, at 209–11. *Contra* GRAY, *supra* note 21, at 153–57; RUYS, *supra* note 3, at 168.

187. See, e.g., Christopher Roberts, *On the Definition of Crimes Against Humanity and Other Widespread or Systemic Human Rights Violations*, 20 U. PENN. J.L. & SOC. 1, 20–21 (2017).

188. There is, of course, a huge literature addressing the debate over the definition of terrorism, and what conduct comes within the scope of terrorist activity. For discussion of the problem, see generally, for example, Sarah V. Marsden, *Typologies of Terrorism and Political Violence*, in THE ROUTLEDGE HANDBOOK OF TERRORISM RESEARCH (Alex P. Schmid ed., 2011); Ben Saul, *Attempts to Define ‘Terrorism’ in International Law*, 52 NETH. INT’L L. REV. 57 (2005).

189. See *supra* text accompanying notes 41–43.

190. GRAY, *supra* note 21, at 163–65; RUYS, *supra* note 3, at 178. *But see* DINSTEIN AGGRESSION, *supra* note 21, at 210–11. Even the British inquiry into the 2015 targeted killing of Reynard Khan questioned the validity of this accumulation of strikes conception of armed attack. JOINT COMMITTEE ON HUMAN RIGHTS, THE GOVERNMENT’S POLICY ON THE USE OF DRONES FOR TARGETED KILLING, 2015–16, HC 574, HL Paper 141, at 44 (UK).

the conception of armed attack articulated in the principles, it does not provide any standards for determining where to draw the line regarding the number of such low-level strikes and what level or scale of violence or damage would be necessary to constitute an armed attack in aggregate. Thus, even as we can appreciate the policy imperatives, we must question the manner in which the doctrine is trying to lower the bar to permit the exercise of self-defense in response to a series of such low-level attacks, because it would lower the bar for all purposes, for the use of force against states as well as against terrorists. A series of small border incidents, or naval incursions by one state into the territorial waters of another, months or even years apart, could all of a sudden be taken in the aggregate as an "armed attack" putatively justifying a use of force in self-defense, leading to an escalating international armed conflict.

Even if one were to accept the proposition that a series of low-intensity terrorist attacks can cumulatively rise to the level of constituting an armed attack, thereby triggering the right of self-defense, it is not consistent with the other way in which the concept of armed attack is used to justify the uses of force against NSAs in nonconsenting states. The United States and other target states have been in the practice of explaining each such use of force as an independent act of self-defense in response to an actual or imminent armed attack.¹⁹¹ In other words, each one of a series of drone strikes some weeks or months apart against NSAs in Pakistan, each of which constitutes a use of force against a nonconsenting state, is separately justified as an act of self-defense against an independent imminent armed attack. But the imminent attacks to which each one of these drone strikes is responding consist of sporadic, low-level terrorist activity, and thus do not individually rise to the level of an "armed attack."¹⁹² As was just examined, it is precisely to get around this apparent problem that target states have argued, backed up by the Bethlehem Principles, that an accumulation of small strikes, no one of which might be sufficient to constitute an armed attack, can form a pattern of strikes that in aggregate can satisfy the test for armed attack.¹⁹³ But even if we were to accept that these two arguments could each be plausible on their own, they become far less so when they are taken together, as they must be. Each use of force in response to a low-scale strike cannot be an individual act of self-defense against an

191. US LEGAL AND POLICY FRAMEWORKS, *supra* note 2; DoJ White Paper, *supra* note 149.

192. For discussion of the threshold for armed attack, see sources cited *supra* notes 182–185 and accompanying text.

193. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 774–75; see, e.g., Taft, *supra* note 181.

imminent armed attack, when at the same time it is argued that each strike that is being defended against does not individually rise to the level of armed attack but constitutes one armed attack when aggregated with others. These two claims are internally incoherent, and yet they are provided for in the principles as a way of lowering the threshold for the exercise of self-defense against NSAs in nonconsenting states.¹⁹⁴ And in the process, they lower the bar to the use of force in self-defense against states as well.

D. *The Conflation of Jus ad Bellum with Other Regimes*

The fourth major problem with the Bethlehem Principles relates to the relationship between *jus ad bellum* and other regimes in international law, specifically IHL and the law of state responsibility. The Bethlehem Principles conflate rules and principles from these different regimes in ways that threaten to undermine the nature of the well-established relationships between *jus ad bellum* and both IHL and the law of state responsibility and weaken the integrity and coherence of all three distinct legal regimes.

As explained above, Principles Six and Seven of the Bethlehem Principles provide that the target state may use force against all those actively planning, threatening, or perpetrating armed attacks, as well as those who are reasonably believed to be taking direct part in those attacks through the provision of material support essential to the attacks.¹⁹⁵ Bethlehem explicitly explained that this is separate and apart from the IHL rule governing the targeting of those taking direct part in hostilities.¹⁹⁶ Yet this proposition, that target states may, in accordance with the authority conferred by principles of *jus ad bellum* under the unwilling or unable doctrine, target individuals who provide “material support essential to armed attacks,” is inconsistent with both the IHL rule itself, and with the long established relationship between

194. The analysis is complicated further by differences in view on the extent to which each strike against the nonconsenting state triggers an international armed conflict. The United States takes the minority view that such strikes do not trigger an international armed conflict at all. See, e.g., Dapo Akande, *When Does the Use of Force Against a Non-State Armed Group Trigger an International Armed Conflict and Why Does it Matter?*, EJIL: TALK! (Oct. 18, 2016), <https://www.ejiltalk.org/when-does-the-use-of-force-against-a-non-state-armed-group-trigger-an-international-armed-conflict-and-why-does-this-matter/> [https://perma.cc/XTX7-EF2L] (archived Feb. 7, 2019) (discussing the debate); Paust, *supra* note 44.

195. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 774.

196. *Id.*

the *jus ad bellum* and IHL regimes. It is the *lex specialis* of IHL, not *jus ad bellum*, that dictates who may be targeted in armed conflict.

Under the rules of IHL, “civilians,” a term of art that includes effectively anyone who is not a “combatant” in an international armed conflict as defined in the Geneva Conventions of 1949, cannot be targeted “unless and for such time as they take a direct part in hostilities.”¹⁹⁷ What precisely constitutes the parameters for defining the scope of “taking a direct part in hostilities,” and thus who can be targeted and when they can be targeted, is heavily debated in academic, operational, and judicial writing.¹⁹⁸ Moreover, there is some debate over whether members of transnational terrorist organizations who are continuously engaged in hostilities should be properly classified as something akin to “combatant”—the leading proposed status is that of “continuous combat function”—rather than as “civilian.”¹⁹⁹ Nonetheless, these are first and foremost concepts of IHL, and these debates must be resolved by reference to principles and concepts internal to the IHL regime.²⁰⁰ Principles Six and Seven of the Bethlehem Principles would purport to trump IHL rules on who may and may not be targeted by providing separate, and presumably superior, authority for targeting members of NSAs who might not be targetable under the rules of IHL.²⁰¹ The idea that individuals could be targeted for providing “material support” would suggest a standard that is far below that which would be acceptable under virtually all conceptions of taking a “direct part in hostilities” in IHL. If this standard from the Bethlehem Principles were to be accepted as law, there would be the risk that IHL and *jus ad bellum* would thus provide conflicting rules on who could be lawfully targeted in the circumstances of uses of force against NSAs in nonconsenting states.²⁰² That would

197. Geneva Conventions, *supra* note 49, at common art. 4; Additional Protocol I, *supra* note 48, arts. 50, 51(3); WILLAM H. BOOTHBY, *THE LAW OF TARGETING* 141–64 (2012); HENCKAERTS, *CUSTOMARY IHL*, *supra* note 47, at 17–19; INT’L COMM. OF THE RED CROSS, *INTERPRETIVE GUIDANCE ON THE NOTION OF DIRECT PARTICIPATION IN HOSTILITIES UNDER INTERNATIONAL HUMANITARIAN LAW* 41–68 (2009).

198. See, e.g., DINSTEIN NIAC, *supra* note 49, at 121–25, 146–52; HENCKAERTS, *CUSTOMARY IHL*, *supra* note 47, at 17–24; Nilz Melzer, *The Principle of Distinction Between Civilians and Combatants*, in *THE OXFORD HANDBOOK OF INTERNATIONAL LAW IN ARMED CONFLICT* 296 (2013).

199. See, e.g., sources cited *supra* note 198; see also The Public Committee Against Torture in Israel v. Israel, HCJ 769/02 (2002) (Sup. Ct. Israel); Curtis A. Bradley, *The United States, Israel, and Unlawful Combatants*, 12 GREEN BAG 2D 397 (2009).

200. See sources cited *supra* note 50 (discussing the relationship between IHL and *jus ad bellum*).

201. DINSTEIN NIAC, *supra* note 49, at 3.

202. Wilmschurst and Wood similarly criticized this aspect of Principle Seven as being in violation of both *jus ad bellum* and IHL. Elizabeth Wilmschurst & Michael Wood, *Self-Defense Against Nonstate Actors: Reflections on the “Bethlehem Principles,”* 107 AM.

be profoundly problematic, not merely because of the resulting erosion of protections for civilians, as well as the uncertainty and instability created around these specific rules of IHL, but because it would potentially undermine the nature of the relationship between *jus ad bellum* and IHL more generally.

There is a similar problem with the manner in which the Bethlehem Principles appear to imply some importation of principles from the law of state responsibility for purposes of assessing lawfulness under the *jus ad bellum* regime. Specifically, Bethlehem in his article suggested that the law of state responsibility may apply, and in Principle Eleven stated that the extent of the responsibility of the harboring or colluding territorial state, for its failure to neutralize the threat emanating from within its territory, may impact whether the target state need seek explicit consent for the use of force.²⁰³ Others have similarly invoked the notion of state responsibility in explaining the doctrine.²⁰⁴ As I will explain later in this Part, even the reliance upon the “inability” of the territorial state to prevent NSA activity as a justification for the use of force is an implicit conflation of state responsibility with *jus ad bellum*.²⁰⁵ This apparent reliance upon state responsibility in elaborating the rationale for the unwilling or unable doctrine would seem to suggest that state responsibility can serve as part of the justification for the use of force. If that is indeed the intended meaning, however, it is not consistent with the doctrine of self-defense or the *jus ad bellum* regime more generally. It either reflects some confusion over the exact nature of the state responsibility in question, or, more seriously, an attempt to use principles of the law of state responsibility to lower the threshold for the justifiable use of force in the doctrine of self-defense.

This does not mean to suggest that the law of state responsibility is not implicated in the circumstances of armed groups launching attacks against a state from within the territory of another state. To be clear, it is indeed the case that a territorial state would incur responsibility for either refusing or failing to prevent NSAs from harming other states from within its territory. There is an international legal obligation on states not to “allow knowingly [their] territory to be used for acts contrary to the rights of other States,”

J. INT'L L. 390, 394 (2013). *But see* Benvenisti, *supra* note 50 (questioning the clear distinction between *jus ad bellum* and IHL in conflicts with NSAs).

203. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 773, 774, 776.

204. *See, e.g.*, Travalio & Altenburg, *Terrorism and State Responsibility*, *supra* note 79, at 110, 116.

205. *See infra* text accompanying notes 224–227.

which would include an obligation not to knowingly allow NSAs within their territory to cause harm to other states.²⁰⁶ It is this obligation that the Bethlehem Principles incorporate for purposes of Principle Nine—which articulates the obligation of states to prevent NSAs from launching attacks from within their territory.²⁰⁷ The violation of this specific obligation will thus attract state responsibility. The failure to prevent harm constitutes a violation attracting state responsibility to the same extent as some unlawful affirmative act of the state.²⁰⁸ There is nothing incorrect or controversial about the Bethlehem Principles articulating this form of state responsibility on the part of a state unwilling or unable to prevent the materialization of threats from an NSA operating within its territory.²⁰⁹

It is important, however, to be very clear that the finding of state responsibility in the situation just described does not depend upon an attribution to the state of any of the unlawful acts of the NSAs. The violation of international law that is at issue here, for purposes of triggering the responsibility of a territorial state that is unwilling or unable to neutralize the threat posed by an NSA, is precisely and only the failure to fulfill its *obligation to prevent* the NSA use of its territory in ways that cause harm to others. And that violation and corresponding responsibility can in no way justify a use of force under well-established understanding of both the law of state responsibility and *jus ad bellum*.

The NSA is itself violating international law in launching attacks against other states, but the territorial state is not responsible for those violations of international law, unless and until the acts of the NSA can be attributed to the territorial state.²¹⁰ The law of state responsibility does provide that states will be responsible when a breach of an international legal obligation by an NSA is attributable to

206. The Corfu Channel Case (UK v. Alb.), Judgment, 1949 I.C.J. Rep. 4, 22 (Apr. 9); see also JAMES CRAWFORD, *STATE RESPONSIBILITY: THE GENERAL PART* 158 (2013). More specifically, with regard to terrorist activity, see S.C. Res. 748 (Mar. 31, 1992); S.C. Res. 1373 (Sept. 28, 2001).

207. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 776.

208. Case Concerning Application of the Convention on the Prevention and Punishment of the Crime of Genocide (Bosn. & Herz. v. Serb. & Montenegro), Judgment, 2007 I.C.J. Rep. 43, ¶¶ 428–38 (Feb. 26) [hereinafter *Genocide Convention Case*]; see also Rep. of the Int'l Law Comm'n on the Work of Its Fifty-Third Session, U.N. Doc. 1/56/10, art. 14(3) [hereinafter *ILC Draft Articles*]; CRAWFORD, *supra* note 206, at 217–18, 226–30; Kristen E. Boon, *Are Control Tests Fit for the Future? The Slippage Problem in Attribution Doctrines*, 15 MELB. J. INT'L L. 330, 363–64 (2014).

209. CRAWFORD, *supra* note 206, at 226–7 and 405–06 (noting the distinction between the standard for liability for failing to prevent and that for complicity, in the context of proving aid or assistance in the commission of a breach).

210. ILC *Draft Articles*, *supra* note 208, arts. 4–11; see also CRAWFORD, *supra* note 206, at 146–56.

the state, and there is a specific test for determining such attribution.²¹¹ If the actions of the NSA *can* be attributed to the territorial state in accordance with the law of state responsibility, then the territorial state is actually responsible for two distinct violations: first, failing to prevent harm to another state from emanating from within its territory as explained above; and second, for the actual violations of international law committed by the NSA. Where the acts of the NSA rise to the level of armed attack on another state, then this will be a violation of the rules of *jus ad bellum*, and if attributable to the territorial state, it will be responsible for such violation.²¹² But it must be emphasized that the state responsibility that is being referred to in the discussion of the unwilling or unable doctrine would appear to be the first, not the second. Because if the acts of the NSA *can* be attributed to the state, in accordance with the rules of the law of state responsibility, then they can also almost certainly be attributed to the territorial state for purposes of the doctrine of self-defense, thus justifying the use of force against that state under the principles of *jus ad bellum*, and reference to state responsibility would be entirely unnecessary.²¹³ So the reference to state responsibility in the Bethlehem Principles would appear to be trying to add some additional basis for the use of force, and it is this that is improper.

To explain this last point further, the tests for attribution under state responsibility and under the doctrine of self-defense are at once quite independent of one another, but nonetheless similar. Both the differences and the similarities matter. The acts of an NSA can be attributed to the state for purposes of state responsibility if, among other things, it is “in fact acting on the instructions of, or under the direction or control of, that State in carrying out the conduct,” or where the NSA is “in fact exercising elements of the governmental authority”

211. *Genocide Convention Case*, 2007 I.C.J. Rep. at ¶¶ 385–415; ILC *Draft Articles*, *supra* note 208, art. 8; CRAWFORD, *supra* note 206, at 146–56.

212. See GRAY, *supra* note 21, at 139–45; RUYS, *supra* 3, at 226 (discussing attribution for purposes of self-defense); *Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.)*, Judgment, 1986 I.C.J. Rep. 14, ¶¶ 194–95 (June 27); *Case Concerning Armed Activities on the Territory of the Congo (Dem. Rep. Congo v. Uganda)*, Judgment, 2005 I.C.J. Rep. 1, ¶¶ 106–47 (Dec. 19); see also Paddeu, *supra* note 45.

213. See ILC *Draft Articles*, *supra* note 208, arts. 5, 8, 9 (discussing attribution for purposes of state responsibility); see also CRAWFORD, *supra* note 206, at 146–56. This assumes, of course, that the NSA is not in fact an organ of the territorial state, or an organ of another state placed at the disposal of the territorial state. ILC *Draft Articles*, *supra* note 208, arts. 5–7.

in place of the government of the territorial state.²¹⁴ The actions of an NSA can also be attributed to the state within which it is operating if the state *ex post* “acknowledges and adopts” the conduct of the NSA as its own.²¹⁵ The fact that the conduct of an NSA can be attributed to the state for purposes of the law of state responsibility, however, does not by itself justify the use of force against the state. The idea that states may use force as a form of countermeasure in response to violations of international law ended with the modern UN system.²¹⁶ The justification for the use of force must be found in the *jus ad bellum* regime, not the law of state responsibility.²¹⁷ The traditional understanding of *jus ad bellum* is that in order to justify the use of force in self-defense against a state for the actions of NSAs acting within its territory, the actions of the NSAs must rise to the level of armed attack and be attributable to the state.²¹⁸ The ICJ has repeatedly held that the test for attribution for purposes of self-defense—that is, for the purpose of justifying the use of force against the state in which the NSAs are located—is that the territorial state was “substantially involved” in the operations of the NSA.²¹⁹ In the case of the use of force against al-Qaeda in the invasion of Afghanistan after 9/11, for instance, both the US letter to the Security Council pursuant to Article 51 of the Charter,²²⁰ and the Security Council resolutions passed in relation to this use of force, characterized the Taliban’s acquiescence in the actions of al-Qaeda and their subsequent refusal to take action against al-Qaeda as having constituted

214. ILC *Draft Articles*, *supra* note 208, arts. 5–9; *see also* CRAWFORD, *supra* note 206, at 146–56.

215. ILC *Draft Articles*, *supra* note 208, art. 11.

216. *See, e.g.*, NEFF, *supra* note 23, at 318; *see also* ILC *Draft Articles*, *supra* note 208, arts. 21–22. *But see* DINSTEIN AGGRESSION, *supra* note 21, at 269 (arguing that “defensive armed reprisals can be a permissible form of self-defence”).

217. CRAWFORD, *supra* note 206.

218. GRAY, *supra* note 21.

219. *Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.)*, Judgment, 1986 I.C.J. Rep. 14, ¶¶ 194–95 (June 27); *Case Concerning Armed Activities on the Territory of the Congo (Dem. Rep. Congo v. Uganda)*, Judgment, 2005 I.C.J. Rep. 1, ¶¶ 106–47 (Dec. 19) (explaining, based on the facts presented by Uganda, that Uganda was not acting in self-defense); GRAY, *supra* note 21, at 139–45, 226. It is worth noting that the “effective control” test that was also elaborated in *Nicaragua* was to assess the attribution of NSA (Contra) violations of IHL to the controlling state (the United States), for purposes of state responsibility—not to determine attribution for purposes of justifying the use of force in self-defense or even the use of force. *Nicaragua*, 1986 I.C.J. Rep. at ¶ 115. Scholars continue to confuse this issue, writing that the “effective control” test was employed by the Court in *Nicaragua* to assess whether US support for the Contras constituted an unlawful use of force. *See, e.g.*, Travalio, *supra* note 79, at 103. This is wrong.

220. Letter from John Negroponte, Rep. of the U.S. to the U.N., to the President of the Security Council (Oct. 7, 2001), http://avalon.law.yale.edu/sept11/un_006.asp [<https://perma.cc/TVM2-ZWEZ>] (archived Feb. 17, 2019).

sufficiently “substantial involvement” so as to ground attribution of al-Qaeda’s actions to Afghanistan.²²¹ As I will return to below, the *jus ad bellum* attribution test is an important factor in considering the extent to which a state is “unwilling,” and the nature of unwillingness that should justify the use of force—for at some point the unwillingness of the territorial state to either deal with the NSA threat or consent to the target state taking such action within its territory, will rise to the level of support that constitutes “substantial involvement” in the NSA’s activity, and thus a traditional justification for use of force in self-defense.

In sum, the two tests of attribution are conceptually distinct, but the tests for each of them are fairly similar.²²² Both are difficult to establish, and the reality is that if one cannot establish attribution of NSA conduct to the state for purposes of state responsibility, it is also unlikely that one will be able to attribute the conduct to the state for purposes of self-defense. On the other hand, if the acts of the NSA can be attributed to the territorial state for purposes of state responsibility, it will almost certainly be the case that the acts will also be attributable to the state for purposes of self-defense. But if the NSA attacks cannot be attributed to the state, then the territorial state will be responsible only for the violation of the obligation not to allow its territory to be used to the detriment of other states—and that state responsibility cannot be any justification for the use of force against the territorial state.²²³

221. See *Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory*, Advisory Opinion, 2004 I.C.J. Rep. 136, ¶ 139 (July 9). The ICJ interpreted these resolutions in this restrictive fashion. See also, GRAY, *supra* note 21, at 193–94, 199. But see DINSTEIN AGGRESSION, *supra* note 21, at 207; LUBELL, *supra* note 3, at 35.

222. Having said that, one difference between them is suggestive of the incoherence of the attempt to rely upon state responsibility. State responsibility is only incurred upon the actual violation of an international legal obligation, not in anticipation of it. Thus, while the *jus ad bellum* is understood by many to permit an exercise of anticipatory self-defense in response to an imminent armed attack from an NSA if the actions of the NSA can be attributed to the state in which it is operating, no state responsibility will be incurred by that territorial state unless and until an armed attack is actually launched. Even if state responsibility could ever justify the use of force, it certainly could not justify the use of force for future violations. This further underlines the confusion created by the apparent attempts to conflate and selectively employ principles from different regimes in an effort to justify responses to terrorist attacks. My thanks to Vladyslav Lanovoy for bringing this point to my attention.

223. SHAW, *supra* note 135, at 605–06. For the ICJ articulation of the law of state responsibility and countermeasures, see *Gabcikovo-Nagymaros Project (Hung. v. Slov.)*, Judgment, 1997 I.C.J. Rep. 7, ¶ 57 (Sept. 25) (explaining Hungary’s failure to comply with treaty obligations and the unavailability of necessity as a defense).

The foregoing discussion of attribution and the distinction between the law of state responsibility and *jus ad bellum* also points to another problem with the unwilling or unable doctrine, which is the extent to which it suggests that inability alone can constitute a justification for the use of force.²²⁴ The very name of the doctrine suggests that the target state may use force against an NSA whenever the territorial state is *either* unwilling *or* unable to itself address the threat. But this cannot be right. It cannot be the case that a target state is entitled to use force against another state (or within the territory of another state, for those who try to make this distinction)²²⁵ exclusively because that state does not have the ability to respond to a threat posed by an NSA within its territory. This would be to import the low standard of due diligence that does operate in the law of state responsibility into *jus ad bellum*.²²⁶ But the *jus ad bellum* regime does not provide for any such near strict liability. There must be something in addition to the mere inability to prevent harm to justify a use of force. As already explained above, the traditional *jus ad bellum* position requires that the armed attacks (or imminent threat of such attacks) by an NSA be attributable to the state in order to justify the target state exercising a right to use force in self-defense against the territorial state.²²⁷ States cannot use force against the blameless (in *jus ad bellum* terms), and to be merely incapable of dealing with the NSA is no violation of the principles of *jus ad bellum*. As indicated above, it may constitute a violation of other obligations, thus attracting state responsibility, but that cannot justify a use of force. This is another way in which the principles implicitly import state responsibility into what should be strictly considerations of *jus ad bellum*. And to be fair, the Bethlehem Principles recognize this initially, by requiring the target state to present the territorial state with a plan of action, and to seek consent to do what the territorial state is incapable of doing, namely neutralizing the threat posed by the NSA.²²⁸ In the event that the unable territorial state so consents, the target state's actions against the NSA do not constitute a use of force

224. Dire Tladi is similarly critical of this aspect of the Principles. Dire Tladi, *The Nonconsenting Innocent State: The Problem with Bethlehem's Principle 12*, 107 AM. J. INT'L L. 570, 576 (2013).

225. See, e.g., Mahmoud Hmoud, *Are New Principles Really Needed? The Potential of the Established Distinction Between Responsibility for Attacks by Nonstate Actors and the Law of Self-Defense*, 107 AM. J. INT'L L. 576, 577–78 (2013).

226. CRAWFORD, *supra* note 206, at 369 (on strict liability in law of state responsibility).

227. On attribution for purposes of state responsibility, see ILC *Draft Articles*, *supra* note 206, arts. 4–11; see also CRAWFORD, *supra* note 206, at 146–156.

228. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 776.

against the territorial state, and the unwilling or unable doctrine is irrelevant.

If the “unable” territorial state refuses to consent, however, then it thereby becomes “unwilling,” and moves along the spectrum toward “colluding” or “harboring.”²²⁹ In this case, such a nonconsenting territorial state moves closer to being a state to which the actions of the NSA can actually be attributed for purposes of *jus ad bellum*, as in the case of the Taliban regime when it refused to address the threat of al-Qaeda after 9/11.²³⁰ In short, the use of force against the NSA within the territorial state cannot be justified by the mere fact that the state is unable to deal with the threat, but may be justified by the fact that the state is unwilling to permit the threat to be addressed by the target state. This is somewhat analogous to the Responsibility to Protect (R2P) doctrine for humanitarian intervention. Just as states may lose some of their sovereign rights under R2P by reason of their causing or failing to address humanitarian disasters within their territory, here the territorial state may be said to forfeit some of its sovereign rights due to its unwillingness to either take action or to consent to such action to prevent terrorist attacks. In *jus ad bellum* terms, it thereby becomes substantially involved in the acts of the NSA—and only then is a use of force against that state in self-defense justified.²³² In essence then, the “unwilling or unable” doctrine should really be understood as being only an “unwilling” doctrine—or the “unwilling, or unable *and* unwilling.” But even if the Doctrine were so interpreted or recharacterized, there remains the problem that the Bethlehem Principles walk back from the presumption that the target state must present to the territorial state an assessment of the threat, a plan of action, and a request for consent. It is to that problem that we turn next.

E. *Self-Judging Armed Attack, Necessity, and Unwillingness*

The Bethlehem Principles do not provide sufficient guidance as to how and by whom a series of determinations are to be made, as part of the process of deciding on the use of force against NSAs located in

229. *Id.*

230. See *supra* text accompanying note 221.

232. Rosa Brooks makes this connection between R2P and the unwilling or unable doctrine in BROOKS, *supra* note 11, at 245–52. See also Paddeu, *supra* note 45, (explaining that if a use of force is justifiable as self-defense in *jus ad bellum* terms, then self-defense also operates as a separate circumstance precluding wrongfulness under the law of state responsibility to justify or excuse the other violations of international law in relation to the rights and interests of the territorial state that will also likely be caused by the use of force – a point that is often overlooked by advocates of the “use of force within” not “use of force against” argument).

another state. By “whom” here I am questioning whether the doctrine does or should authorize that the decisions be made by the target state alone, or by the target state after consultation with and input from the territorial state. And by “how,” I am questioning whether there are standards, beyond the bare “reasonable and objective basis,” according to which decisions are to be made, and later assessed for legitimacy. Before turning to these issues of “how” and “by whom,” however, it is important to first unpack and identify each of the more significant determinations that have to be made—for as I will argue later, these decisions should not all be subject to the same standards. One of the weaknesses of the Bethlehem Principles is that they aggregate most of these decisions and imply that they are all subject to a single “reasonable and objective basis” standard (adding only “strong” to the one decision of whether there is excessive risk in asking for consent).²³³ The distinct decisions buried in the principles include: (i) that the NSA has launched, or poses a threat of launching, strikes that rise to the level of an armed attack; (ii) that the use of force is necessary to prevent the launching or continuation of attacks, which, in the event of future attacks, rests on a related determination that the threatened attack is imminent; (iii) that the territorial state is unwilling to prevent the attacks—which in turn requires a more specific determination that either (a) the territorial state is both unable to address the threat itself and is unwilling to provide consent for the target state to take the necessary action, or (b) that the territorial state, while able to do so, is unwilling to prevent the attacks or consent to the target state doing so, and is thereby actually “harboring” or “colluding with” the NSA.²³⁴ What is more, the “unwillingness” determination needs to be unpacked further: the principles contemplate determinations of unwillingness after a request for action or consent has been made and denied, and unilateral determinations of unwillingness or consent by way of inference.²³⁵ Each of these may require different standards. As will be discussed in more detail in Part IV below, a far more stringent standard should be required to justify an inference of unwillingness than for determining that a request for consent has been unreasonably denied.

The inferences of “unwillingness” or “consent” also involve some prior judgments. That is, the principles suggest that a prior determination must be made by the target state as to whether it is required to explicitly request that the territorial state take action or

233. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 775–76. The “strong, reasonable, and objective basis” standard is also used in Principle 7 as the basis for targeting those providing material support for armed attacks. *Id.*

234. See *supra* Part II.D.

235. See Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 775–76.

consent to such action in the circumstances, or whether it is permitted to forgo such requests and unilaterally make inferences as to the unwillingness of the territorial state.²³⁶ The inference of unwillingness thus theoretically involves two distinct determinations—namely the determination by the state that it is entitled in the circumstances to make such an inference, and then the drawing of the substantive inference itself. But in practical terms these two distinct determinations begin to collapse into one. Both decisions are based on the same evidence and conclusions. The evidence that forms a sufficient basis for the decision that the territorial state is unwilling to either take action or consent to action is going to also form the basis for the decision that the target state is entitled to make that very inference.²³⁷ The same is true of the inference of consent. Thus, the question will be what standard should apply to the combined determination that the territorial state is unwilling to assist or consent, and that such a unilateral inference is permitted in the circumstances. But that standard arguably should be different from the standard for determining whether the target state can reasonably determine that the territorial state is unwilling, based on its explicit denial of a request for action or consent to permit a use of force against the NSA. For short, I will refer to the categories of substantive decision points as (i) “armed attack,” (ii) “necessity/imminence,” (iii) “explicit unwillingness,” and (iv) “inference of unwillingness/consent.”

A problem with the Bethlehem Principles is that they do not sufficiently consider the differences between these distinct decisions, and thus purport to subject them all to a similar standard. Part IV will take up the issue of how we might think about developing different standards to govern these distinct decisions, but it is worth pausing here to consider why varied standards are theoretically necessary. To begin with the categories of “armed attack” and “necessity/imminence,” these are already part of the doctrine of self-defense, and yet they are modified in the context of the unwilling or unable doctrine.²³⁸ The *jus ad bellum* regime implicitly contemplates that states are entitled to determine for themselves that they are the victim of an actual or imminent armed attack, and that the use of force in response is necessary.²³⁹ The presumption is that states may and will make these

236. *Id.* at 776 (explaining in what circumstances the requirement for consent is inapplicable).

237. *Id.*

238. See *supra* Part II.A.

239. This can be traced back to the exchange of diplomatic notes during the negotiation of the Kellogg-Briand Pact. See FORCESE, *supra* note 65, at 185–86, and

decisions unilaterally. Such decisions are, however, typically subject to *ex post* assessment of one kind or another, and the defending state usually feels compelled to furnish the international community with arguments and evidence in justification of its actions.²⁴⁰ The determination of these very same issues is slightly different, however, in the context of the unwilling or unable doctrine, because there is the explicit presumption, provided for in Principle Ten, that the target state should first request that the territorial state either take action or consent to the target state doing so.²⁴¹ Notwithstanding how the principles walk back from this presumption, it should be understood to alter the nature of the standards that govern the decision-making. For it needs to be said yet again that the presumption reflects the fact that the territorial state may be “innocent” in the sense of having done nothing to encourage or assist the NSA action—and so the evidentiary burden for establishing that it is in some way culpable should be high for justifying any violation of its sovereign rights for the purpose of attacking the NSA without the state’s consent. It certainly cannot be the case, therefore, that the standards governing these decisions would be lower than those that apply to a defending state in a straightforward exercise of self-defense against an attack by another state.

Turning to the third category, of “explicit unwillingness,” recall again that Principle Ten establishes a baseline presumption that the target state must obtain consent for the use of force from the territorial state, and that “all reasonable good faith efforts be made to obtain” such consent.²⁴² Principle Twelve, the “unable” Principle, even provides that if consent is requested, the target state must provide to the territorial state a reasonable plan of action to deal with the NSA, and seek its consent to the execution of the plan.²⁴³ But Principles Eleven, Twelve, and Thirteen then go on to significantly hollow out that presumption. What happens if the target state does request consent but the territorial state refuses? The principles do not actually contemplate this circumstance explicitly, but they would seem to imply that such refusal would by definition make the territorial state “unwilling,” thereby “colluding” with or “harboring” the NSA. But this too needs to be unpacked. What if the territorial state does not agree with the target state’s assessment that the NSA poses a risk of

nn.40–42. For the notes exchanged, see DENYS P. MYERS, ORIGIN AND CONCLUSION OF THE PARIS PACT 34–56 (1929) (providing a history of the negotiations); see also U.S. DEP’T OF STATE, TREATY FOR THE RENUNCIATION OF WAR (1933) (containing reproductions of all the associated documents, including the US note of June 23, 1928).

240. Consider the examples of Colin Powell addressing the Security Council on the eve of the invasion of Iraq in 2003. Colin Powell, U.S. Sec’y of State, Address to the U.N. Security Council (Feb. 5, 2003), <https://www.theguardian.com/world/2003/feb/05/iraq.usa> [<https://perma.cc/ETD9-96WZ>] (archived Feb. 17, 2019).

241. *Id.*

242. *Id.* at 776.

243. *Id.* at 775–76.

imminent armed attack? Or that the use of force is necessary to prevent such attack? Indeed, what evidence is the target state required to provide to the territorial state in the process of requesting action or consent? Which judgment should prevail in making the determinations of armed attack, imminence, and thus necessity? This raises both “how” and “by whom” issues. If the target state has complete authority to decide to act regardless of the territorial state’s position on these issues, then the request for consent is not a consultation in any meaningful sense, but merely an ultimatum: “take action or consent immediately, or we will undertake strikes in your territory regardless.”

It is worth recalling that the Taliban requested, as a precondition for turning over Osama Bin Laden, evidence from the United States to support its claim that al-Qaeda was responsible for the 9/11 attacks, and that the United States refused to provide any such evidence.²⁴⁴ The fact that it is now almost universally accepted that al-Qaeda *was* responsible for the 9/11 attack, and that the 9/11 attacks did constitute an armed attack on the United States, does not mean that the request for evidence was unjustified, or that the United States had no obligation to provide such evidence before using force against Afghanistan. Yet while the Bethlehem Principles again do not address these questions directly, it seems quite clear that the doctrine is understood to permit the target state to make most if not all of these determinations for itself unilaterally, each one governed by the same relatively low reasonableness standard.²⁴⁵ There seems to be little recognition that territorial states may have a reasonable basis for protesting that a use of force is not necessary or justified in the circumstances. Moreover, there is no contemplation of any *ex post* assessment of which view was the more reasonable view in the circumstances, if the target state makes its own determination and acts upon it over the objections of the territorial state.

This brings us to the final category, the “inference of unwillingness/consent.” Recall that Principle Twelve provides that consent to the use of force by the target state is not required where the territorial state is unable to itself address the threat, and there is a “strong, reasonable, and objective basis” for concluding that

244. See, e.g., *Bush rejects Taliban offer to hand Bin Laden over*, THE GUARDIAN, (Oct. 14, 2001), <https://www.theguardian.com/world/2001/oct/14/afghanistan.terrorism5> [<https://perma.cc/24QJ-EFBJ>] (archived Mar. 13, 2019); President George W. Bush, Address to a Joint Session of Congress and the American People (Sept. 20, 2001), <https://georgewbush-whitehouse.archives.gov/news/releases/2001/09/20010920-8.html> [<https://perma.cc/VG3U-H8JW>] (archived Feb. 17, 2019) [hereinafter Bush]

245. See generally *Bethlehem Self-Defense Against NSAs*, *supra* note 16.

attempting to seek consent would materially increase the risk that the use of force will be ineffective (primarily due to suspicion that the territorial state will tip off the NSA).²⁴⁶ This suggests that the target state is entitled to both infer that the territorial state is “unable” and to determine that requesting consent poses too great a risk. As discussed above, the idea that this alone would permit the use of force is entirely inconsistent with the traditional understanding of *jus ad bellum*,²⁴⁷ and so to interpret it in a manner that is consistent with established law, the Principle should also be understood to mean that the target state has the basis for making the further inference that the unable state is also unwilling to permit the target state to take action. As we have seen, Principle Eleven provides for just such an inference, stating that the target state need not obtain consent where there is a “reasonable and objective basis” for concluding that the territorial state is either “harboring” or “colluding with” the NSA, and is thus unwilling to either take action or provide consent to the target state’s action (though, again, the Principle does not explicitly indicate who is to so conclude).²⁴⁸ Finally, Principle Thirteen adds that the territorial state’s consent to a target state use of force within its borders may be implicit, being inferred from past acquiescence.²⁴⁹ And all of these decisions are permitted without any communication with the territorial state regarding its ability to act, its willingness to act, or its willingness to consent to the target state’s action. As with other aspects of the principles, the grant of unilateral discretion to make these decisions overly privileges the target state, allowing it to reduce its risk to the greatest extent possible at the cost of the rights and interests of the territorial state, and at the cost of distorting the principles that govern the doctrine of self-defense.

These unilateral decisions or inferences should surely be understood to be governed by stricter standards than the standard that applies to self-judgment in the traditional doctrine of self-defense. What is more, in contrast to the operation of the traditional doctrine of self-defense, in which there is typically some form of explanation and disclosure as part of a process of justification, the unwilling or unable doctrine is typically invoked in circumstances in which target states make virtually no disclosure. The United States has refused to disclose the legal analysis justifying its targeted killing program in general, and in the specific instances in which it has invoked the unwilling or unable doctrine, it has done so with bald assertions of the imminence

246. *Id.*

247. *See supra* text accompanying notes 224–231.

248. *Bethlehem Self-Defense Against NSAs*, *supra* note 16, at 775.

249. *See supra* Part II.D.

of attack and necessity of self-defense, with no supporting evidence.²⁵⁰ The greatest disclosure has been the legal analysis justifying the killing of Anwar al-Awlaki,²⁵¹ but it was both reluctant and was the exception, with both the analysis and the reasons for its disclosure being driven primarily by the fact that al-Awlaki was an American national.²⁵² This lack of transparency regarding the decision-making suggests that there should be additional standards regarding levels of disclosure and *ex post* justification, as well as an understanding that the less transparency there is, the stricter the standards governing the decision-making itself should be.

The failure to more fully address how and by whom all of these distinct determinations are to be made, and the strong bias in favor of the interests of powerful target states over the rights of weak territorial states, is a structural weakness that has already manifested itself in uses of force that are very difficult to reconcile with the traditional doctrine of self-defense.²⁵³ As Deeks wrote before the publication of the Bethlehem Principles, “where the test is not clear, a victim state’s claim that a territorial state is unwilling or unable to act is easy to make, relatively hard to disprove, and at least superficially useful in concealing an incursion based on other motivations.”²⁵⁴ The Bethlehem Principles have not added sufficient clarity to resolve this problem. Notwithstanding the recognition that there should be a presumption of consultation and request for consent, the principles entirely undermine this presumption. They tend to lump all the determinations together, entirely obscuring some of them, treating all of them as being within the sole preserve of target state self-judgment, and suggesting that all but one of them is subject to the low standard of “objective and reasonable basis.”²⁵⁵ The principles provide no guidance as to how that standard should be actually applied, or to what extent it should be susceptible to *ex post* review.

250. See BROOKS, *supra* note 11, at 104–28 (discussing the lack of transparency and secrecy of the policy).

251. This attack was arguably undertaken with Yemini consent, or at least without objection, and so did not require invocation of the unwilling or unable doctrine—yet the analysis did nonetheless rely on self-defense arguments. President Ali Saleh, who had resisted US air operations in Yemen, had been the victim of rocket attack in June, 2011. Al-Awlaki was killed in the months that followed, during which there was something of a vacuum in Yemini government leadership. See MAZZETTI KNIFE, *supra* note 156, at 305–08.

252. See, e.g., MAZZETTI KNIFE, *supra* note 156, at 299–320; SCAHILL, *supra* note 156, at chs. 44, 56, 57.

253. See RUYSS, *supra* note 3, at 10.

254. Deeks *Unwilling*, *supra* note 7, at 507.

255. Bethlehem *Self-Defense Against NSAs*, *supra* note 16.

IV. REFINING THE DOCTRINE

How then should the doctrine be refined? There have been various suggestions in response to the publication of the Bethlehem principles. Some scholars have written that communication between the target and territorial states is necessary, and that the target state must be ready to provide *ex post* any evidence supporting its determination that the territorial state was indeed unwilling or unable.²⁵⁶ Ashley Deeks elaborated a new factor-based test to assist in these determinations as part of moving to a clearer rule-based system, which still has much to commend it as a basis for improving the Bethlehem Principles.²⁵⁷ Others have suggested that some independent international organizations should be involved in the determination *ex ante*.²⁵⁸ But the target states that have most commonly relied upon the doctrine—namely the United States, the United Kingdom, and Australia in recent years—are highly unlikely to accept any delegation of the decision-making process to an international body *ex ante*, and will be quite resistant to most of the other proposals.²⁵⁹ What is more, most of the proposals focus on the important but narrow issue of how and by whom the determinations get made, with much less attention on addressing the inconsistency with the *jus ad bellum* regime. But, as I have argued above, the principles not only excessively privilege and favor the interests of the target state over the rights and interests of the territorial state in the manner in which these determinations are dealt with, but aspects of the principles are inconsistent with *jus ad bellum* and its relationship with other regimes. All of these problems taken together serve to dangerously undermine the prohibition on the use of force more generally, and thus all need to be addressed. They need to be revised or refined in order to bring them into closer compliance with the established understanding of *jus ad bellum* and its relationship with both IHL and the law of state responsibility, and to establish clearer standards on how and by whom the relevant decisions are made. This Part is devoted to exploring some suggestions on how this might be done.

A. Restoring the Integrity of *Jus ad Bellum*

The initial step is to revise those aspects of the Bethlehem Principles that are inconsistent with the *jus ad bellum* regime and its

256. See, e.g., Irene Couzigou, *The Right to Self-Defence against Non-State Actors: Criteria of the "Unwilling or Unable" Test*, 1 HEIDELBERG J. INT'L L. 47, 47–49 (2017).

257. Deeks *Unwilling*, *supra* note 7, at 519.

258. Ahmed, *supra* note 14, at 36.

259. See, e.g., Wright *Self-defense*, *supra* note 57; Moorehead, *supra* note 128.

relationship with other bodies of international law. As explained in Part III, while these inconsistencies pose serious threats to the *jus ad bellum* regime and the integrity of its relationships with other regimes, these aspects of the principles yet are not central to the operation of the doctrine, so could be eliminated or revised without compromising the doctrine. Thus, a first step would be to reject entirely those parts of the Bethlehem Principles that interfere with or purport to trump the rules of IHL.²⁶⁰ This can be solved by the deletion or revision of those clauses in Principles Six and Seven that purport to establish separate and independent authority for the targeting of individuals who are planning, threatening, or providing support essential to a terrorist attack.²⁶¹ It must be recognized that even in the event that an international armed conflict is not already in existence between the target state and territorial state, the very act of using force in self-defense triggers the operation of IHL, and thus IHL governs the conduct of armed forces in all aspects of targeting.²⁶² The principles should be revised to clarify that the doctrine does not purport to interfere in any way with IHL rules. Similarly, the references to the law of state responsibility should be deleted, as they run the risk of improperly suggesting that state responsibility may contribute to the justifications for the use of force in some circumstances.²⁶³ The essential question should be whether the unwillingness of the territorial state rises to the level of collusion with, control over, or substantial involvement with the NSA, so as to justify the use of force in self-defense against the state for the actions of the NSA. The principles should make clear that this is exclusively a *jus ad bellum* question, not a state responsibility issue.²⁶⁴

The meaning of the concept of imminence also needs to be revised. First, the current formulation needs to be adjusted so as not to undermine the integrity of the principle of necessity in the doctrine of self-defense.²⁶⁵ But at the same time, the revision could take seriously the concerns that the Bethlehem Principles and target state policy seek to address. The doctrine could be revised to adopt some version of the

260. See *supra* Part III.D.

261. *Id.*

262. DINSTEIN AGGRESSION, *supra* note 21, at 156–68; NEFF, *supra* note 23, at 340–46, 366–69. IHL would apply regardless of whether one argued the strikes were part of an international or a non-international armed conflict. Indeed, if there is no armed conflict and IHL is not operating, then the target state cannot make recourse to IHL for legal authority to engage in lethal operations in the territorial state at all, and any such operations would be governed by the stricter standards of domestic criminal law and international human rights law. MELZER, TARGETED KILLING, *supra* note 3, at 80–81.

263. See *supra* Part III.E.

264. See *supra* Part III.E.

265. See *supra* Part III.B.

true “last clear chance” formulation, such that target states would not always be required to have proof that an attack was immediately pending before being justified in responding with force. But this revised formulation would still have to include a temporal concept tied directly to the principle of necessity. And it would not overly privilege target states at the expense of the rights of territorial states or operate in ways that would likely lead to unnecessary uses of force. It should require compelling evidence that: (i) the NSA in question had decided upon and was capable of implementing a specific armed attack, such that there was considerable certainty that the attack would be executed; (ii) notwithstanding that the attack was not immediately pending, the immediate opportunity to use force that presented itself was truly the last and only chance to prevent the threat from materializing; and therefore (iii) there were no other alternatives remaining to prevent the attack, and the immediate use of force was necessary. That may seem like a high bar, but it does take seriously the concerns reflected in the Bethlehem Principles, yet not at the expense of the essential temporal meaning of the concept of imminence, or its integral relationship to the principle of necessity.²⁶⁶

The problems surrounding the definition and use of the concept of “armed attack” should also be resolved. First, the issue of treating a series of small-scale strikes as together constituting an armed attack should be addressed. Some may argue that the current ICJ position on the scale and intensity of force required to constitute an “armed attack” is too high, and needs to be modified, particularly given the kinds of threats posed by transnational terrorist organizations.²⁶⁷ But such arguments either disregard the fact that the modified concept of armed attack will apply in all contexts of self-defense, or they assume without any explanation that there should be a different and unique definition for the concept when applied to the NSA context.²⁶⁸ If there is to be such a unique, context-specific definition, the rationale and parameters of the concept should be clearly developed. In the absence of any such bifurcation of the concept, any modification made for purposes of addressing NSA threats will affect the operation of the doctrine of self-defense in all other circumstances. And permitting the use of force generally against another state in response to a series of sporadic, low-scale, and low-intensity uses of force, such as border skirmishes, could dangerously lower the threshold for war.²⁶⁹ As examined above, the *jus ad bellum* regime does not yet recognize the

266. See *supra* Part III.B.

267. See, e.g., DINSTEIN AGGRESSION, *supra* note 21, at 210.

268. There is currently no such distinction. RUY, *supra* note 3, at 143–74.

269. See GRAY, *supra* note 21, at 155–57 (explaining which mix of events have justified the use of force by the ICJ); RUY, *supra* note 3, at 532–35.

idea that a series of small-scale and low-intensity uses of force can in the aggregate constitute an armed attack justifying the exercise of self-defense,²⁷⁰ and thus, in my view, the doctrine should be revised to comply with the established law.

As to the second problem in relation to the concept of armed attack, the unwilling or unable doctrine should not employ the concept in ways that are internally incoherent. As explained above, one cannot define armed attack as an accumulation of strikes that do not individually rise to the level of armed attack, yet support the use of force against each such strike as an act of self-defense against an independent imminent armed attack.²⁷¹ For the doctrine to have a chance of being accepted widely, champions of the doctrine must develop rationales that are internally coherent and consistent with the *jus ad bellum* regime more generally, and its relationship with IHL and other regimes. And again, the operation of the doctrine need not depend upon such incoherent positions. Under existing IHL, for instance, once a target state has used force against an NSA within an unwilling state, it thereby triggers an international armed conflict with that state.²⁷² If the NSA is preparing to launch further attacks, the situation is akin to the continuation of hostilities in a more traditional international armed conflict. As such, target state responses to those strikes would be a continuation of the act of self-defense, and the *jus ad bellum* regime would require an assessment of necessity and proportionality in deciding whether and how to respond to such pending strikes, but it would not require establishing anew that the pending action by the NSA poses an “imminent armed attack” justifying a use of force in self-defense. There is no need to justify each use of force against NSA action as an independent act of self-defense. And the threshold for determining the unwillingness of the state (a topic addressed in more detail below) might be lower once it has already been established to have been unwilling in the context of the initial attacks.

It is true that so revising the principles to bring the treatment of armed attack back into compliance with the established doctrine of self-defense, and in accord with the relationship between *jus ad bellum* and IHL, might disqualify some of the target state actions that have been previously justified under the doctrine. But that merely highlights the extent to which those strikes were likely illegitimate. Revising the principles in this way would still allow the doctrine to

270. See sources cited *supra* notes 182–186 and accompanying text.

271. See *supra* Part III.C.

272. Prosecutor v. Tadić, Case No. IT-94-1-A, Judgment, ¶ 84 (Int'l Crim. Trib. for the Former Yugoslavia July 15, 1999).

operate, albeit within narrower parameters, but also with greater legitimacy and wider acceptance.

B. *Differentiated Standards for Distinct Target State Decisions*

As explained in Part III, the aspect of the principles relating to how and by whom each of the relevant determinations are to be made is seriously underdeveloped, and thus revision requires elaboration as well as refinement. Recall that these decisions were grouped into four broad categories in Part III, namely: (i) armed attack; (ii) necessity/imminence; (iii) explicit unwillingness; and (iv) an inference of unwillingness/consent.²⁷³ The principles do not sufficiently provide for differentiated standards to govern the decision-making process underlying each of these determinations, in terms of the degree of consultation and disclosure that is required, and on what criteria and what levels of evidence the determination of unwillingness in particular should be made. So how should they be refined?

It was noted in Part III that the determinations as to “armed attack” and “necessity/imminence” are already part of the established doctrine of self-defense and would be at a minimum subject to the same standards as applies in the broader context. As will be explored in more detail below, this includes being subject to a good faith standard, in that these decisions must be made honestly and on an objective and reasonable basis—yes, the same language used in the Bethlehem Principles. But, as I have argued above, in contrast to the regular state-on-state exercise of self-defense, in the context of the unwilling or unable doctrine the determinations of “armed attack” and “necessity/imminence” are not alone sufficient to justify the use of force—there must also be the determination of territorial state “unwillingness” in order to justify the use of force in the absence of consent.²⁷⁴ The determinations of unwillingness ought to be subject to different and stricter standards.

Beginning with the issue of explicit unwillingness, recall that Principle Ten provides for a presumption that the target state must request assistance or consent prior to using force.²⁷⁵ There are two aspects to this presumption that might impact how we think about the standards applicable to the determination of unwillingness.²⁷⁶ First, in the scenario in which the target state does request assistance or consent and the territorial state refuses, there have to be (i) standards regarding the sufficiency of evidence that is provided by the target

273. See *supra* Part III.E.

274. See sources cited *supra* notes 224–231 and accompanying text.

275. See *supra* Part II.D.

276. For a useful discussion of presumptions, albeit in the context of evidentiary presumptions in adjudication, see FRÉDÉRIC GILLES SOURGENS, KABIR DUGGAL & IAN A. LAIRD, *EVIDENCE IN INTERNATIONAL INVESTMENT ARBITRATIONS*, ch. 6 (2018).

state as to the armed attack and necessity for the use of force (disclosure standards), and (ii) standards for governing the basis upon which the target state can disregard or effectively overrule the territorial state's objection or denial (consultation standards). With regards to the first, the target state must be required to provide sufficiently compelling evidence to demonstrate that (i) its determinations as to the threat of armed attack and necessity for the use of force were made honestly, on a reasonable and objective basis, and (ii) that there is a high probability that the conclusion is accurate—not merely likely. The target state may be reluctant to share information that could reveal intelligence-gathering sources or methods,²⁷⁷ but on the other side of the equation is the territorial state's sovereign right to be free from the use of force and foreign intervention. It does not seem a reasonable balance that the more powerful target state should be entitled to violate those sovereign rights, often killing innocent civilians among the population of the territorial state,²⁷⁸ merely to minimize the risk of compromising its own intelligence collection methods. The obligation to provide sufficiently compelling evidence to ground the honesty and reasonableness of the target state's determinations, and to demonstrate that there is a high probability that such determinations are accurate, should be made an explicit standard for the operation of the doctrine where assistance or consent is requested.

The second aspect of the presumption to obtain consent, the consultation standards, governs the extent to which the target state has a duty to consider the territorial state's views. The principles must consider what happens if the target state makes the request and provides sufficient evidence, but the territorial state interprets the evidence differently and reaches a different conclusion. As discussed earlier, the target state should not be entitled to entirely ignore the territorial state's input, but neither would it be reasonable for the territorial state to have a "veto" over target state action.²⁷⁹ The standard should reflect a position that is somewhere in between the two extremes. Thus, the target state should be required to provide evidence and afford an opportunity to respond, following which the target state must in good faith consider the territorial state's response.

277. BROOKS, *supra* note 11, at 124–28.

278. There is considerable debate over this issue. See, e.g., *Drone Warfare*, THE BUREAU OF INVESTIGATIVE JOURNALISM, <https://www.thebureauinvestigates.com/projects/drone-war> (last visited Feb. 6, 2019) [<https://perma.cc/4ZW5-J32F>] (archived Feb. 17, 2019).

279. See *supra* Part III.E.

For instance, if the territorial state itself has evidence that may cast a different light on how the activity of the NSA should be understood, it should be entitled to have that evidence reviewed and factored into the analysis, and the target state should be obligated to so consider it. There are far too many known examples of the United States mistakenly killing and torturing the wrong people since 9/11 for there to be any argument that target states always get the analysis right, or would not benefit from input from territorial states.²⁸⁰ Finally, this middle position also means that the target state must be prepared to demonstrate, after the fact, that it did indeed consult with the territorial state, and did consider its position and evidence, before final decisions were made.

Next, we come to the inferences of unwillingness and consent, which is an alternative to the target state explicitly seeking consent or assistance. The treatment of these inferences is among the most serious weaknesses of the Bethlehem Principles, and requires the greatest revision. I would argue that since there is no explicit basis in treaty or otherwise for self-judgment in this regard, and that the inference is being relied upon to justify what would otherwise be a violation of a *jus cogen* norm, there should at minimum be a presumption against such unilateral inferences. Such a presumption should be rebuttable only by satisfying significantly more stringent standards than the other unilateral determinations. The idea that the target state can decide on a mere “reasonable and objective basis” that it is entitled to simply infer that the territorial state is unwilling, or that it has implicitly consented to a use of force, makes a mockery of the presumption that the target state must obtain consent. This is particularly so when one considers that in practice, target states relying on the doctrine have typically failed to provide any disclosure *ex post* regarding the evidence and analysis upon which these decisions have been made.²⁸¹ If states are entitled to make these inferences, and never have to explain the basis upon which they were made, it is naïve in the extreme to expect that such decisions will always be made honestly and on valid and persuasive evidence. It is an invitation for action on the basis of bare supposition, suspicion, and speculation, and indeed for use of the doctrine as a pretext for ulterior motives.²⁸² What then should be the standard?

280. On errors committed in the CIA torture program, see S. REP. NO. 113-288 [Senate Select Committee on Intelligence, Study of the Central Intelligence Agency's Detention and Interrogation Program], at 14–17, 133 (2014).

281. THE DRONE MEMOS: TARGETED KILLING, SECRECY, AND THE LAW 23–35 (2016); see also Alston *Study*, *supra* note 3, at Part III (“Policy continues to be characterized primarily by secrecy and opacity, combined with occasional leaking of figures that are not explained in any detail and that generate more questions than they answer.”).

282. Deeks noted this in her earlier work, Deeks *Unwilling*, *supra* note 7, at 28.

In order for the target state to depart from the presumption that it must obtain consent, and instead draw any one of the three inferences—namely: that it is too risky to request consent; that consent has been implicitly provided; or that the territorial state is unwilling—the target state should be required to have sufficiently compelling evidence to support the conclusion that the territorial state is implicitly consenting, or that it is unwilling to consent, with a high probability of correctness. And this should be on the basis of the perspective of an objective reasonable observer. Moreover, not only should the possession of such evidence be a condition of the state acting on such an inference, but the target state should be required to disclose such evidence after using force against the NSA, both to the territorial state and to relevant international institutions. The standard should not be merely objective and reasonable evidence that the conclusion be more probable than not, and it should not permit the target state to merely assert conclusions and avoid *ex post* disclosure of the evidence. The same standard should apply to the inference that the territorial state has implicitly consented, whether the evidence be of explicit consent in the past in similar circumstances (not mere silence in the wake of past uses of force), or documentary evidence of informal back-channel consent in the current situation. It will be objected that such informal back-channel consent cannot be subsequently disclosed without embarrassing the territorial state government (or one of its agencies), thereby undermining the very idea of informal or implicit consent. But more harm is done to the international rule of law, and to the integrity of the *jus ad bellum* regime in particular, by instances of target states using force against territorial states on the basis of asserted implicit consent, but in the face of explicit and loud objections and protests by the territorial state.

A further distinction has to be made with respect to the inference of unwillingness. It will be recalled that in Principle Eleven of the Bethlehem Principles, a distinction is made between unwillingness that can be established because the territorial state is “colluding” with the NSA, and other situations in which the territorial state is “otherwise unwilling to effectively restrain the activities of the NSA,” which is referred to as “harboring.”²⁸³ The principles are not explicit on whether both of these are subject to the unilateral “inference of unwillingness.” Is it open to the target state to infer both? Is one more likely to be the result of an inference? Do they imply the same level of culpability? It is not at all clear how the principles respond to such questions. But I would argue that our understanding of the unilateral inference of “unwillingness” should be informed by the more traditional

283. See *supra* Part II.D, specifically text accompanying notes 108–111.

test for attribution in the doctrine of self-defense. It will be recalled that the traditional test, at least as articulated by the ICJ, is that there be sufficient evidence that the territorial state is or has been substantially involved in the activities of the NSA.²⁸⁴ The ICJ ruled out the notion that the mere “harboring” of armed groups constituted a sufficient nexus to support attribution to the state of the acts of the armed groups.²⁸⁵

It is true that the *Nicaragua* standard has been increasingly challenged in the post-9/11 era. As discussed above, the Taliban’s acquiescence in al-Qaeda’s actions and subsequent refusal to address the threat was seen as being sufficient “involvement” in and *ex post* “adoption” of al-Qaeda’s actions to justify the use of force against Afghanistan.²⁸⁶ In the age of increased threats from more sophisticated transnational terrorist organizations, there are certainly good reasons to think that the traditional attribution standard in the doctrine of self-defense may need to be modified. But even so, the underlying rationale, namely that the territorial state must have sufficient involvement to justify a use of force against it, does and should continue to govern the test for attribution—and the same rationale should inform the standards for when a target state may infer that a territorial state is “unwilling” to address the threat posed by an NSA. Looked at through the lens of the traditional test for attribution, an inference of “collusion” is not so far removed from a determination of “substantial involvement.” If there is sufficiently compelling evidence to support the conclusion, with a high probability of accuracy, that the territorial state is “colluding” with an NSA that is launching armed attacks, that would also likely satisfy the traditional test for “substantial involvement,” such that the actions of the NSA could be attributed to the state for purposes of self-defense.

In contrast to collusion, however, the Bethlehem Principles’ vague definition of “harboring” is far removed from the traditional test for attribution, and it could even cover situations that involve less culpability than the “harboring” that was rejected by the ICJ.²⁸⁷ If there is to be some reconciling of the standards for “unwillingness” and the test for “attribution,” there should be more focus on the *unreasonableness* of the territorial state’s unwillingness to either deal

284. *Military and Paramilitary Activities in and Against Nicaragua* (Nicar. v. U.S.), Judgment, 1986 I.C.J. Rep. 14, ¶¶ 194–95 (June 27); *Case Concerning Armed Activities on the Territory of the Congo* (Dem. Rep. Congo v. Uganda), Judgment, 2005 I.C.J. Rep. 1, ¶¶ 106–47 (Dec. 19) (explaining sufficient evidence was not presented by Uganda to constitute acting in self-defense against the DRC); GRAY, *supra* note 21, at 139–145, 226.

285. *Nicaragua*, 1986 I.C.J. Rep. at ¶¶ 194–95; see RUYS, *supra* note 3, at 528–29.

286. Letter from John Negroponte, *supra* note 220 (explaining the United States’ plan to exercise self-defense against al-Qaeda); see also GRAY, *supra* note 21, at 193–94, 199.

287. See generally *Bethlehem Self-Defense Against NSAs*, *supra* note 16.

with the threat itself or provide consent to the target state to do so. Such unwillingness, in the face of sufficient evidence of the necessity for action to defend against the threat, starts to move that state down the spectrum toward the standards for traditional attribution—that is, substantial involvement in the operations of the NSA. It may not be “involved” in operational terms, but unwillingness to act in the face of compelling evidence of the threat begins to implicate and involve the territorial state in the NSA’s activity in moral terms, and the concepts of fault and blame begin to attach. But to permit the target state to merely infer that the territorial state is “otherwise unwilling to effectively restrain the activities of the NSA,” without having first provided it with evidence of the threat or made any direct request for its action or consent, is to radically depart from the idea that there be some level of involvement and culpability sufficient to justify the use of force. I would argue that target states should only be permitted to make inferences of “unwillingness” where there is sufficiently compelling objective evidence to support the conclusion, with a high probability of accuracy, that the territorial state is “colluding” with or substantially involved in the operations of an NSA. In the absence of such evidence, the request for consent must be made and rejected in order to ground the determination of unwillingness.

Finally, given this requirement that there be some level of involvement and thus culpability, which is qualitatively and morally different from a mere inability to prevent the operations of the NSA, the doctrine should be revised not only to reflect this distinction on unwillingness, but also on the issue of inability. As explained in Part III above, the doctrine cannot be understood to justify the use of force against a territorial state based solely on its *inability* to prevent NSA attacks. Granted, from the target state’s perspective, it may make little difference whether the territorial state is “unwilling” or “unable,” so long as the armed attacks continue to rain down upon its head. But from the perspective of whether the territorial state has done something to forfeit its right to full respect for its sovereignty, the difference is salient.²⁸⁸ Thus, mere inability to deal with the threat posed by an NSA should not be grounds for a use of force against a state. The onus is on the target state to approach the unable territorial state for consent to take action. It is only when the state that is unable to deal with the threat posed by an NSA also refuses its consent to allow the target state to do so that it becomes “unwilling,” such that that the use of force may be justified against the NSA and the territorial state in which it is operating as an exercise of self-defense.

288. See *supra* text accompanying notes 229–231.

In essence then, the “unwilling or unable” doctrine should be more properly understood as the “unwilling” doctrine.

C. *Minimum Good Faith Standard for All Self-Judgment*

In addition to the foregoing argument that there should be some differentiation in the standards that apply to the different determinations that the doctrine authorizes the target state to make, there is also an argument that all of these determinations should be additionally subject to a robust and well-developed good faith standard. The Bethlehem Principles themselves suggest the application of a good faith standard—Principle Ten ends with the admonition that “[w]here consent is required, all reasonable good faith efforts must be made to obtain consent.”²⁸⁹ Strangely, it only explicitly imposes a good faith requirement when consent is actually requested, and not when unilateral inferences are being made by the target state. But the language “reasonable and objective basis” applied to those determinations does also suggest an element of good faith. Yet the principles do not go further to explain the nature of the good faith standard that should govern all these unilateral decisions, particularly those that depart from the presumption of requesting either assistance or consent. In suggesting that these determinations all should be subject to a good faith standard, it is important to unpack that ubiquitous concept and establish what precisely it should mean in this context. In thinking about what that should entail, it is helpful to examine other areas of international law in which the exclusive exercise of self-judgment is understood to be subject to a robust and well-developed standard of good faith. Such good faith requirements operate to both limit the discretion of the state engaged in self-judgment and provide a standard for an *ex post* assessment of the validity of the judgment made.

One area in which this has been examined in some detail is that of investor-state arbitration under the terms of bilateral investment treaties (BITs). Many BITs include provisions that allow the state to depart from its obligations for reasons of national security, maintenance of public order, or to respond to some other emergency.²⁹⁰ Such provisions are also common in the many bilateral friendship, commerce, and navigation (FCN) treaties concluded in the period just

289. Bethlehem *Self-Defense Against NSAs*, *supra* note 16, at 776.

290. See generally, e.g., Jürgen Kurtz, *Adjudging the Exceptional at International Law: Security, Public Order and Financial Crisis* (Jean Monnet Prog. Working Paper No. 06/08, 2008).

after World War II.²⁹¹ They are commonly referred to as “non-precluded measures” provisions (NPMs). Many such NPMs include language indicating that the state is not precluded from taking measures (which would otherwise violate the treaty) that it “considers necessary” in order to protect its national security and other vital interests.²⁹² This language is intended, so the argument goes, to signal that such determinations are the exclusive preserve of the state, and not subject to challenge by the investor or the investor’s home state (in the case of BITs), nor subject to judicial review.²⁹³ This immunity from second guessing and judicial review is argued to flow from the fact that this right to exercise self-judgment is explicitly provided for in the treaty, and is thus quite distinct from customary international law defenses under the law of state responsibility.²⁹⁴ In this sense, the exercise of self-judgment pursuant to NPMs in BITs and FCNs is afforded the greatest deference, permitting states to unilaterally make decisions without condition or oversight.²⁹⁵ And yet, even here, it has been argued that the exercise of such judgment is bounded by, and subject to institutional review on the basis of, the requirements of good faith.²⁹⁶ What is more, while no single test for assessing good faith has been generally accepted,²⁹⁷ there are two minimum requirements that have been widely recognized.²⁹⁸

This minimum two-part test for good faith requires that (i) the state engage in honest and fair dealing in the exercise of self-judgment; and (ii) the state have a rational basis for the judgment that is made. Beginning with the first prong of the test, the idea that states must in

291. Such FCN treaties were the basis for the cases in both *Nicaragua* and the *Oil Platforms* cases. See generally *Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.)*, Judgment, 1986 I.C.J. Rep. 14 (June 27); *Case Concerning Armed Activities on the Territory of the Congo (Dem. Rep. Congo v. Uganda)*, Judgment, 2005 I.C.J. Rep. 1 (Dec. 19).

292. William Burke-White & Andreas von Staden, *Investment Protection in Extraordinary Times: The Interpretation and Application of Non-Precluded Measures Provisions in Bilateral Investment Treaties*, 48 VA. J. INT’L L. 307, 311 (2008); see also Michael D. Nolan & Frédéric Sourgens, *The Limits of Discretion? Self-Judging Emergency Clauses in International Investment Agreements*, in *YEARBOOK ON INTERNATIONAL INVESTMENT LAW AND POLICY 2009–2010* (Karl P. Sauvant ed., 2010).

293. Burke-White & von Staden, *supra* note 292, at 371 (“It is difficult to subject highly policy-relevant terms like public order, health and morality, or essential security interests, to judicial evaluation by an *ad hoc* tribunal in the same way as other, more technical legal terms, such as “most-favored-nation” treatment.”).

294. *Id.* at 373.

295. *See id.*

296. *Id.* at 377.

297. See generally JOHN O’CONNOR, *GOOD FAITH IN INTERNATIONAL LAW* (1991).

298. Burke-White & von Staden, *supra* note 292, at 379–81.

good faith act honestly and fairly in performing treaty obligations is simply an essential component of the principle of *pacta sunt servanda*.²⁹⁹ The 1935 Harvard Draft Convention on the Law of Treaties articulates this aspect of good faith in these terms:

[t]he obligation to fulfill in good faith a treaty engagement requires that its stipulations be observed in their spirit as well as according to their letter, and that what has been promised be performed without evasion or subterfuge, honestly, and to the best of the ability of the party which made the promise.³⁰⁰

Thus, in the investor state context, it has been argued that it is open to review whether the state has acted honestly and to the best of its ability in invoking an NPM clause. Where there is evidence that the state has used the NPM as a pretext for ulterior motives, or where the connection between the state action and the alleged national security rationale is tenuous, it is open for a tribunal to decide that the invocation of the NPM was not undertaken in good faith.³⁰¹

The second element of the test shifts the focus from the motive and honesty underlying the decision-making process to the objective grounds for the judgment made, requiring that there be an objectively rational basis for the decision made in the circumstances.³⁰² This is the commonly understood reasonableness test, which requires that a reasonable person in the circumstances of the decision maker, with the information available at the time, could have reasonably reached the same conclusion. Such a good faith test has been applied in contexts other than the BIT and FCN context. It was considered by the ICJ, for instance, in a case involving a bilateral agreement for the provision of judicial assistance, in *Certain Questions of Mutual Assistance in Criminal Matters (Djibouti v. France)*.³⁰³

How should this good faith test apply in the context of the unwilling or unable doctrine, and moreover, what is the legal basis for such a requirement? To begin with the first question, in the circumstances we are addressing here, the good faith requirement would arguably apply to both the initial determination that self-judgment is appropriate, as well as to the exercise of judgment on the various substantive issues in question. For instance, if the target state determines that it need not seek consent from the territorial state because the risk is too high or because it has concluded that the territorial state is colluding with the NSA, that preliminary

299. O'CONNOR, *supra* note 297.

300. *Research in International Law*, 29 AM. J. INT'L L. 1, 981 (Supp. 1935).

301. Burke-White & von Staden, *supra* note 292, at 379.

302. *Id.* at 380.

303. *Certain Questions of Mutual Assistance in Criminal Matters (Djibouti v. Fr.)*, Judgment, 2008 I.C.J. Rep. 939, ¶ 145 (June 4). For analysis of the case, see Robyn Briese & Stephan Schill, *Djibouti v. France: Self-Judging Clauses Before the International Court of Justice*, 10 MELB. J. INT'L L. 308, 308 (2009).

determination should be subject to a good faith requirement. But in addition, the good faith test should operate to govern, and be the basis for the subsequent substantive assessment, that: (i) an NSA poses a threat of armed attack; (ii) the threat is imminent and that the use of force in self-defense is thus necessary to respond to it; and (iii) the territorial state is unwilling to either take action or consent to the target state doing so. Each of these judgments should be honestly and fairly reached, and each conclusion should be based on reasonable and objective evidence. What is more, these requirements further imply that there must be some degree of disclosure, at least after the fact, so as to permit an assessment of honesty and reasonableness. And these should be understood as threshold minimum requirements, before applying the differentiated standards discussed above, some of which impose higher evidentiary standards and higher probabilities of accuracy.

Turning to the second question, regarding the legal basis for a good faith requirement in these circumstances, this would appear to be less obvious than in the investor-state or FCN treaty context. This is because in the latter situation there are specific treaty provisions that provide for non-precluded measures, some of which explicitly reserve to the state the right to unilaterally decide when the provision is to apply.³⁰⁴ These, it has been argued, are clearly subject to the good faith requirement in Article 26 of the Vienna Convention on the Law of Treaties,³⁰⁵ and to the general principle of good faith that is part of the law of treaties.³⁰⁶ It might be objected that there are no such specific treaty provisions governing the judgments at issue in the context of the unwilling or unable doctrine, and thus no Article 26 good faith requirement can be said to apply. This objection could of course cut two ways, since it is the explicit language of the NPMs in BITs and FCN treaties that confers the right of self-judgment upon states, and so the absence of any such treaty authority could be said to militate against claims of a legal authority for target states to unilaterally make all the determinations at issue here. Indeed, this second argument is supported by language in the *Nicaragua* judgment that suggests that the unilateral invocation of NPMs will be afforded less deference in the absence of explicit treaty language signaling a state's reservation of the right to make such unilateral decisions.³⁰⁷ But assuming for the

304. Burke-White & von Staden, *supra* note 292, at 324–29 (explaining the different countries that have adopted non-preclusive measures as well as the language of these measures).

305. Vienna Convention on the Law of Treaties, art. 26, May 23, 1969, 1155 U.N.T.S. 331.

306. Burke-White & von Staden, *supra* note 292, at 379–81.

307. *Military and Paramilitary Activities in and Against Nicaragua* (Nicar. v. U.S.), Judgment, 1986 I.C.J. Rep. 14, ¶ 141 (June 27).

moment that we agree that the target state has the authority to exercise self-judgment in making some or all of these determinations, the principle of good faith in the law of treaties would then arguably still apply.

First, while there are no specific NPM-type provisions at issue here, the determinations are being made in relation to the exercise of a right under the UN Charter. The implicit claim of the unwilling or unable doctrine—that the target state is entitled to unilaterally make a number of key determinations regarding the justification for the use of force against NSAs in a nonconsenting state—is based on an interpretation of the treaty right of self-defense provided for in Article 51 of the UN Charter, and the related principle of customary international law.³⁰⁸ That being so, it follows that the unilateral determinations so authorized by Article 51 must be similarly subject to the principle of good faith provided for in Article 26 of the Vienna Convention on the Law of Treaties, in much the same way as the self-judgments authorized by NPMs in BITs and FNCs. And to the extent that the parallel doctrine of self-defense in customary international law is also said to similarly authorize these unilateral determinations, then the parallel obligation to perform customary international law obligations in good faith would apply.³⁰⁹

This may lead to the objection that if the foregoing argument is correct, then it must also follow that the “normal” self-judgments undertaken in the course of the exercise of self-defense against another state—that is, that the state is the victim of an actual or imminent armed attack, and that a use of force is necessary to respond to such attacks—would also be subject to good faith requirements. The short answer to this objection is simply “yes, that is correct.” While it has not been the explicit focus of the ICJ decisions regarding the doctrine of self-defense, there can be little question that the determinations of the defending state are subject to the requirement that the decisions be honestly made, and that they be rationally made on the basis of objective evidence. The debates surrounding the legitimacy of actual claims of justifiable self-defense frequently center on questions of whether the defending state had a legitimate purpose or was using self-defense as a pretext to disguise nefarious motives, and whether the defending state had a reasonable basis for determining that it was under attack—reflecting precisely the elements of the good faith

308. See *id.* ¶¶ 175, 177 (discussing the close overlap of treaty and custom as it relates to the doctrine of self-defense).

309. On the link between customary international law obligations and rights and the requirement to perform them in good faith, see *North Sea Continental Shelf Cases* (Ger. v. Den.; Ger. v. Neth.), Judgment, 1969 I.C.J. Rep. 1, ¶ 85 (Feb. 20); *Nuclear Tests Case* (Austl. v. Fr.), Judgment, 1974 I.C.J. Rep. 253, ¶ 46 (Dec. 20).

test.³¹⁰ What is more, as discussed earlier, in the operation of the broader doctrine of self-defense, there is typically an expectation that the defending state will provide disclosure of its reasons and the evidence upon which its decisions were made.³¹¹ There is always the prospect of *ex post* review, if not in the ICJ, then in the court of public opinion in the international community.³¹²

D. Transparency and Accountability

This last point brings us back to the issue of transparency and accountability. There has been a widely noted lack of transparency surrounding the American drone-based targeted killing policy.³¹³ Indeed, it has been shrouded in secrecy. For many years the US government would not even confirm that it was engaging in targeted killing, and even now the legal arguments, rules of engagement, evidentiary standards, and many other aspects of the policy remain classified.³¹⁴ There have been increasing arguments that greater transparency and accountability are essential to the legality of the operations against NSAs more generally.³¹⁵ The arguments for transparency and accountability have been specifically made with respect to IHL and international human rights law aspects of targeted

310. See, e.g., FRANCK, *supra* note 21, at 53–108; RUYS, *supra* note 3, at 83–98. For typical analysis of the ICJ into the validity of claims of self-defense, see *Nicaragua*, 1986 I.C.J. Rep. ¶¶ 229–38.

311. See sources cited *supra* note 239.

312. See sources cited *supra* note 240; see also generally U.K. COMM. OF PRIVY COUNSELLORS, THE REPORT OF THE IRAQ INQUIRY (2016), <https://ia800406.us.archive.org/0/items/ChilcotReportOneVolume/ChilcotReportOneVolume.pdf> [<https://perma.cc/FFM8-YSJC>] (archived Mar. 13, 2019).

313. See, e.g., BROOKS, *supra* note 11, 104–28; CHARLIE SAVAGE, POWER WARS: INSIDE OBAMA'S POST-9/11 PRESIDENCY chs. 6, 9 (2015); THE DRONE MEMOS, *supra* note 281, at 23–35.

314. What has been disclosed include S. REP. NO. 113-7 (2013); US LEGAL AND POLICY FRAMEWORKS, *supra* note 2; DoJ White Paper, *supra* note 71; OLC AL-AULAQI MEMO, *supra* note 148; PROCEDURES FOR APPROVING DIRECT ACTION AGAINST TERRORIST TARGETS LOCATED OUTSIDE THE UNITED STATES AND AREAS OF ACTIVE HOSTILITIES (May 22, 2013), <https://fas.org/irp/offdocs/ppd/ppg-procedures.pdf> [<https://perma.cc/3SPN-339T>] (archived Feb. 17, 2019).

315. Ben Emmerson, *Third Annual Rep. of the Special Rapporteur on the Promotion and Protection of Human Rights and Fundamental Freedoms While Countering Terrorism*, ¶ 35, U.N. Doc. A/68/389 (Sept. 18, 2013); BROOKS, *supra* note 11, at 34; NILS MELZER, HUMAN RIGHTS IMPLICATIONS OF THE USAGE OF DRONES AND UNMANNED ROBOTS IN WARFARE 37 (2013); Alston *CIA*, *supra* note 176, at 308. For my addition to these arguments, see Craig Martin, *A Means-Method Paradox and the Legality of Drone Strikes in Armed Conflict*, 19 INT'L J. HUM. RTS. 142, 152–54 (2015).

killing operations.³¹⁶ These claims are grounded in treaty provisions that implicitly require transparency for purposes of monitoring compliance, and accountability in the form of preliminary *ex post* assessment of all operations, followed by thorough investigation of those actions that raise questions of lawfulness.³¹⁷ But similar arguments should apply to the *jus ad bellum* aspects of operations against NSAs in nonconsenting states. Transparency and accountability are integral to the very essence of the rule of law, since the assessment of fair implementation of, compliance with, and enforcement of the law is impossible without both. The good faith obligations considered above, as well as the differentiated higher standards articulated to govern each of the determinations made by the target state, all contemplate and require the possibility of *ex post* assessment of whether target states have made the decisions leading to such operations honestly, and on the basis of the requisite level of evidence, satisfying the probability of accuracy required by the particular standard in question. The increased standards that I have argued for would require disclosure in order for the international community to monitor compliance, and for a more formal process of procedural review in the event of a challenge to the legality of a use of force.

In concluding this discussion of self-judgment and the differentiated standards that should apply to the target state's unilateral determinations, it may be helpful to distill and summarize here the standards that should apply. There are four decisions implicated, namely whether: (i) the threat posed rises to the level of an armed attack; (ii) a use of force in response is necessary (which may require a determination of imminence); (iii) the territorial state has explicitly refused assistance or consent (explicit unwillingness); and (iv) the target state may infer, and does infer, that the territorial state is implicitly consenting, or is unwilling to consent (an alternative to option (iii)). All of these determinations, as forms of self-judgment, are subject to a minimum standard of good faith, requiring honest dealing and a reasonable objective basis for decisions. But they are also subject to further differentiated standards.

Decisions as to armed attack and necessity are similar to a state-on-state attack, but affirmative determinations are not sufficient to justify the use of force—either consent or unwillingness is also required. There should be a presumption that target states will explicitly seek assistance or consent from territorial states. The determination of explicit unwillingness requires that the target state: (a) make disclosure to the territorial state, providing sufficiently compelling evidence to demonstrate that its determination that the

316. Alston *CIA*, *supra* note 176, at 310, 317.

317. See, e.g., Emmerson, *supra* note 315 (discussing the principle of accountability in international human rights law).

NSA poses a risk of imminent armed attack, which makes the use of force necessary, is very likely accurate; and (b) to meaningfully consult with the territorial state, giving reasonable and good faith consideration to alternative views and evidence the territorial state may share.

There should be a presumption against unilateral inferences of either consent or unwillingness. Rebutting this presumption requires that the target state have sufficiently compelling evidence to support that conclusion with a high probability of correctness, from the position of an objective and reasonable person, and it should be required to disclose this evidence *ex post* to justify the action. In the event of inferring “unwillingness,” the test is “collusion,” “substantial involvement,” or “adoption,” in line with the established *jus ad bellum* attribution test, not mere “harboring.” And finally, the inability of the territorial state to address the threat, by itself and without denial of consent (and thus unwillingness), cannot justify a use of force by the target state. All of these standards require greater transparency and accountability, and the inferences of unwillingness or consent, in particular, require the target state to provide significant *ex post* disclosure.

V. CONCLUSION

Out of the ashes of two catastrophic world wars, the United Nations system was designed with the primary objective of enhancing international peace and security. The modern *jus ad bellum* regime was developed at the core of this system for the purpose of reducing the scourge of war among states. Notwithstanding the prevalence of civil wars and insurgencies in recent years, the system has been more successful than is often realized—so successful perhaps that we have become complacent about the risk of armed conflict among powerful states.³¹⁸ Yet at a time when the entire system of international law and institutions is under strain, and instability in the relations among the great and emerging powers is increasing, we should not be too sanguine about this risk of war, or indifferent to threats to the legal regime that was designed to reduce that risk.

The unwilling or unable doctrine, embodied and embraced in the form of the Bethlehem Principles, poses just such a risk to the *jus ad bellum* regime. While the doctrine is relied upon for the narrow purpose of using force against NSAs within the territory of nonconsenting states, primarily as an aspect of counterterrorism

318. See generally HATHEWAY & SHAPIRO, *supra* note 23.

operations, it is claimed to be part of the doctrine of self-defense in the *jus ad bellum* regime. The doctrine perverts the concept of imminence, and distorts the definition of armed attack, in ways that would lower the threshold conditions for the justifiable use of force in self-defense for all purposes, not just against NSAs in nonconsenting states. This is exacerbated by implications that the law of state responsibility could bolster claims of justification for the use of force in ways that confuse the relationship between *jus ad bellum* and the law of state responsibility, and again cause confusion regarding the justification for the use of force. Similarly, aspects of the doctrine destabilize the well-established relationship between the *jus ad bellum* and IHL regimes, weakening both in the process. Finally, by providing weak standards to govern when the target state may make unilateral determinations as to the necessity for the use of force, and the consent or unwillingness of the territorial state regarding such use of force, the doctrine grossly over-privileges the interests of the powerful target states at the expense of the rights and interests of typically weaker territorial states. The perception that the law is being developed yet again to instrumentally serve the interests of the powerful states only serves to further erode the legitimacy of the *jus ad bellum* regime and the international rule of law more generally. This perception is exacerbated by the fallacious claims that the doctrine has achieved the status of customary international law, on the basis of the practice of a handful of powerful Western states.

Having said all of this, it must be recognized that states are unlikely to accept the proposition that they cannot use force against NSAs that pose a legitimate threat to their national security simply because the state in which they are operating is unwilling to provide consent. The efforts to develop the unwilling or unable doctrine into a more robust and law-based set of principles, so as to create a more realistic and substantive framework to govern the use of force against NSAs in nonconsenting states, are to be applauded. But those efforts have been insufficient, given the problems addressed above. The doctrine can and must be brought into greater compliance with the established principles of the *jus ad bellum* regime. And this can be done without overly weakening the effectiveness of the doctrine in addressing the real threat of transnational terrorism. Several of the more problematic aspects of the doctrine, such as those perverting the concept of imminence and distorting the definition of armed attack, can be adjusted without nullifying the core operation of the doctrine. Similarly, the establishment of differentiated and more stringent standards for governing the self-judgment of target states in determining the necessity and justification for the use of force would not preclude the operation of the doctrine. These adjustments will of course shift some of the risks, and provide a greater balance as between the rights and interests of target states and territorial states. But it will also reduce the risk of error, increase the legitimacy of the

doctrine, and diminish the very real threat that the current formulation of the doctrine poses to the *jus ad bellum* regime. Just as this Article was going to press, India responded to a terrorist suicide bombing by conducting air strikes against NSA targets within Pakistan, which India claimed had been at best unwilling to suppress the terrorist groups.³¹⁹ One of India's planes was shot down. The world held its breath at the prospect of war between nuclear-armed states, an escalation facilitated in part by operation of the unwilling or unable doctrine. This is not to trivialize or dismiss the very real risk posed by transnational terrorism—but efforts to reduce that risk cannot displace all other considerations. At the end of the day, there has to be a recognition that a doctrine developed to address the narrow and specific threat posed by transnational terrorism cannot be allowed to undermine the regime designed to maintain international peace and security. War among states poses far greater risks than that posed by sporadic terrorist attacks.

319. Srinivas Burra, *Legal Implications of the Recent India-Pakistan Military Standoff*, OPINIO JURIS (Mar. 8, 2019), <http://opiniojuris.org/2019/03/08/legal-implications-of-the-recent-india-pakistan-military-standoff/> [https://perma.cc/A3T2-877P] (archived Mar. 13, 2019).
