NOTES AND COMMENTS

CLARIFYING NECESSITY, IMMINENCE, AND PROPORTIONALITY IN THE LAW OF SELF-DEFENSE

*By Dapo Akande and Thomas Liefländer**

The concepts of necessity, imminence, and proportionality play a central part in Daniel Bethlehem's sixteen proposed principles regulating a state's use of force against an imminent or actual attack by nonstate actors.[1] While all three are requirements that must be considered in the law of self-defense, their exact content remains somewhat unclear. In this comment, we examine how each one is conceived in Bethlehem's principles and review the questions that remain unanswered.

In our view, Bethlehem's principles would benefit from greater conceptual clarity and precision with regard to necessity, imminence, and proportionality, especially in relation to anticipatory self-defense. As we discuss below, there are good reasons to think that reactive and anticipatory self-defense are not very different conceptually. However, that an armed attack has already occurred in one case and is merely expected in the other makes a substantial difference in the assessment of these three requirements. While such an assessment can be and is treated more leniently in reactive self-defense, a strict application of the rules governing their use is crucial in anticipatory self-defense. Accepting vague general principles, rather than precise standards, weakens the law's power to impose meaningful restraints in this area.

In particular, the use of self-defense against nonstate actors further militates for more precision. That a host state may be subjected to defensive force by another state even where it has not colluded in the nonstate actors' activities (and may even have done all it can to suppress their activities) is significant. In the classical state-to-state scenario of self-defense, the state subjected to defensive force has either already attacked the defending state or is about to do so. In that scenario, the attacking state has rendered itself liable to defensive force, which somewhat lowers the burden of justification for the defending state. Where a host state has not made itself liable to defensive force and may not have even been asked for consent, the burden that must be overcome should be greater. Only precise limits on the force that can lawfully be used against that host state impose meaningful legal restraints. Moreover, the contribution (or lack thereof) of the host state may subtly change the necessity and proportionality calculi. Where these concepts lack precision, subtle but important differences might be overlooked. If the host state that is merely unable or even entirely innocent is to receive the protection under international law

* Dapo Akande is University Lecturer in Public International Law, Co-Director of the Oxford Institute for Ethics, Law and Armed Conflict, and Yamani Fellow at St. Peter's College, University of Oxford. Thomas Liefländer is a doctoral candidate at the Faculty of Law, University of Oxford.

[1] Daniel Bethlehem, *Self-Defense Against an Imminent or Actual Armed Attack by Nonstate Actors*, 106 AJIL 769 (2012).

https://doi.org/10.5305/amerjintelaw.107.3.0563 Published online by Cambridge University Press

that it deserves in comparison to the colluding state, imprecision with regard to necessity, imminence, and proportionality is hardly tolerable.

Necessity and Imminence

Of the three concepts, necessity is the least problematic. Bethlehem's allusion thereto in principle 2 essentially coincides with the requirement that self-defense be used only where no peaceful alternative exists.[2] In theory, this requirement applies equally to any action in self-defense. However, practice indicates that conformity with necessity is assessed in different ways depending on the sort of actual or threatened attack to which a state is responding. When self-defense is used in immediate response to an ongoing attack by another state, practice indicates that a state is not required to seek or use alternative means. In this scenario, even when a state's use of force in self-defense would be futile, or when it has a more realistic chance of achieving a cessation of the attack by other means, practice does not deny that the state has an "inherent" right to use force to try to defend itself. In other words, there seems to be an almost irrebuttable presumption here that such a use of force in self-defense would pass the test of necessity.[3]

By contrast, the necessity test is no mere formality when the use of force is in response to an attack by a nonstate actor or when there is no ongoing use of force at that moment, either because no actual armed attack has yet occurred or because the armed attack appears to have ceased.[4] In such cases, necessity becomes a critical gateway for considering whether a forcible response is permitted at all. Particularly in these circumstances, it is important to clarify what exactly necessity requires. For example, is an active exhaustion of peaceful alternatives necessary, or does the mere appreciation that an alternative is not equally effective suffice?

Conceptually more challenging is the relationship between necessity and imminence. Imminence describes a certain pressing quality that a threat must have for anticipatory self-defense to be lawful. Bethlehem rightly notes that there is "little scholarly consensus on what is properly meant by 'imminence' in the context of contemporary threats" and that "the concept needs to be further refined and developed,"[5] which he attempts to do in principle 8.

To clarify imminence, we first consider the concept of threat. A threat is a situation where a causal chain can lead from the status quo (no attack) to an undesired future (attack). The essential components of a threat are fourfold: (1) type—what kind of attack is threatened? (2) likelihood—how probable is it that the attack will occur? (3) gravity—how severe will the attack be? and (4) timing—when will the attack occur? Bethlehem lists all four factors,[6] but further questions must be asked: what exactly is required of each of these elements, and how

[2] *See* 2 ARNOLD MCNAIR, INTERNATIONAL LAW OPINIONS 221, 222 (1956) (noting the *Caroline* incident of 1837); Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), 1986 ICJ REP. 14, para. 237 (June 27) [hereinafter *Military and Paramilitary Activities*].

[3] *See* Christian J. Tams & James G. Devaney, *Applying Necessity and Proportionality to Anti-terrorist Self-Defence*, 45 ISR. L. REV. 91, 97 (2012); JAMES GREEN, THE INTERNATIONAL COURT OF JUSTICE AND SELF-DEFENCE IN INTERNATIONAL LAW 84 n.122 (2009) (with further references).

[4] The use of force by the United States against two Iranian oil platforms is such a case. The International Court of Justice here held that "the requirement of international law that measures taken avowedly in self-defence must have been necessary for that purpose is strict and objective." Oil Platforms (Iran v. U.S.), 2003 ICJ REP. 161, para. 73 (Nov. 6). While this dictum, on its face, applies to all uses of force in self-defense, wider state practice indicates that it should be read with a grain of salt and viewed in light of the particular circumstances.

[5] *Bethlehem*, *supra* note 1, 773–74.

[6] *Id.* at 775, princ. 8, pts. a, b, d.

https://doi.org/10.5305/amerjintelaw.107.3.0563 Published online by Cambridge University Press

do they interact? There is a far-reaching consensus that a threat with a very low likelihood of materialization does not allow anticipatory self-defense. But where is the cutoff mark? Similarly, according to traditional doctrine, only a threat reaching some threshold of gravity qualifies. Again, where is this threshold? Even more interesting from a conceptual point of view is whether the requirements of likelihood and gravity are independent or whether they must be aggregated, in the sense that a higher probability is required where a threat is relatively mild and that a much lower probability is acceptable where a threat is extremely severe. These questions are at the core of the debate about preemptive versus anticipatory self-defense.

Likelihood and gravity of an attack are also connected to another factor that Bethlehem mentions with regard to imminence: "whether the anticipated attack is part of a concerted pattern of continuing armed activity."[7] While this factor is important, it does not seem to add something not already captured by likelihood or gravity. The only difference that a continuing series of attacks makes is that establishing the threat becomes much easier, as identifying "a concerted pattern of continuing armed activity" usually allows a more reliable prediction that a threat within this pattern will materialize with sufficient gravity.

In addition to likelihood and gravity, imminence seems, on first sight, to require that the threat will materialize within a short time frame. However, there are good reasons to suggest that imminence entails no independent temporal requirement, even though this claim may sound like a paradox. For one thing, such an independent temporal limitation would mean that where a highly probable and severe threat exists, whose realization is temporally remote, no action could be taken even where no future opportunity will arise to eliminate the threat. The better argument is that where a threat is sufficiently probable and severe, the mere fact that it is still temporally remote should provide no independent injunction against action where that action is necessary and proportionate. The International Court of Justice (ICJ) interpreted imminence in this manner in the context of necessity as a circumstance precluding wrongfulness, essentially eliminating any independent requirement of temporal proximity.[8]

However, denying the existence of an *independent* temporal limitation does not mean that temporal factors are unimportant. They have a heavy impact on the possibility of making accurate predictions about both the likelihood and the gravity of a threat.[9] The shorter a causal chain is, the easier it becomes to predict what will occur. It will be harder to establish that a threat is sufficiently probable and severe if such a threat is still very far away in a temporal sense. In addition, and more importantly, the temporal dimension affects necessity. As noted, necessity allows using force only where no peaceful alternative is available. Thus, where a military option will be available for some time because the threat is temporally remote, other options should be tried first. Other scholars, including those involved in the Chatham House principles,[10] have read the imminence requirement similarly, focusing on the last point in time at which an effective responsive action is possible, rather than temporal proximity per se.[11] What is really at stake is whether some sort of self-defense action is demonstrably necessary—without any alternative, including later in time—rather than how temporally remote the threat is.

[7] *Id.*, princ. 8, pt. c.

[8] Gabčíkovo-Nagymaros Project (Hung./Slovk.), 1997 ICJ REP. 7, para. 54 (Sept. 25).

[9] *See* NOAM LUBELL, EXTRATERRITORIAL USE OF FORCE AGAINST NON-STATE ACTORS 53–54 (2010).

[10] Published as Elizabeth Wilmshurst, *The Chatham House Principles of International Law on the Use of Force in Self-Defence*, 55 INT'L & COMP. L.Q. 963, 967–68 (2006).

[11] Vaughan Lowe, *'Clear and Present Danger': Responses to Terrorism*, 54 INT'L & COMP. L.Q. 185, 192 (2005).

https://doi.org/10.5305/amerjintelaw.107.3.0563 Published online by Cambridge University Press

Necessity is likewise at the core of a defending state's obligation to seek an unable host state's consent to military action. In taking defensive actions against nonstate actors, there is no need to violate the territorial sovereignty of another state and to resort to a self-defense justification where effective action can be taken with the consent of the host state. Bethlehem asserts that a defending state is released from the obligation to seek such consent where doing so "would be likely to materially undermine the effectiveness of action in self-defense . . . or would increase the risk of armed attack [or] vulnerability to future attacks."[12] Thus, the state acting in self-defense can only be released from the obligation to seek consent when it is necessary to act without consent. In circumstances where seeking consent from the host state would materially undermine the effectiveness of defensive action or increase the defending state's vulnerability to the attack, it may be necessary to use force in self-defense without such consent because the defending state should not have to adopt a course of conduct that would be ineffective in countering the threat posed to it.

However, it is problematic to dispense with the requirement to seek consent from the host state simply because there will be an increased risk of armed attack. It should not count *as such* whether the risk is enhanced, but rather whether the ability to respond efficiently is reduced. Imagine a scenario where terrorists are hiding in the host state's mountains. The defending state knows that these terrorists are plotting an attack but are not yet ready. Assume further that the threat is sufficient to justify anticipatory self-defense and that an air strike would eliminate it. Seeking consent to carry out the air strike may take time, during which the terrorists grow more numerous and get closer to readiness. In abstract terms, the threat has increased. Nevertheless, the air strike would still be fully effective. Would the defending state be justified in omitting attempts to obtain the host state's consent? The central consideration is again one of necessity: would seeking consent deprive the defending state of the possibility to act effectively? Only if the answer is yes is immediate action justified.

What we are left with is that only the likelihood and gravity of the attack are genuine elements of the concept of imminence, and their actual content is mostly left vague. Apart from these factors, much of what is often considered under imminence—notably the temporal factor—more properly belongs to necessity.

Proportionality and Self-Defense

There is a profound lack of clarity and consensus as to the test to be applied with regard to the proportionality requirement in the *jus ad bellum*. At least three different conceptions of proportionality have been associated with self-defense, and each has different implications. First, proportionality may simply be used to describe the requirement that the defending state use no more force than is necessary. Second, proportionality might mean that the defensive action must be quantitatively "commensurate" either with the attack to which it is responding or with the threatened attack. Third, proportionality may require that the damage inflicted in self-defense not be disproportionate in comparison to the pursued objective. Bethlehem is not clear as to which conception he seeks to incorporate, stating only that use of force "must be

[12] Bethlehem, *supra* note 1, at 776, princ. 12.

https://doi.org/10.5305/amerjintelaw.107.3.0563 Published online by Cambridge University Press

limited to what is necessary to address an imminent or actual armed attack and must be proportionate to the threat that is faced."[13]

Since Bethlehem lists the requirement that armed action in self-defense "must be limited to what is necessary" separately from "must be proportionate,"[14] it is not clear whether he deems the former to be an aspect of proportionality. Indeed, the inclusion of proportionality—in addition to the requirement of doing no more than is necessary—suggests that he considers that international law imposes two further requirements beyond the necessity requirement of "no choice of means," namely, a requirement to use no more force than necessary and a requirement that such use be proportionate, however that term is defined. Or perhaps he considers "no more force than necessary" as simply an aspect of necessity more generally. On this view, necessity would impart two requirements: to use no force where a peaceful alternative exists, and to use no more force than necessary where that is not the case. It is certain that international law imposes both conditions when a state is evaluating the use of force in self-defense,[15] particularly where a state is using force in self-defense against a nonstate actor.[16]

If the requirement not to do greater harm than necessary is considered to be part of the *jus ad bellum* concept of necessity, separate content must be given to the requirement of proportionality, as Bethlehem, indeed, appears to do. What would that additional requirement entail? One would be left with the second and third variants outlined above. However, it is not at all clear that the law of self-defense endorses either of these conceptions of proportionality, or that it should. Although the ICJ has, on occasion, appeared to suggest that the *jus ad bellum* proportionality calculation involves a comparison of the use of force in self-defense with the original attack,[17] some have rejected this view and have asserted that proportionality simply requires that a state acting in self-defense do no more harm than is necessary.[18] However, some dicta of the Court can also be read as going in the direction of comparing the harm done and the objective pursued.[19]

A decision as to which notion of proportionality is required has fairly significant implications for self-defense operations. Imagine a state faced with a threat of an armed attack that just barely reaches the threshold of an armed attack (e.g., a series of suicide bombers). Let's say that the state under threat knows that a full-fledged invasion of the country in which the attackers are hiding is the only way to prevent the attack. If the first reading of proportionality ("do no more harm than necessary") is adopted, the only limitation that it imposes is whether the same objective could be achieved by using less force, which is not the case here. The second reading of proportionality (commensurability of harm in self-defense and of harm caused or threatened

[13] *Id.* at 775, princ. 3.

[14] *Id.*

[15] *Cf. Oil Platforms*, *supra* note 4, Sep. Op. Kooijmans, J., para. 62.

[16] *See* Kimberley N. Trapp, *Back to Basics: Necessity, Proportionality, and the Right of Self-Defence Against Non-state Terrorist Actors*, 56 INT'L & COMP. L.Q. 141 (2007); Tams & Devaney, *supra* note 3.

[17] *Military and Paramilitary Activities*, *supra* note 2, para. 237; *Oil Platforms*, *supra* note 4, para. 77 (majority opinion).

[18] Legality of the Threat or Use of Nuclear Weapons, Advisory Opinion, Diss. Op. Higgins, J., 1996 ICJ REP. 226, 583, para. 5 (July 8) [hereinafter *Nuclear Weapons*]; *see also, e.g.*, JUDITH GARDAM, NECESSITY, PROPORTIONALITY AND THE USE OF FORCE BY STATES 16, 156–67 (2004).

[19] *Nuclear Weapons*, *supra* note 18, paras. 41–44, 97 (majority opinion); Armed Activities on the Territory of the Congo (Dem. Rep. Congo v. Uganda), Sep. Op. Kooijmans, J., 2005 ICJ REP. 168, 306, para. 34 (Dec. 19); Enzo Cannizzaro, *Contextualizing Proportionality:* Jus ad Bellum *and* Jus in Bello *in the Lebanese War*, 88 INT'L REV. RED CROSS 779, 792 (2006).

https://doi.org/10.5305/amerjintelaw.107.3.0563 Published online by Cambridge University Press

in original attack) would bind the state to use only the same quantum of force as the attacker is threatening and would clearly disallow the invasion. Under the third reading, the answer is not as clear since it effectively requires a comparison of the harm threatened and the benefit pursued, which, in turn, requires complex value judgments. The resulting question becomes: What weighs more in the particular circumstances of the case, the interest in preventing the attack or the interests that are violated by using massive counterforce?

In the context of self-defense against nonstate groups, the distinction between a colluding state and an unable state further helps to see more clearly the difference between the variants of proportionality. In considering that consent is not required where a state hosts terrorists and colludes in their activities, Bethlehem concludes that "the extent of the responsibility of that state for aiding or assisting the nonstate actor in its armed activities may be relevant to considerations of the necessity to act in self-defense and the proportionality of such action . . . ."[20] While the impact on necessity stems from the fact that collusion by the host state reduces the array of possible alternative measures[21] and may necessitate action against the host state itself, it is not immediately clear how such collusion relates to the proportionality calculation: it all depends on which variant of proportionality is adopted. While the host state's contribution affects the first conception of proportionality by changing the array of available options, the second variant is oblivious of the circumstances of the (threatened) attack that provokes defensive action. Only the quantum of force used or threatened, rather than the identity of the aggressor involved, matters in this option, and, consequently, the involvement of the host state does not factor here into the amount of force that may be used in self-defense. In the third variant, the host state's actions inform the evaluation of the damage inflicted by the defending state. In our scenario (where massive counterforce is needed to prevent a relatively minor attack), the third conception of proportionality may allow defensive force when the host state is colluding in the threatened attack but would likely render it disproportionate where the host state is merely unable or innocent.

Choosing one conception of proportionality over another involves heavy normative implications, as the second and third conceptions may render an act of self-defense unlawful even where the defending state could not have achieved its objectives by using less force. In addition, the second variant may exclude action simply because it goes quantitatively beyond the attack or threat provoking it, even where good reasons exist to conclude that inflicting harm on a colluding state should count less than preventing harm to the (innocent) defending state. In contrast, the third variant may allow defensive action to go far beyond the action provoking it, based on nothing but complicated evaluative questions of balancing interests.

Self-Defense and Permissible Objectives

A final significant problem underlying much of the law of self-defense (and not addressed in Bethlehem's principles) is the lack of clarity regarding the permissible objectives of action

[20] Bethlehem, *supra* note 1, at 776, princ. 11.

[21] For example, important options—such as seeking consent for intervention, convincing the host state to act against the nonstate actor, or undertaking some joint operation—are not available where the host state itself is involved in the terrorist activities. Indeed, even exploring such options might not be realistic where information might be leaked to the terrorist forces. In contrast, where the host state is merely unable to curb the terrorist activities, a wider array of options can generally be explored.

https://doi.org/10.5305/amerjintelaw.107.3.0563 Published online by Cambridge University Press

in self-defense. There is a remarkable paucity of guidance on what objectives a state is allowed to pursue in self-defense. It is not sufficient to assert that the state must not do more than is necessary to "address an imminent or actual armed attack."[22] Indeed, defining the legitimate aims of self-defense is perhaps the crucial question in this area of the law.[23] The classic view holds that self-defense is entirely defensive (i.e., preventive).[24] Nevertheless, the language of deterrence has crept into state practice, raising the question of precisely what objectives a defending state is entitled to pursue. Even more significantly, especially as regards self-defense against nonstate actors, occasional claims of "retaliatory self-defense" have emerged.[25]

Knowing what objectives a defending state may permissibly pursue in self-defense is central to applying the concepts of necessity and proportionality. Necessity relates to the means required to achieve an objective, and, in law, only a legally permissible objective counts. Where the permissible objectives are unclear, how can the necessity of an action to achieve such an objective ever be judged? Similarly, under a conception of proportionality that relates the objective pursued to the damage inflicted, how can we compare a vague objective to very concrete damage? Without a good idea of the permissible objectives of self-defense, necessity and proportionality remain, even if conceptually clear, difficult to apply.

The choice of permissible objectives has further implications for conceptualizing the difference between anticipatory and reactive self-defense. If one accepts that self-defense is purely preventive and not retaliatory, the conceptual difference between anticipatory and reactive self-defense fades away. Both are essentially about preventing a threat from materializing—either a continuing threat from an actual armed attack or from an imminent threat. The difference would be merely evidentiary or epistemic: where an armed attack has already occurred and is continuing, there is usually no doubt that a sufficient continuing threat exists. Where an armed attack has not yet taken place, establishing a threat is more difficult.

Even assuming that the purpose of self-defense is exclusively preventive, questions remain as to the legitimate objectives of preventive action. For example, where a state is faced with an ongoing pattern of attacks by a nonstate group acting from a territory across its border, the state is entitled to take defensive action, but with what objective? Is the state only entitled to act to stop the threat of immediate future attacks, or may it take action to prevent these attacks over the long run? The answer to that question will determine whether, for example, the state is only entitled to go across the border to destroy rocket launchers used to initiate the attacks, to destroy the base where the nonstate groups are camped, or, instead, to seek to change the government of the host state to prevent the territory from being used for future attacks.

Conclusion

Despite their constant invocation, the notions of necessity, imminence, and proportionality are fraught with conceptual ambiguity and are notoriously difficult to apply in practice. By relying heavily on necessity, imminence, and proportionality, Bethlehem imports these difficulties into his principles. Certainly, Bethlehem cannot be criticized for failing to solve all the

[22] *Id.* at 775, princ. 2.

[23] *See* David Kretzmer, *The Inherent Right to Self-Defence and Proportionality in* Jus ad Bellum, 24 EUR. J. INT'L L. 235 (2013).

[24] *See* VAUGHAN LOWE, INTERNATIONAL LAW 278 (2007).

[25] *See* Christian J. Tams, *The Use of Force Against Terrorists*, 20 EUR. J. INT'L L. 359, 391 (2009).

https://doi.org/10.5305/amerjintelaw.107.3.0563 Published online by Cambridge University Press

concerns that stem from the general law of self-defense. However, it may be that a set of principles that are designed, explicitly, not just to state the existing law but rather to "formulate principles that reflect, as well as shape, the conduct of states"[26] should strive for a higher degree of precision. Bethlehem is right to lament the "paucity of considered and authoritative guidance on the parameters and application" of the right of self-defense against imminent or actual armed and to stress the need for legal principles that are "capable of objective application."[27] However, more attention needs to be paid to what necessity, imminence, and proportionality actually require.

THE NONCONSENTING INNOCENT STATE: THE PROBLEM WITH BETHLEHEM'S PRINCIPLE 12

*By Dire Tladi**

In a recent issue of this *Journal*, Daniel Bethlehem proposed a set of principles on the scope of a state's right of self-defense against an imminent or actual armed attack by nonstate actors.[1] In response, this essay seeks to assess Bethlehem's proposition in principle 12 that a state may use force against nonstate actors on the territory of another state without the consent of the territorial state when the territorial state is not responsible for the initial attack and when the attack cannot be imputed to the territorial state. This description might be termed the *nonconsenting innocent state problem*.

The dilemma that I take on here arises when Bethlehem's principles are applied to the use of force against nonstate actors without the consent of the territorial state. Principle 1 indicates that "[s]tates have a right of self-defense against an imminent or actual armed attack by nonstate actors."[2] Principle 10 provides that, as a general rule, the consent of the territorial state is required whenever force is used against nonstate actors.[3] Yet principle 11 states that consent is not required when the territorial state colludes with or is "*unwilling to effectively* restrain the armed activities of the nonstate actor" carrying out armed attacks against another state.[4] While this proposition should be fairly uncontroversial, it is unclear what is meant by that quoted phrase. To the extent that this choice of language—in particular, the term *effectively*—is meant to suggest that consent may be dispensed with even in those instances where the territorial state takes reasonable but ineffective measures to prevent its territory from being used to launch attacks, problems arise that parallel those that I discuss in relation to principle 12. Principle 12, the second exception and the focus of this paper, provides that consent can be dispensed with

[26] Bethlehem, *supra* note 1, at 774.

[27] *Id.*

* Principal State Law Adviser (International Law), Department of International Relations and Cooperation of South Africa; Extraordinary Professor, Department of Public Law, University of Stellenbosch. The views expressed in this article are not attributable to the Government of South Africa or the Department of International Relations and Cooperation.

[1] Daniel Bethlehem, *Self-Defense Against an Imminent or Actual Armed Attack by Nonstate Actors*, 106 AJIL 769 (2012).

[2] *Id.* at 775, princ. 1.

[3] *Id.* at 776, princ. 10.

[4] *Id.*, princ. 11 (emphasis added) (footnote omitted).

https://doi.org/10.5305/amerjintelaw.107.3.0563 Published online by Cambridge University Press