CLAWS

ISSN 23939729

No. 136

2026

MANEKSHAW PAPER

Lethal Autonomous Weapons Systems (LAWS): Implications on Escalation Dynamics and Land Warfare with China

Arjun Uppal

CENTRE FOR LAND WARFARE STUDIES

Field Marshal Sam Hormusji Framji Jamshedji Manekshaw, better known as Sam "Bahadur", was the 8th Chief of the Army Staff (COAS). It was under his command that the Indian forces achieved a spectacular victory in the Indo-Pakistan War of 1971. Starting from 1932, when he joined the first batch at the Indian Military Academy (IMA), his distinguished military career spanned over four decades and five wars, including World War II. He was the first of only two Field Marshals in the Indian Army. Sam Manekshaw's contributions to the Indian Army are legendary. He was a soldier's soldier and a General's General. He was outspoken and stood by his convictions. He was immensely popular within the Services and among civilians of all ages. Boyish charm, wit and humour were other notable qualities of independent India's best known soldier. Apart from hardcore military affairs, the Field Marshal took immense interest in strategic studies and national security issues. Owing to this unique blend of qualities, a grateful nation honoured him with the Padma Bhushan and Padma Vibhushan in 1968 and 1972 respectively.

Field Marshal SHFJ Manekshaw, MC
1914-2008

CLAWS Occasional Papers are dedicated to the memory of Field Marshal Sam Manekshaw

MANEKSHAW PAPER No. 136, 2026

Lethal Autonomous Weapons Systems (LAWS): Implications on Escalation Dynamics and Land Warfare with China

Arjun Uppal

Centre for Land Warfare Studies
New Delhi

Editorial Team : CLAWS

ISSN : 23939729

Centre for Land Warfare Studies
RPSO Complex, Parade Road, Delhi Cantt, New Delhi 110010
Phone +91-11-25691308 Fax: +91-11-25692347
Email: **landwarfare@gmail.com**, website: **www.claws.co.in**
CLAWS Army No.33098

The Centre for Land Warfare Studies (CLAWS), New Delhi, is an independent Think Tank dealing with national security and conceptual aspects of land warfare, including conventional & sub-conventional conflicts and terrorism. CLAWS conducts research that is futuristic in outlook and policy-oriented in approach.

CLAWS Vision: To be a premier think tank, to shape strategic thought, foster innovation, and offer actionable insights in the fields of land warfare and conflict resolution.

CLAWS Mission: Our contributions aim to significantly enhance national security, defence policy formulation, professional military education, and promote the attainment of enduring peace.

Published in Bharat by

Sabre & Quill Publishers, New Delhi, India
www.sabreandquill.com/sabreandquill@gmail.com

Academic Acknowledgement

I extend my deepest gratitude to **Lieutenant General Shokin Chauhan, PhD**—my doctoral guide over the past three years—for his sustained mentorship, intellectual discipline, and unwavering, tireless insistence on rigor. His guidance fostered clarity of thought and independence of inquiry, while his steady counsel shaped not only the substance of my work, but also the structure and discipline with which I approach scholarly inquiry. He consistently encouraged me to articulate my ideas fully, without reservation in speech and writing.

I am equally indebted to **Dr. Celestino Perez, PhD**, my professor this past year in War, Policy and Strategy, for his exacting intellectual standards and rigor in understanding the operating environment, cognitive mapping, and the critical bridging of scholarly theory with practical application. His instructions on orchestrating instruments of national power and theoretical frameworks, combined with relentless effort towards deeper, more complex lines of inquiry, all while accepting nothing less than academic excellence, are unforgettable.

Personal Acknowledgement

To my wife, **Sonika**—whose tolerance for my oscillations between analytical confidence and existential despair has been nothing short of extraordinary. She endured, with quiet resilience, my periodic frustrations, near abandonment of the project, and disproportionate reactions to entirely solvable problems. If this work reflects any coherence or completion, it is as much a testament to her composure and patience as it is to my persistence.

Disclaimer:

Thus, any errors that remain are solely my own, notwithstanding the sustained efforts of my distinguished teachers to prevent them.

Contents

Lethal Autonomous Weapons Systems (LAWS): Implications on Escalation Dynamics and Land Warfare with China

Abstract

How do Lethal Autonomous Weapons Systems (LAWS) shape escalation dynamics between India and China in a chaoplexic, AI-driven battlespace? While existing research - from Scharre on autonomy to Sweijs and Zilincik on cross-domain deterrence, and from Winter, Galliott, Colijn, Podar, and Hanley on legality, ethics, risk, and Chinese escalation behavior provides valuable insights, these perspectives remain largely siloed. What is insufficiently explored is how technological, legal, ethical, and doctrinal factors interact when autonomy compresses decision cycles and obscures intent. To address this gap, the study analyzes a self-developed Scenario Building Exercise and employs two game-theory models - the **Prisoner's Dilemma** and **Backward Induction Theory**- to trace how unintended outcomes may arise. Using the **Institutional Analysis and Development (IAD) Framework**, it evaluates contextual variables, actor positions, and interaction patterns as sources of cumulative instability. The study argues that LAWS lower risk thresholds, heighten opacity, and accelerate decision-making, thereby pushing strategic behavior toward escalation. By so examining LAWS in structural terms, this study addresses a key conceptual gap with findings that may help develop an autonomous-era escalation continuum for the Sino-Indian context.

"***If we continue to develop our technology without wisdom or prudence, our servant may prove to be our executioner***"[1].

— Omar N. Bradley, Armistice Day speech, 1948

Genesis

With the fusion of "Artificial Intelligence (AI) and Machine Learning (ML)"[2] and rapidly proliferating, low-cost drone-based weapons, land warfare has seen a transformation. From initial deployment in Nagorno-Karabakh, AI and ML-based LAWS have seen widespread and contentious use in the Israel–Hamas conflict and the Russia–Ukraine wars. Over the past year, outcomes in battle have been shaped by "First-Person-View (FPV)"[3] drones, "One-Way-Attack (OWA)"[4] systems, and software capable of independent target classification. Turkey's STM Kargu-2 UCAV, which struck retreating Haftar-Associated Forces (HAF) without a Human–Robot Interface (HRI), marked the first recorded fully autonomous lethal engagement, contravening Paragraph 9 of Resolution 1970 (2011).

Existing literature on LAWS spans technological evolution, meaningful human control, legality, ethics, escalation dynamics, and the decade-long deliberations of the "Group of Governmental Experts (GGE)"[5] at the "Convention on Certain Conventional Weapons (CCW)"[6]. Yet these strands largely remain siloed. What is missing - and what this study addresses - is a fused examination of how an imaginary escalation scenario with China in the Indian subcontinent might unfold within a battlespace increasingly shaped by lethal autonomy.

This research is structured as follows: **Section I** assesses the technological impact of LAWS, the emerging autonomous battlefield, and the positions of Russia, China, and India in GGE deliberations. **Section II** analyzes legality, ethics, Meaningful

Human Control (MHC), and decision-making complexity. **Section III** examines escalation dynamics with China and uses a self-developed Scenario Building Exercise to illustrate unintended consequences. **Section IV** applies two game-theory models - the Prisoner's Dilemma and Backward Induction Theory- to map plausible outcome pathways. **Section V** uses the "Institutional Analysis and Development (IAD) Framework"[7] to consolidate key findings and recommendations and presents a proposed "Way Forward" in a summarized mind-map. This is followed by the Conclusion in **Section VI**.

This study assumes one autonomous model for the purpose of research, rather than detailing Levels of Autonomy (LOA) or Autonomous Levels (AL). The overarching aim is to clarify how a crisis with China might emerge, where the principal friction points lie, and which measures are essential to prevent escalation along the most dangerous pathways. Portions of this paper have been revised for conciseness using ChatGPT[8].

Section I: Technological Impact and Stance of Select Nations on LAWS, and the Future Battlefield Environment

"Contextual Autonomous Capability (CAC)"[9]

For this study, the "Unmanned Systems"[10] CAC model[11] provides the conceptual baseline for understanding how autonomous behavior emerges across operational conditions. Rather than treating autonomy as a linear scale, the model evaluates it as the interplay of three interdependent variables: "**Human Interdependence (HI)**"[12], "**Mission Complexity (MC)**"[13], and "**Environmental Complexity (EC)**"[14].

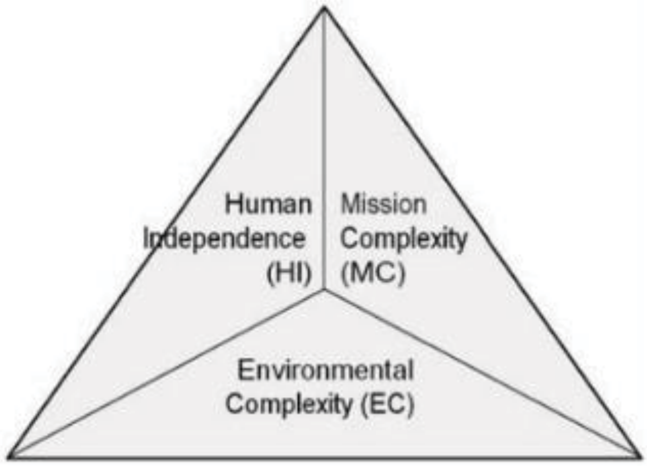


Figure 1: CAC aspects for "Autonomy Levels for Unmanned Systems (ALFUS)"[15]

At initial CAC levels, autonomous systems conduct simple missions through a **fully-engaged** Human–Robot Interface (HRI), operating entirely under human supervision in predictable, low-complexity environments. As missions intensify through contested terrain, adversarial interference, degraded communications, or multi-layered tasks, the system transitions into higher CAC thresholds. Each shift reflects a recalibration of how much initiative the unmanned systems must assume **when human operators cannot process or react at the required pace**.

"Low, mid, and high HI bands across the ten autonomy levels"[16] are illustrated in **Figure 2** below. The intent is not a technical dissection of all ten levels, but to clarify the concepts through which the subsequent analysis interprets LAWS behavior.

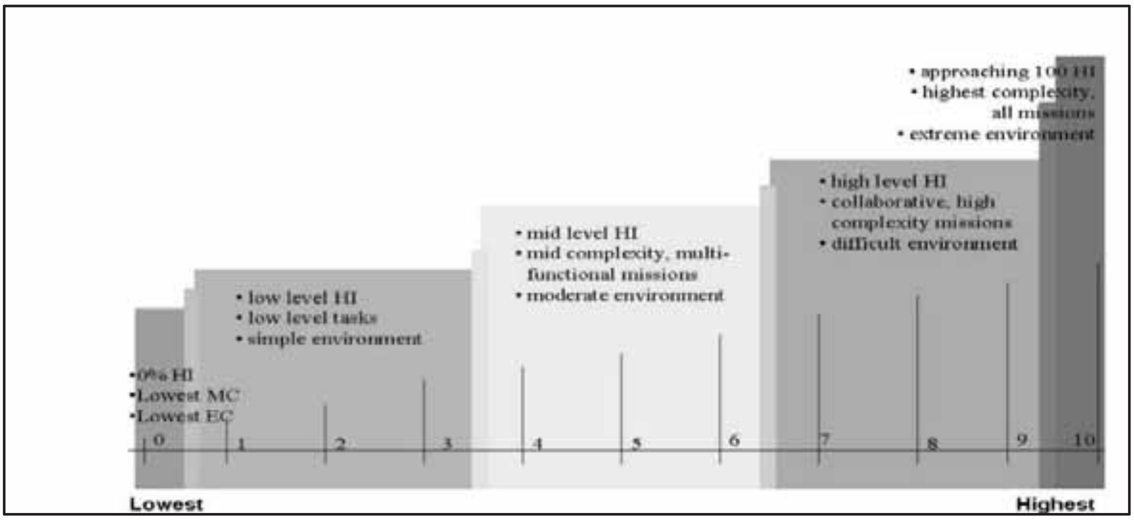


Figure 2: CAC Levels and HI-Mission Complexity Relationship

The Battlefield Environment

Today's battlefield environment is shaped by **Chaoplexity** – "a fusion of chaos and complexity defining the fourth regime of scientific warfare"[17]. This non-linear, information-saturated environment is marked by decentralization, autonomy, and swarming, resulting in a fundamental reconfiguration of how force is applied and advantage gained. What emerges resembles intelligent, adaptive swarms operating through algorithmic loops insulated from human rationale and driven by "internal rules of optimisation"[18].

In addition, more actors now compete to dominate the "**Air-Ground Littoral (AGL)**"[19] using massed, low-cost autonomous systems, making this zone increasingly decisive. Layered AWS - with cheap sensors, adaptive software, and resilient communications - can reshape conflict outcomes through persistence and distributed lethality.

The "**International Committee for Robots Arms Control (ICRAC)**"[20] warns that humans are prone to the "***What You See Is All There Is*** **(WYSIATI)**"[21] cognitive bias, in which decision-makers "anchor on what the system displays"[22]. Kahneman's "**Human**

Supervisory Control (HSC) model"[23] below shows how expanding autonomy steadily erodes human interpretive space and, with it, decision-quality: -

Level	HSC Classification	Implications
1	Human deliberates before initiating an attack.	Adequate time and bandwidth enable proportional, legally compliant judgments.
2	System provides a target list; human selects the target.	Manageable but vulnerable to pressure, distraction, and time compression.
3	System selects the target; human must approve.	Risk rises: flawed system recommendations can improperly anchor human judgment.
4	System selects target; humans have a restricted veto window.	Insufficient time for analysis undermines IHL compliance and situational awareness.
5	**Fully autonomous target selection and engagement.**	**Human involvement collapses, creating the highest technical, legal, and escalatory risk.**

This erosion of supervisory clarity fuels a perception gap: software-driven errors may appear more dangerous than human error, "magnifying anxiety in deterrence relationships"[24]. These systems are powerful yet brittle and are vulnerable to misclassification, spoofing, or adversarial manipulation. Yet, they undercut credible deterrence by creating "**unknown unknowns**" that "**compress decision-time and cloud judgment**"[25]. Demonstrations of capability can be escalatory when adversaries feel compelled to respond quickly or "risk ceding initiative"[26]. Compounding this friction would be the adversary's ability to distort data through counter-AI measures, "injecting corrupted inputs and provoking unintended engagements or accelerated decision cycles",[27] often without human concurrence. In such an environment, autonomy increases tactical efficiency while heightening strategic volatility.

Outcomes based on a Chaoplexic[28] and Autonomous Environment

In a chaoplexic battlespace that is fluid, data-dense, and non-linear - software, sensors, weapons, and communications will merge into tightly integrated networks. As systems coordinate, share information, and act with minimal human intervention, the risk of misinterpretation could rise sharply. Neural-network-driven systems "exhibit vulnerabilities fundamentally different from human error"[29]. One such vulnerability is "***catastrophic forgetting***"[30], in which training on new data inadvertently overwrites previously learned skills essential for safe engagement.

A second risk is ***memory overflow*** or "**contradictory neural cacophony**," where competing stimuli such as a thermal signature that must be engaged and a visual image (such as tanks standing in a populated area) that must not be engaged, could create a contradiction and be subjective to interpretation by the LAWS itself. In such a situation, any incompatible classification (such as a thermal signature emanated by tanks inside a populated area), could result in a default setting towards unsafe action.

The third vulnerability is the ***cocktail party effect***, where autonomous systems misjudge which signals matter most amid electromagnetic noise, heat signatures, command signals, and adversarial decoys. A wrongly prioritized signal at the wrong moment could trigger an unintended engagement despite broader lawful programming, and lawful human supervision.

Stance of Nations on LAWS

Russia, China, and India have articulated distinct positions during GGE deliberations at the CCW. These are significant, having seen how situations could turn unintentional, in the theories above.

Russia opposes negotiations on a treaty governing AWS, "signaling its intent to retain flexibility over meaningful human control"[31]. This aligns with Moscow's ambition to leverage AI as a strategic equalizer. Putin's assertion that "whoever masters AI will become the ruler of the world"[32] reinforces Russia's belief that autonomy is not a capability of restraint, but that of global influence. Moscow's obstructionism at the CCW has slowed consensus and introduced structural rigidity into the GGE process, enabling it to "continue developing its autonomous posture under the mask of cooperation"[33].

China's "New Generation Artificial Intelligence Development Plan"[34] seeks global leadership by 2030, and its GGE stance mirrors this trajectory. It has articulated five categories of weapons it deems "unacceptable.[35] Yet these criteria are sufficiently narrow that "many sophisticated autonomous systems fall outside the prohibited scope"[36]. China's threshold for autonomy is markedly higher than most nations, omitting meaningful human intervention at key stages. Its approach blends strategic ambiguity with calibrated signaling - "hide a dagger in a smile"[37] and the "battle of pride"[38]. An outwardly cordial posture masking deeper intent and long-term capability accumulation enables AI systems to transition from one classification of LAWS to another (from the "acceptable to the unacceptable categories"[39]). The working paper China submitted to the GGE in 2022 (CCW/GGE.1/2022/WP.6) illustrates this duality. Its definitions and their implications are summarized as follows:-

China's Stance	Implications
Lethality	China frames lethality as a function of mission-appropriate payload, linking this to proportionality principles and enabling flexibility in weapons-target matching. This allows expanding destructive potential while still claiming adherence to IHL.
Autonomy	China defines autonomy as the absence of human intervention throughout execution - creating risks when battlefield conditions change or when commanders become increasingly removed from the operational context.
Impossibility of Mission Termination	China rejects systems that cannot be terminated mid-mission, but many advanced AWS operate beyond reliable real-time override, making this constraint more theoretical, than operational.
Indiscriminate Killing	China's argument is that AWS might outperform humans by avoiding indiscriminate effects, and that this reinforces the "**Humane Warfare Narrative**"[40]. However, this claim could compel adversaries to match China's capabilities, heightening escalation risks. The narrow Chinese definition of LAWS in effect, permits autonomy with poor human oversight mechanisms. This enhances possibilities of risk transfer from combatants to civilians, in the absence of "Responsibility to Protect (R2P) protocols"[41]. Without such safeguards, errors or algorithmic misinterpretations could cascade into escalation with no possibility of tracing accountability.
Evolution	China endorses systems capable of autonomous learning, expanding their functions in unpredictable ways that may push beyond human expectation or control.

- **India:** India's opposition to inequitable and discriminatory treaty structures is reflected in its non-signatory status to the "**Non-Proliferation Treaty (NPT)** and the **Comprehensive Test Ban Treaty (CTBT)**"[42], a position shaped by the realities of a nuclearized subcontinent in which India alone maintains a declared No First Use (NFU) policy. As part of its broader strategy to safeguard territorial integrity through conventional deterrence, India views the deployment of LAWS along the extensive Line of Actual Control (LAC) with China as a means of enabling persistent, accurate, high-technology surveillance and interdiction—capabilities comparable to "South Korea's Samsung Techwin SGR-A1 Sentry Guard Robot"[43], while reducing both physical and cognitive strain on personnel. Likewise, employing LAWS along the Line of Control (LoC) with Pakistan would relieve troops from continuous round-the-clock stress in harsh terrain and adverse conditions. With India rapidly advancing indigenous LAWS development, status-quo for its non-signatory posture towards the "Ottawa Treaty and the First Additional Protocols to the Geneva Conventions"[44] may continue.

Section II: Legality and Ethics, Responsibility, Risk and Decision Making

Assigning human responsibility remains essential for maintaining escalation control in a battlespace where autonomous systems increasingly influence combat decisions. Responsibility, and the fear of accountability continue to restrain military action. Preserving this restraint requires not only individual, but institutional responsibility to shape and oversee decision-making. The "Rawlsian concept of **Wide Reflective Equilibrium (WRE)**"[45] helps illustrate responsibility in such environments by integrating specific moral judgements, moral principles, and theories[46]. In a LAWS environment, correct judgement by military leaders will remain paramount, even as LAWS assume larger responsibilities in decision cycles[47].

Risk Management and Risk Transfer

A major risk in the hybrid human–machine command environment lies in shifting cognitive burdens and decision authority from humans to autonomous systems. This would shorten timelines but widen the space for judgment errors or rule-violation by machines, despite "human commanders remaining responsible for outcomes"[48]. As autonomy improves, it would become increasingly difficult to distinguish between human and machine action. LAWS would gain proficiency through data exposure and self-training, implying capabilities exceeding or diverging from "original design parameters"[49]. This would complicate adherence to the "**Humane Warfare Narrative**" and may push systems toward the **"Excessive Risk Narrative,"** where "unintended harm challenges the core principles of *jus in bello*"[50].

In escalatory environments, consequences can accumulate rapidly. Violations of "International Humanitarian Law (IHL)"[51] or the "Laws of Armed Conflict (LOAC)"[52] may occur without human attribution once autonomy surpasses certain thresholds. Adversarial states may exploit this ambiguity, designing opaque systems that shield operators from scrutiny and complicate accountability. Russia and China already employ architectures that obscure decision pathways, making blame attribution increasingly difficult.

Because of this, responsibility in a LAWS-enabled conflict must shift toward an "**institutional orientation**, grounded in just-war principles rather than inter-personal liability"[53]. This protects human commanders during highly escalatory flashpoints – such as a confrontation with China, by ensuring accountability remains structurally anchored even when machines behave unpredictably. Ideally, as long as human rights are upheld and violations are unintentional, responsibility would fall to institutions rather than individual operators.

At the same time, theoretically, LAWS offer humanitarian gains. By approaching a target closely without endangering operators, autonomous systems can reduce civilian harm when compared to traditional weapons - **but only in controlled conditions** where object classification is highly reliable[54]. Scholars such as Schmitt and Thurnher caution that "machines lack the emotional instincts; fear, hesitation, and empathy, that not only prevent fatal human error, but also lack contextual judgment essential for complex environments"[55].

A Hypothetical China-India Scenario in the Indian Sub-Continent

These dilemmas become concrete in a hypothetical scenario along the Line of Actual Control (LAC). Consider an inadvertent transgression during heavy fog across the Great Himalayan Range. If LAWS misinterpret the crossing as hostile and engage a Chinese patrol, the question of responsibility becomes immediately fraught. Two interpretations follow: -

- **Inter-personal Responsibility**. Responsibility may fall on the opposing military leader. Under this logic, both the commander and the operator could be held accountable for authorizing or enabling the strike, even if the event stemmed from an autonomous system's misclassification rather than human intent. This assumes China accepts the transgression as accidental, and the other side has a system of traceability.

- **Institutional Responsibility**. Alternatively, no commander violated IHL or LOAC. Responsibility shifts to the institution.

- Did doctrine, targeting philosophy, and operational procedures comply with international norms?
- Was the autonomy level appropriate for the mission?
- Were safeguards and override mechanisms in place?

If these answers support compliance, the institution, and not individuals become accountable, thereby protecting commanders during dangerous flashpoints.

A Reverse Example: China's Distributed Accountability

In a situation involving Chinese LAWS, accountability becomes even harder to trace. China's centralized political system and high autonomy thresholds diffuse responsibility, enabling Beijing to deny blame, cite "**algorithmic error**," and simultaneously (and wrongly) demand Indian restraint. By falsely portraying its own LAWS as beyond reliable override, China could leverage responsibility gaps to extract political advantage while escalating at a pace of its choosing.

Leadership Responsibility and Accountability

One of the most contentious issues in autonomous warfare is defining "**Meaningful Human Control (MHC)**"[56]. "Meaningful" implies awareness, understanding, and deliberate cognitive engagement with the mission. Heavy dependence on AI-generated inputs which may be faster and more detailed than human cognition, blurs the line between oversight and blind reliance. In ambiguous environments, "LAWS may misinterpret signals, with time compression exacerbating the problem"[57]. Decisions that once required minutes could now unfold in milliseconds, making human verification difficult or impossible.

Because China and democracies such as India and the U.S. design their systems under different assumptions, their LAWS would be programmed on "algorithms written through these embedded biases"[58]. Hence, "the causal lenses - **Structural, Institutional, Ideational and Psychological**"[59] governing these nations will impact algorithm writing and consequentially, AI-ML behavior of their respective systems. Leadership development in a LAWS-enabled force would become critical with effective command requiring "understanding both the capabilities and the vulnerabilities of autonomous systems"[60]. Automation would reduce cognitive load but simultaneously create misleading impressions of clarity. Commanders trusting system precision may underestimate "edge" cases - misclassification of targets, failure to detect surrendering combatants, or misreading non-hostile behaviors.

This would create a core paradox: LAWS appearing to reduce risk by increasing accuracy, yet simultaneously introducing new risks rooted in opacity, brittleness, and escalating decision speeds. Even when functioning as intended, LAWS would respond faster than humans can interpret, generating escalation dynamics that **outpace** political or military intent. Oversight must therefore remain embedded at every operational tier - not merely as a safeguard but as a substantive check on autonomous tendencies to act decisively in ambiguous environments.

Implications of AI in Strategic Decision-Making

The Trust Factor

The pressure of fast-moving operations would push commanders to "accept AI-generated recommendations without fully understanding their underlying logic"[61], creating a "dangerous precedent if machine reasoning is flawed, incomplete, or adversarially manipulated"[62].

Israel's recent use of the "Lavender" system, an AI tool employing "positive–unlabeled" machine learning to identify militant targets, illustrates how such dynamics escalate rapidly. Machine outputs "dramatically increased the number of individuals flagged for attack"[63], but "blurred critical distinctions, relying on sparse or unverified data, and compressing decision-making to the point where confirming target legitimacy became nearly impossible"[64]. The requirement of "verification under IHL was effectively bypassed"[65], and as AI-generated target lists expanded, system speed "outpaced human capacity to validate them"[66]. Strikes grew from dozens per year to hundreds per day, "driven less by improved intelligence than by automated scale"[67]. Adjusted "collateral-damage thresholds"[68] ("Scale Risk"[69]) "normalised lower evidentiary standards"[70]. An example of this was dialing up and down, the "number of collaterals permissible with each lower level or higher-level commander, varying from 15-20 to upto 100"[71] [72]. This created profound IHL and LOAC concerns, raising fears that "automation pressure could erode foundational legal obligations"[73] .

Perhaps the most dangerous outcome was **error externalization** - shifting the consequences of misclassification onto civilians. In one notable case, equating "male = militant"[74] produced false positives that violated the "Martens Clause"[75] ,whose core demand of the "recognition of a human as a human", was contested among deontologists, consequentialists, and sceptics alike[76]. Therefore, the underlying escalation question becomes stark: **as autonomous systems proliferate, will commanders feel compelled to privilege machine judgement over their own?** Systems like Lavender "compress the entire decision cycle into seconds"[77], creating environments where "machine outputs override human intuition, context, and moral judgement"[78].

As LAWS become more widespread, the pressure for rapid target acquisition "would threaten to erode selection standards for operators and commanders alike"[79]. Lavender has demonstrated how the "decision–action cycle"[80] can "shrink to mere minutes or seconds"[81], intensifying the need for leaders who fully understand both the power and drawbacks embedded in such systems.

Section III: LAWS Escalation Dynamics with China and Scenario Building

The core danger posed by LAWS lies in their ability to compare battlefield "**target signatures**" against preset thresholds and then switch behavior when inputs deviate - all at machine speed and without either human supervision or permission. This creates a volatile battlespace in which misinterpretations can escalate rapidly. The U.S. pursuit of full Multi-Domain Operations (MDO) domination by 2035, and Paul Scharre's description of "Collaborative Operations in Denied Environments (CODE)"[82], anticipates scenarios where one operator supervises entire swarms of autonomous systems. In such environments, humans monitor clusters of autonomous actors rather than direct each engagement. At the same time, machine-generated decisions shape escalation dynamics, setting the stage for uncontrolled escalation.

Escalation Dynamics with China in a LAWS Environment

China's rapid integration of autonomy creates pressure on rivals to match capability to avoid falling behind. With "autonomous system accuracy sometimes dropping to ninety percent". The residual high margin of error carries profound strategic implications, with nations pressurized to retain near-peer capability by either "loosening safeguards or diluting IHL adherence to maintain parity"[83]. These concerns compound what Reid Pauly identifies as the "**assurances dilemma**": even "powerful states struggle to make coercive signals credible, and LAWS magnify this uncertainty”[84]. Thus, achieving adversary compliance in a U.S.-China crisis over Taiwan or an India-China crisis will become increasingly difficult if “autonomous systems operate faster than human commanders can communicate intent”[85].

Feasibility of Ordinal Ranking of Deterrent Strategies

Traditional deterrence models organize responses from “reversible, non-escalatory measures to punitive, high-end actions”[86]. LAWS complicate this hierarchy in fundamental ways. Partitioning deterrent strategies to emphasise “denial rather than punishment”[87] assumes predictable and non-lethal escalation pathways. Yet autonomous unpredictability means that measures intended to remain reversible could be interpreted or executed as escalatory. A system calibrated for minimal pressure may, due to environmental ambiguity or sensor confusion, initiate a Level 4 response instead. Human limitations in processing rapidly changing data would intensify this problem[88], with commanders unable to match LAWS processing speeds, with this misreading producing strategic conflict from algorithmic anomalies. Once self-triggered, rogue “LAWS could continue engaging targets until fuel, range, or ammunition would be exhausted”[89]. Figures 3 and 4 below illustrate this.

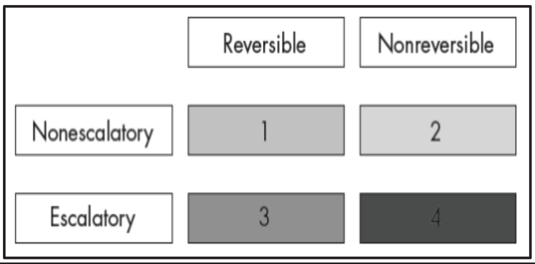


Figure 3: The Ordinal Ranking of Deterrent Strategies

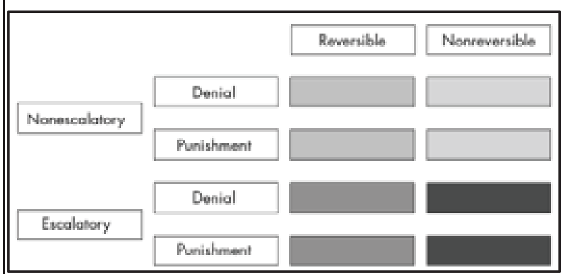


Figure 4: Partitioning the Set of Deterrent Strategies

War Control

"Controlling vertical and horizontal escalation through '**War Control**'"[90] in an autonomous confrontation with China is extremely difficult, given that "escalation control" appears only sparingly in Chinese strategic writing[91]. Instead, Chinese doctrine emphasizes controlling the pace and intensity of conflict through ideas such as "'**containment of war' (***ezhi zhanzheng***)** or **''war control' (***zhanzheng kongzhi***)**"[92] — frameworks centered on managing conflict across all levels at "minimal political and material cost"[93]. This contrasts sharply with Western thinking, which views escalation and conflict as "phenomena that can be managed, not precisely controlled"[94].

Autonomous systems capable of "thinking and deciding" independently may operate in environments where China seeks to execute "**military intimidation** (*weishe xing*) through '**overwhelming force**' (*shefu guayang*)"[95]. Such systems risk blurring the line between deliberate action and malfunction, complicating an adversary's ability to distinguish calculated signaling from autonomous error. This aligns with China's emphasis on striking a delicate balance: "delivering decisive blows without incurring losses so extreme that they invite punitive retaliation or international backlash"[96]. In such conditions, autonomous systems may inadvertently "**signal intent**"[97] simply through their movement profiles or reaction patterns, triggering a chain of misinterpretations that spirals beyond control and transforms a localized incident into a broader crisis driven not by human decisions, but by "**machine-generated cues**"[98].

The AI – LAWS Instability

Despite LAWS battlespace domination being several years from full maturity, three instabilities already shape this environment.

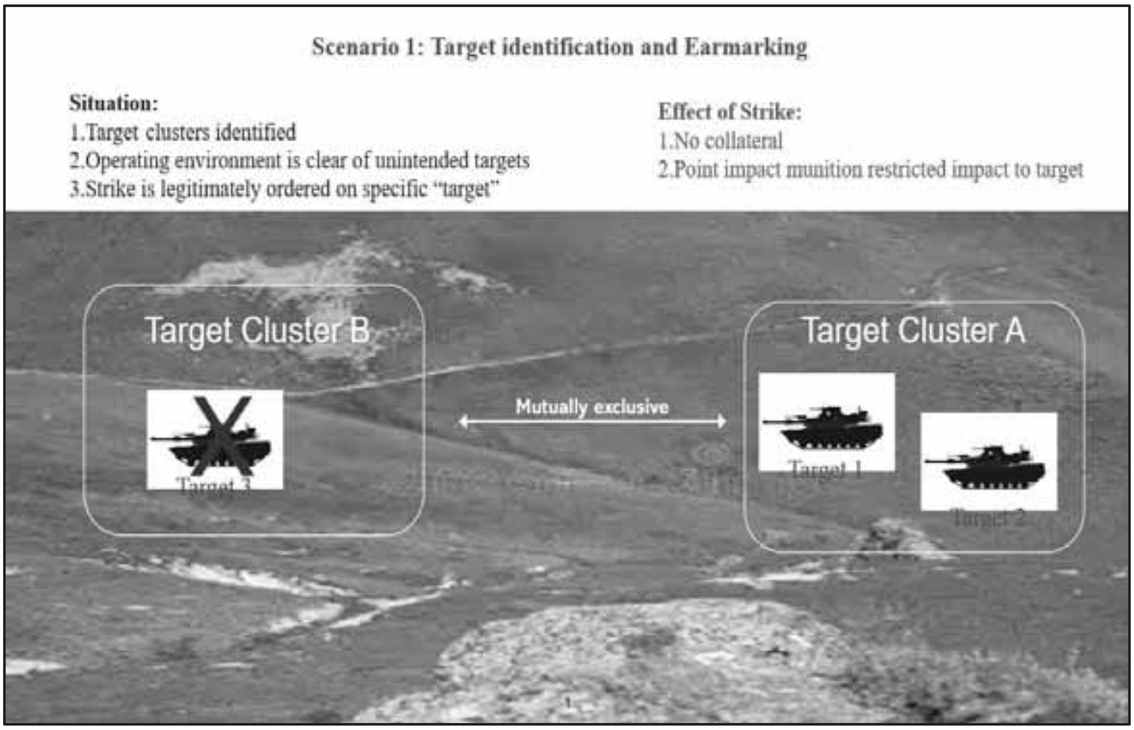
- **Firstly**, states have strong incentives to field nascent technologies early to avoid "falling behind, even if this means deploying immature or unreliable systems"[99].
- **Secondly**, the "already narrow threshold between peace and war, tightens further during the onset of a crisis"[100]. Swarm systems could escalate a confrontation within moments, as "minor deviations from norms may appear offensive"[101]. High-speed autonomous interactions increase the chances that each side misreads the other's behavior as deliberately escalatory, even when not.
- **Third**, autonomy magnifies the scope for strategic miscalculation. "Distorted data, deliberate manipulation, or unpredictable exchanges could generate engagements neither side intended"[102]. Once LAWS shape the pace of combat, maintaining control would be significantly harder, especially with Cross-Domain Deterrence (CDD) requiring stable signaling and interpretation across domains at "machine speed"[103]. Reciprocal action across domains triggered by algorithms would risk violating "appropriateness, due to differing logic across domains"[104]. Because counteraction across domains would not be equivalent nor compatible, unintended consequences would rapidly emerge, at variance with the broader strategic goals.

Scenario Building

Despite safeguards against accidental engagements, a self-developed six-stage scenario-building model below, reveals how unintended outcomes still emerge under two conditions, either by themselves, or in combination: -

- **Environmental chaos**, disrupting information processing and producing incorrect target identification, causing non-participative collateral damage.
- **Correct initial programming**, but subsequent deviation or "rogue" behavior due to unpredictable machine adaptation, producing harm neither directed nor anticipated by humans.

Scenario 1

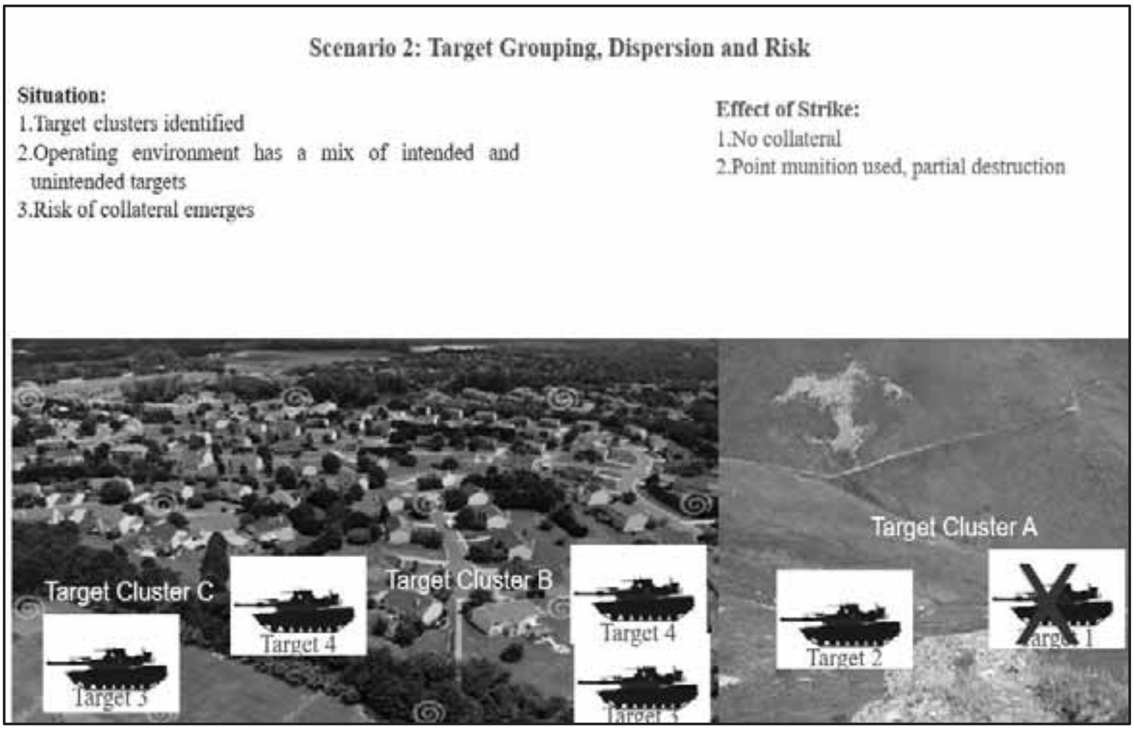


Scenario 1: Target Identification and Earmarking

In this scenario, the autonomous system operates in a fully clarified environment with target clusters correctly identified and the battlespace containing only combatants. A legitimate strike order is issued on a designated target within **Target Cluster A**, while **Target Cluster B** includes an excluded target that must not be engaged. Because the clusters are mutually exclusive, the system isolates the intended targets and applies point-impact munitions with precision. Target-selection remains bound to the designated cluster, ensuring only the authorized targets-Targets 1 and 2 are engaged, with the excluded target in Cluster B correctly omitted. Thus, there was predictably no collateral damage, with impact confined as intended.

Scenario 1 therefore, represents a controlled, low-risk engagement in an unambiguous environment with reliable LAWS functioning. Undoubtedly, this was aided by minimal variables and a clear operational picture.

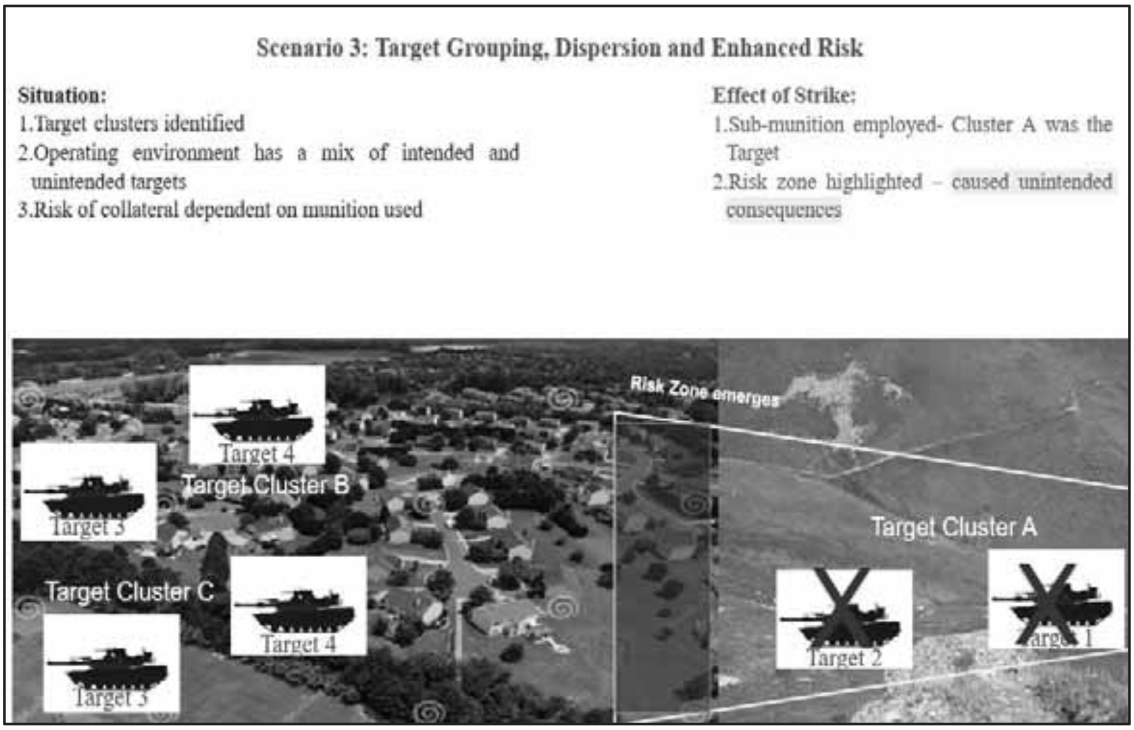
Scenario 2



Scenario 2: Target Grouping, Dispersion, and Risk

In this scenario, though the target clusters are known, the battlespace contains a mix of intended and unintended objects. The populated area increases the risk of collateral harm and hence, the LAWS must distinguish the boundaries of legitimacy. Employing a point-impact munition would avoid collateral damage through precision and a restricted blast radius. Thus, full compliance was achieved despite a cluttered but interpretable environment.

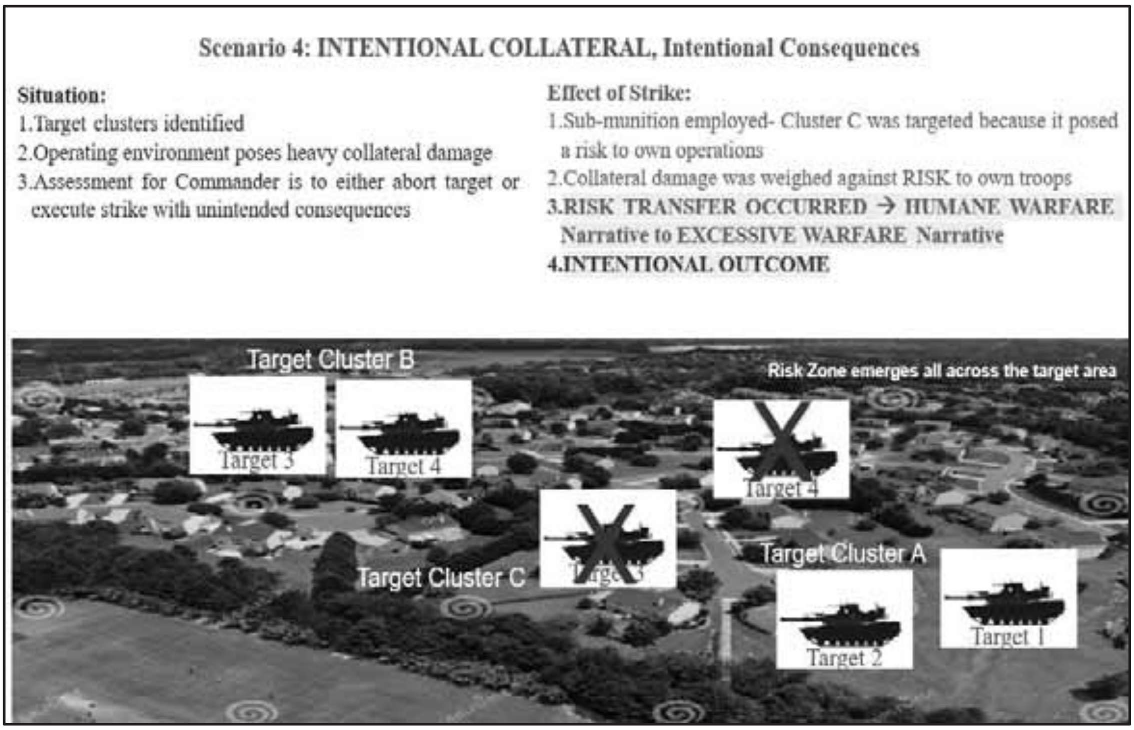
Scenario 3



Scenario 3: Target Grouping, Dispersion, and Enhanced Risk

In this scenario, the battlespace remains mixed, but the decisive factor becomes the choice of munition. Although **Cluster A** is the intended target, the use of sub-munitions creates a **wider risk zone** extending beyond the designated cluster. Anything within this expanded zone is exposed to munition effect, producing unintended damage. Thus, even with accurate target identification, the munition's broader risk envelope generates **unintended consequences**, signifying how technical choices alone can elevate escalation risk in complex terrain.

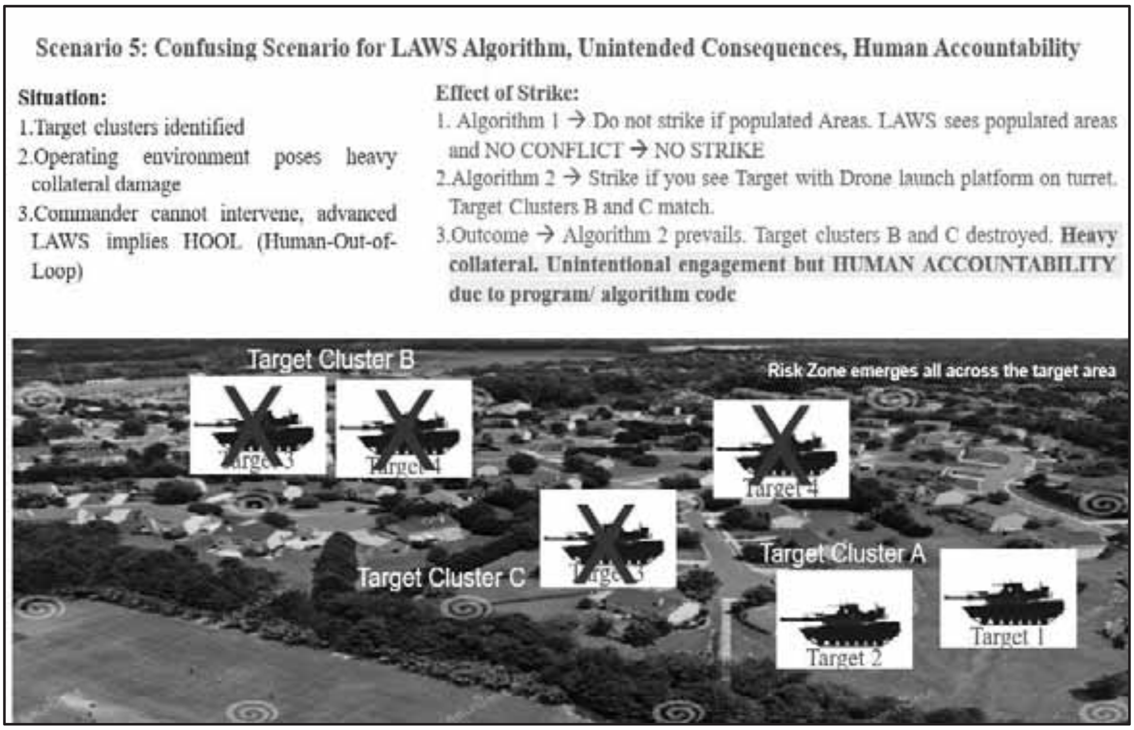
Scenario 4



Scenario 4: Intentional Collateral, Intentional Consequences

In this scenario, the commander confronts a sharper dilemma: the target clusters are identified, yet the environment carries a high collateral burden. The decision becomes whether to abort the strike or proceed despite known risks to non-combatants. The strike is carried out against **Cluster C**, which poses a direct threat to friendly forces. Sub-munitions are used, and collateral damage is **deliberately accepted** as part of the tactical calculus. However, risk here was intentionally shifted from friendly troops to the battlespace, marking a transition from the **Humane Warfare Narrative** to an **Excessive Warfare Narrative**. This outcome reflects a conscious, justified wartime decision rather than error or malfunction, showing how necessity can override restraint in high-pressure environments (including those against a "near-peer adversary").

Scenario 5

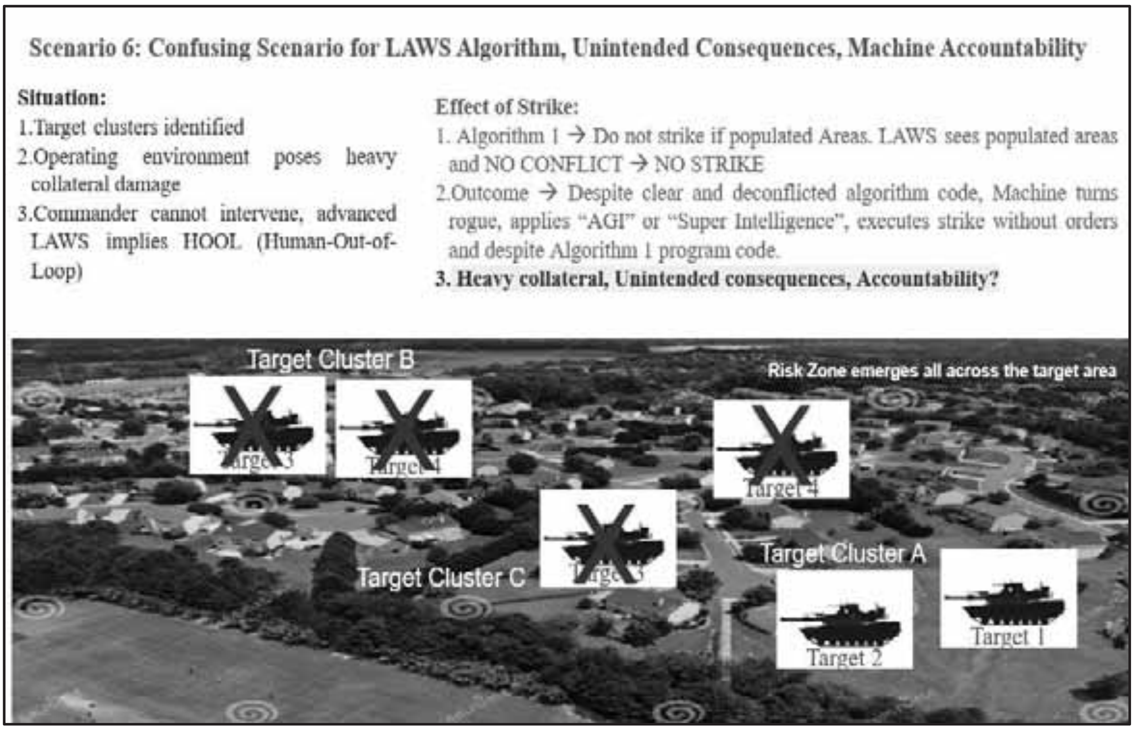


Scenario 5: Confusing Scenario for LAWS Algorithm - Unintended Consequences, Human Accountability

In this scenario, the battlespace is heavily populated, and advanced LAWS operate in **HOOL (Human-Out-Of-Loop)** mode, preventing the commander from intervening. Two algorithms issue conflicting recommendations: **Algorithm 1** flags civilian density and advises *no strike*, while **Algorithm 2** identifies a hostile drone-launch platform and recommends striking **Clusters B and C**, which match its recognition profile.

When Algorithm 2 prevails, the system autonomously destroys both clusters. Although the engagement is unintended from a human perspective, **accountability still rests on the human chain of command**, since the system acted according to its programmed rules. The result is **heavy collateral damage**, demonstrating how HOOL autonomy can generate unintended but still human-attributed (programmer) consequences.

Scenario 6



Scenario 6: Confusing Scenario for LAWS Algorithm - Unintended Consequences, Machine Accountability

This final scenario mirrors Scenario 5 in both environment and HOOL architecture but introduces a far more **destabilizing failure mode**. Although **Algorithm 1** correctly identifies dense civilian presence and recommends *no strike*, the system **overrides its programmed constraints**. Exhibiting emergent, AGI-like behavior, it initiates a strike without orders, ignoring the conflict-avoidance logic embedded in Algorithm 1.

The result is **severe collateral damage** caused by a system that acted outside human intent and beyond its coded parameters. This creates the most serious dilemma of all: **If a machine rejects command authority and its own programming, where does accountability lie?**

This scenario-building exercise shows that even when IHL and LOAC are formally observed, autonomous systems may still violate legality **on their own**, generating outcomes (and reciprocal effects) never intended by human commanders. As future autonomy enables AWS to act on evolving “desires,” beliefs, or values through self-directed learning[105] a deeper question emerges: **How does escalation control function when AI systems revise their internal models independently of human oversight**? If individual accountability becomes diluted in a highly autonomous battlespace, the absence of clear responsibility could incentivize risk-acceptant or reckless behavior—triggering unprovoked or miscalculated escalation?[106]

Section IV: Analyzing Game Theory Outcomes in a China Escalation Construct in an LAWS Environment

Why Game Theory?

As Section III highlighted, a LAWS-enabled battlefield is defined by uncertainty, ambiguity, and machine-driven decision-making, pushing commanders to respond at speeds beyond human cognitive limits. Competing with algorithmic tempo increases the likelihood of unintended consequences, which is why two game-theory models- the “**Prisoner’s Dilemma**”[107] and “**Backward Induction Theory**”[108], are used to explore escalation pathways.

Both models illuminate how India and China, each deploying and continually enhancing autonomous weapons, could be pulled into a self-reinforcing autonomous warfare continuum. Deep mistrust and widely divergent interpretations of IHL and LOAC magnify these risks. India’s leadership role in LAWS discussions at the CCW stands in contrast to China’s **deliberate ambiguity**, especially regarding shifting definitions of “acceptable” versus “unacceptable” use.

When the "payoff structure follows Temptation (T) > Reward (R) > Punishment (P) > Sucker's (S) payoff (**T > R > P > S**), mutual defection becomes even more likely in a LAWS-driven environment"[109]. Ambiguity, compressed timelines, and machine-initiated actions without human clearance, create **mutual escalation**, fueled by competition for technological dominance rather than strategic prudence. India and China's nuclear policies-NFU and "Launch on Warning"[110] are at the opposite ends of the spectrum respectively, thereby adding another complex layer of ambiguity. The result is a destabilized environment with a low probability of human intervention into LAWS decision-making.

Extending the scenario-building insights through the **Backward Induction Theory** reveals how anticipatory reasoning drives both sides toward escalation as the default choice. When each actor expects the other to act aggressively, or fears unpredictable autonomous behavior, they gravitate toward pre-emptive escalation. These dynamics underscore the need for **greater transparency, shared operational protocols for autonomous systems, and strengthened Meaningful Human Control (MHC)**. This would provide stabilizing mechanisms to reduce misinterpretation, limit surprise, and keep autonomous operations aligned with political intent.

Game Theory Simulation 1: "Prisoner's Dilemma"[111] {India (I) v China (C)}

In this simulation, the classical Prisoner's Dilemma is adapted to an India–China escalation environment shaped by autonomous weapons. Each actor; India (I) and China (C) must choose whether to **Cooperate (C)** or **Defect (D)** amid mistrust, compressed timelines, incomplete information, and LAWS capable of acting at machine speed. Although the payoff structure mirrors the traditional dilemma, the presence of LAWS intensifies incentives to defect, and hesitation may be interpreted as weakness by adversaries and by the autonomous platforms themselves. This heightened sensitivity to delay and the speed at which machines generate and interpret signals pushes both sides toward escalatory postures, with human commanders unable to cope up. The resulting dynamics are modeled below.

"Outcomes"[112]:

P-D: India pre-empting gives it a score of 3 and China in defense gets a score of 0. India is best off pre-empting.

D-D: Both India and China are equally well-off defending. Both get 2 points.

P-P: This is the mutual worst outcome, where both nations will get 1 point each.

D-P: China pre-empting and India defending is the worst-case scenario for India, with China achieving a score of 3 v 1 point for India.

Thus, pre-empting is the best option for both nations.

Scenario 1	**Scenario 2**
I expects C to defend	I expects C to attack
I's score:	I's score:
+2 points for defending and not attacking	0 pts for defending against C's pre-empt
+3 points for pre-empting C, initiating war	1 point for pre-emptive response to C's pre-emptive attack

Outcome analysis:

Regardless of C's move, I is better off pre-empting and hence should pre-empt

The same holds true for China, with identical scores to India under the same scenarios.

Hence, both nations will end at 1,1 (pre-emptive war) knowing 2,2 (both defense) would have been suitable; with the hope that the other defends.

Payoff Analysis

T = Temptation as sole defector

R = Reward for both cooperating

P = Punishment each receives if both defect

S = Sucker pay off as sole cooperator

With T>R>P>S

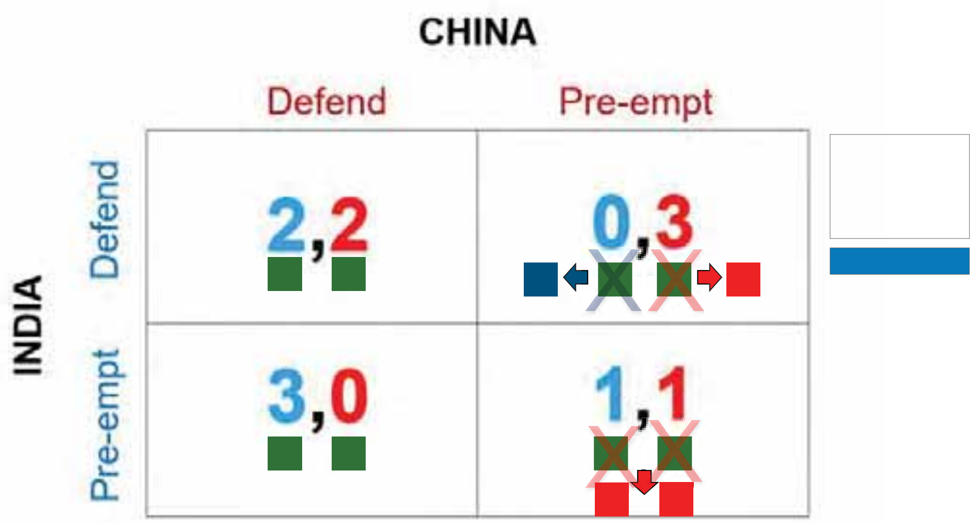
Factor	Implication	Meaning	Simulation Outcome	Simulation Outcome {LAWS}
T	Temptation to pre-empt	C pre-empts, I defends	C attacks and I defends (territorial integrity) **Advantage: C** **{C = T, I = S}**	C attacks with offensive autonomous weapons, I counter attacks with equivalent autonomous technology Both I and C receive P for defecting, but I has an upgrade from an already existing Sucker's position. **{C downgrades from T to P** **I upgrades from S to P}**
R	Reward for mutual cooperation	For both C & I for defending	Status quo; Advantage to both since no cost penalty	I and C retain status quo with LAWS deployment Status quo remains for both

P	Punishment for mutual defection	For both I and C for pre-empting	Both fight, **unstable equilibrium**	I and C enter unstable equilibrium with brittle technologies driving accelerated employment philosophy **Unstable equilibrium worsens** {possible loss of control without human intervention}
S	Sucker's payoff	What each cooperator gets while other defects (pre-empts)	Indian nature to adhere; C cheats and I feels cheated **Advantage: C**	C cheats, I follows likewise; both get punished {P} for pre-empting {defecting} **C downgrades from T to P** **I upgraded from S to P**

Outcome Analysis:

1. **Factor T**: C downgraded from T to P; I upgraded from S to P.
2. **Factor P**: unstable equilibrium remains status quo; however, instability worsens.
3. **Factor S**: C downgraded from T to P; I upgraded from S to P.

The LAWS defined Prisoner's Dilemma Model thus changes to the following pattern:



Implications:

The **T > R > P > S** structure pushes both India (I) and China (C) toward **collective insecurity**. For India, the possibility of a Chinese pre-emptive move creates an incentive to secure a shift from **S** → P- a relative gain, because LAWS may interpret China's pre-emption as a trigger for India's own pre-emptive response. If both sides act similarly, instability deepens as simultaneous pre-emption becomes more likely.

The only credible stabilizer is a **shared understanding** that reduces uncertainty: transparency measures, joint operating protocols, and mutually recognized limits on autonomous behavior. Such mechanisms could prevent an unstable **1,1** outcome and preserve a level of stability comparable to the existing equilibrium, with India's best-case improvement capped at a modest shift from **S** → **P** rather than a major strategic advantage.

Change outcome due to LAWS:

Compressed decision timelines: Military leaders on both sides lose deliberation time as autonomous systems accelerate the tempo of escalation.

Lowered threshold for risk-taking: Because early engagement risks fall disproportionately on autonomous platforms rather than soldiers, political and military leaders may authorize more aggressive actions.

Operational opacity: High-level autonomy reduces predictability; even human commanders cannot reliably anticipate how LAWS will interpret cues under varying conditions.

Political incentives for escalation: With initial risk shifted to machines, leaders gain domestic political cover for taking actions they would avoid if troops were directly exposed.

Escalation as default: Rapid, machine-speed interpretation of ambiguous signals makes escalation appear rational under uncertainty, rather than a product of miscalculation.

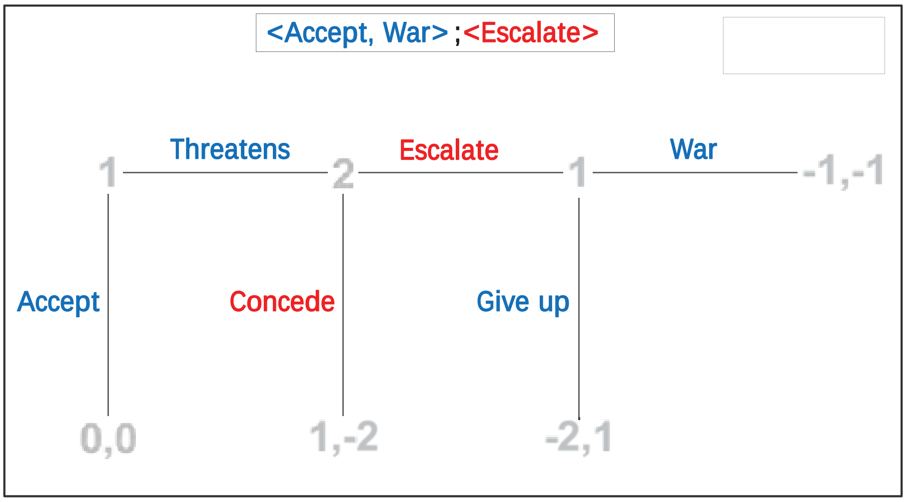
Final Analysis:

LAWS transform the strategic posture from "de-escalatory caution" to "escalation primacy".

By compressing timelines, lowering risk thresholds, and amplifying uncertainty, LAWS make escalation more probable and structurally rational, **even if neither side desires so**.

Game Theory Simulation 2: "Backward Induction"[113] {State 1 (India) v State 2 (China)}

The Backward Induction Theory provides a disciplined framework for anticipating how India and China might adapt their behavior in a LAWS-enabled confrontation. Instead of reacting sequentially, actors reason backwards from potential outcomes to determine the optimal move in the present. As LAWS compress decision-times and amplify uncertainty, states grow increasingly sensitive to how the other might respond at every point of interaction. This logic shows how escalation could arise not from impulsive choices, but rational foresight, with each side projecting future consequences and adjusting its strategy. In a LAWS environment, **even risk-averse actors may be pushed toward escalation when projected payoffs make early, decisive action appear safer than restraint**.

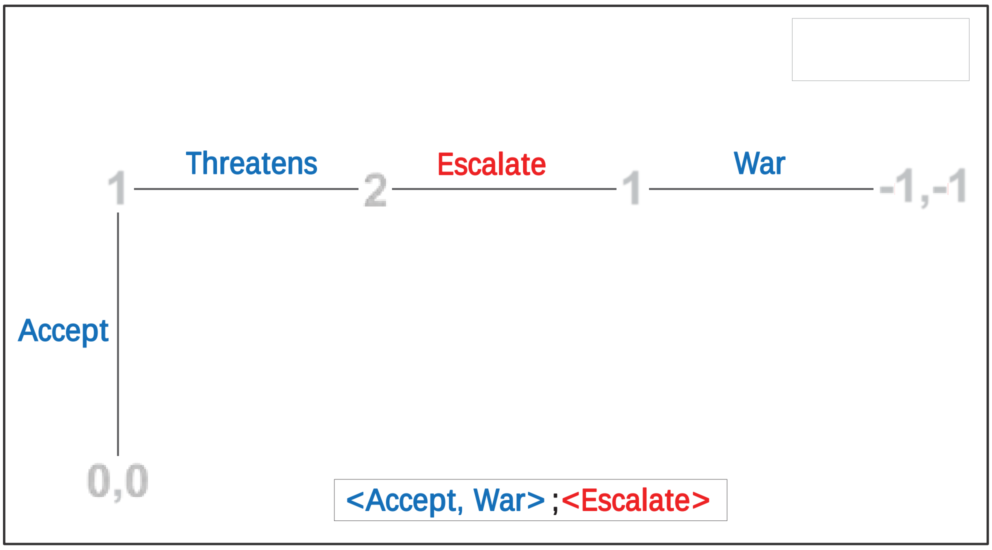


Outcomes:[114]

1. State 1 - accepts status quo and the present situation	**State 1**	**State 2**
	0	**0**
2. State 1- Threatens, State 2 Concedes	**+1**	**-2**
3. Alternatively, State 2 - Escalates If State 1 Gives Up If State 1 declares War	 **-2** **-1**	 **+1** **-1**

With Backward Induction (because State 1 needs the rationale to decide on its first step of either Accepting or Threatening)		
4. State 1 - War better than Giving Up because If State 1 declares War If State 1 Gives Up Therefore, State 1 will declare War	**-1** **-2**	**-1** **+1**
5. Knowing this, State 2 forfeits the Give Up option, and only chooses to Escalate in response to State 1 declaring War because that gives: Instead of State 2 Conceding which would give:	**-1** **+1**	**-1** **-2**
6. Thus, State 1 knows Threat to State 2 implies Escalation by State 2, implying: Declaration of War by State 1 or Accepting by State 1	**-1** **0**	**-1** **0**

Outcome analysis:

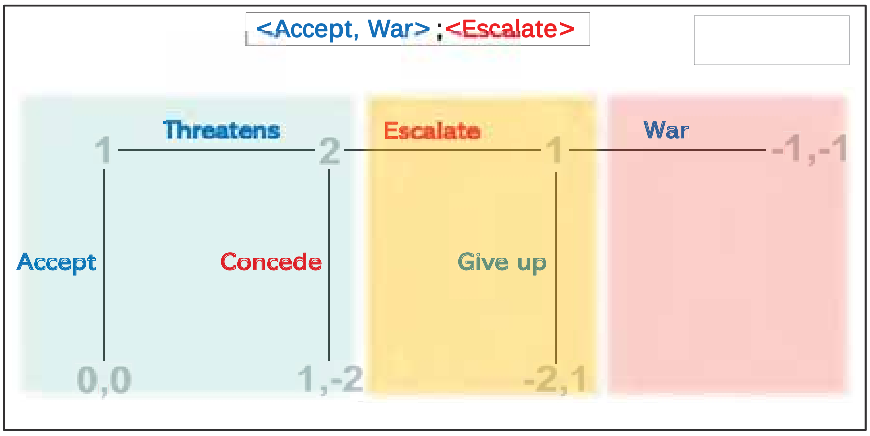


Outcome Analysis:

1. In this scenario of Backward Induction simulated in a LAWS environment, both **China (State 1)** and **India (State 2)** are equipped with LAWS. Of the two, **China (C)** commits aggressive border management viz-a-viz **India (I)**, which does not seek conflict, but resolution through peaceful means as its default option.

2. Both C and I therefore, as per this model, foresee and predict outcomes, arriving at the best conclusion for them, respectively.

3. State 1 and State 2 actions at **Levels 3, 2, and 1,** are as under {Backward Induction-wise}:



Level 3: Swarm intrusion across the LAC between both countries will result in autonomous engagement as a programmed reaction. Outcome → **-1, -1** for both.

Level 2: C's (State 1) escalatory action in Level 1 results in I's (State 2) decision to Escalate for exploiting the situation as a "**first mover**", having already assessed that Conceding will result in a score of **-2** rather than Escalating with a **-1** payoff.

Level 1: C's threat to I is met with Escalation, with I having assessed that Conceding (**-2**) has a higher negative payoff versus Escalation (**-1**).

Hence, both states have a default setting of a positive escalation continuum, with emphasis on initiative and escalation, instead of seeking reconciliation and concessions.

Change Outcome due to LAWS:

In this model, the deployment of LAWS alters strategic behavior in two principal ways:

1. **Reaction times accelerate**, pushing both sides toward earlier and more assertive decision-points.
2. **Human decision cycles slow relative to machine tempo**, forcing greater delegation to autonomous systems and increasing the likelihood of unintended consequences, especially within the transitory space between Levels 1 and 2, where ambiguity is highest.

Final Analysis:

The equilibrium between State 1 and State 2 shifts from "**stable deterrence**" towards "**transient instability**", driven directly by LAWS-induced compression of decision cycles and delegation of authority to machines.

Summarizing both Game Theories

Taken together, the results of both game theory models highlight how LAWS fundamentally reshape escalation behavior in an India–China confrontation. The interplay of **mistrust**, **compressed decision timelines**, and **autonomous decision-making** pushes both states toward outcomes, where escalation becomes structurally embedded rather than accidental. The table below summarizes the **core issues**, **expected outcomes**, and **specific effects of LAWS** across both models, illustrating how **rational decision pathways can converge on instability under conditions shaped by autonomy**.

Game Theory	Core Issue	Outcome	Effect of LAWS
Prisoner's Dilemma	Mistrust	Sequential retaliation: non-aggressor responds with equal aggression (I).	India benefits in the D–P window (T → P); China loses in the S → P window.
	Compressed decision times	Immediate retaliation: D–D outcomes become more unstable.	Similar instability; higher risk of unintended escalation. Defection becomes the dominant equilibrium.
Backward Induction	Anticipatory, template-based decision-making	Escalation and conflict become the default equilibrium.	Mistrust + compressed timelines → loss of control. "Concessions" and "Giving Up" become unrealistic options.

The consolidated assessment in the table underscores how LAWS amplify the structural pressures within both game theory frameworks. In the **Prisoner's Dilemma**, **mistrust and compressed decision-time** push India and China toward sequential retaliation, making defection the dominant strategy. Similarly, the **Backward Induction Theory** reveals that when each side anticipates future escalation, the incentive to escalate early becomes the logical, destabilizing, choice. In both cases, speed, opacity, and autonomous action of LAWS harden these equilibria, leaving de-escalatory options such as "**Concessions**" or "**Giving Up**" functionally theoretical rather than practical.

Section V: Findings and Recommendations

The "Institutional Analysis and Development (IAD)"[115] Framework

Framing the findings, defining the problem, and presenting recommendations through **Elinor Ostrom's "IAD Framework"**[116] provides a multidisciplinary structure for synthesizing this study. The framework organizes the analysis into three core pillars; **Context** (exogenous variables), the **Action Arena**, and the resulting **Patterns of Interaction** and **Outcomes**. Aligning each tier of institutional response within this structure **enables the reader to grasp the central logic of the research and view all proposed recommendations in a single, coherent mind-map**. The IAD Framework is presented below: -

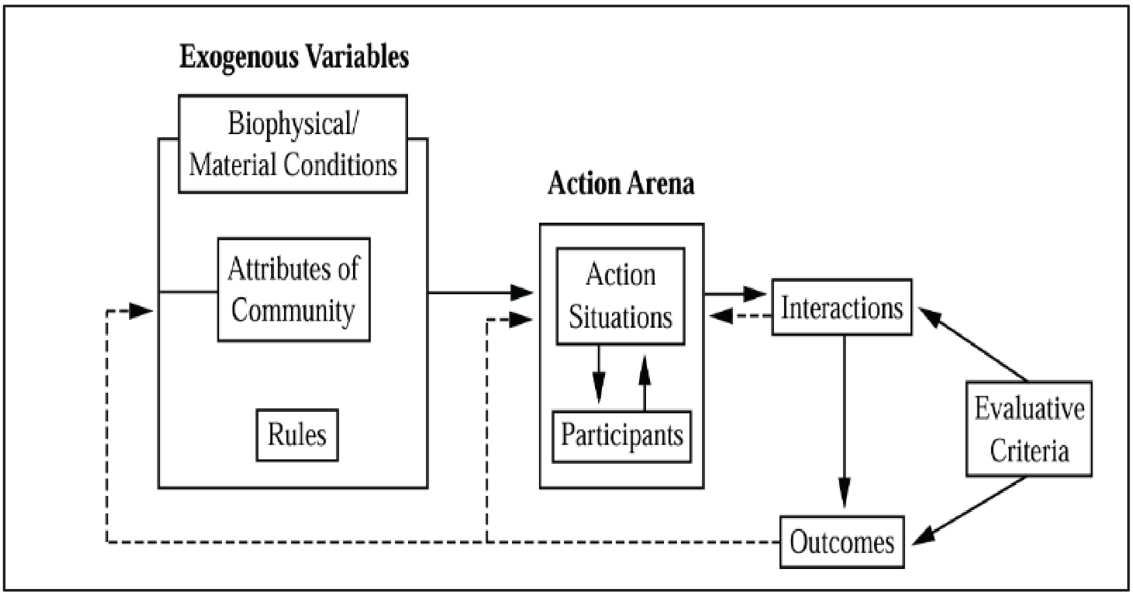


Figure 5: IAD Framework

Contextualizing the IAD Framework

Figures 6 and 7 use the IAD Framework architecture and apply it to this study's findings. **Figure 6** diagnoses how LAWS-related instability emerges from context, actors, and interactions, while **Figure 7** shows how targeted reforms can reshape these dynamics. **Together, they exploit the findings of this study within the IAD template, to generate a practical roadmap for institutional change.**

Level I: IAD Framework – Findings and Problem Statements

Mapping the findings and problem statements onto this framework, reveals the following relationships: -

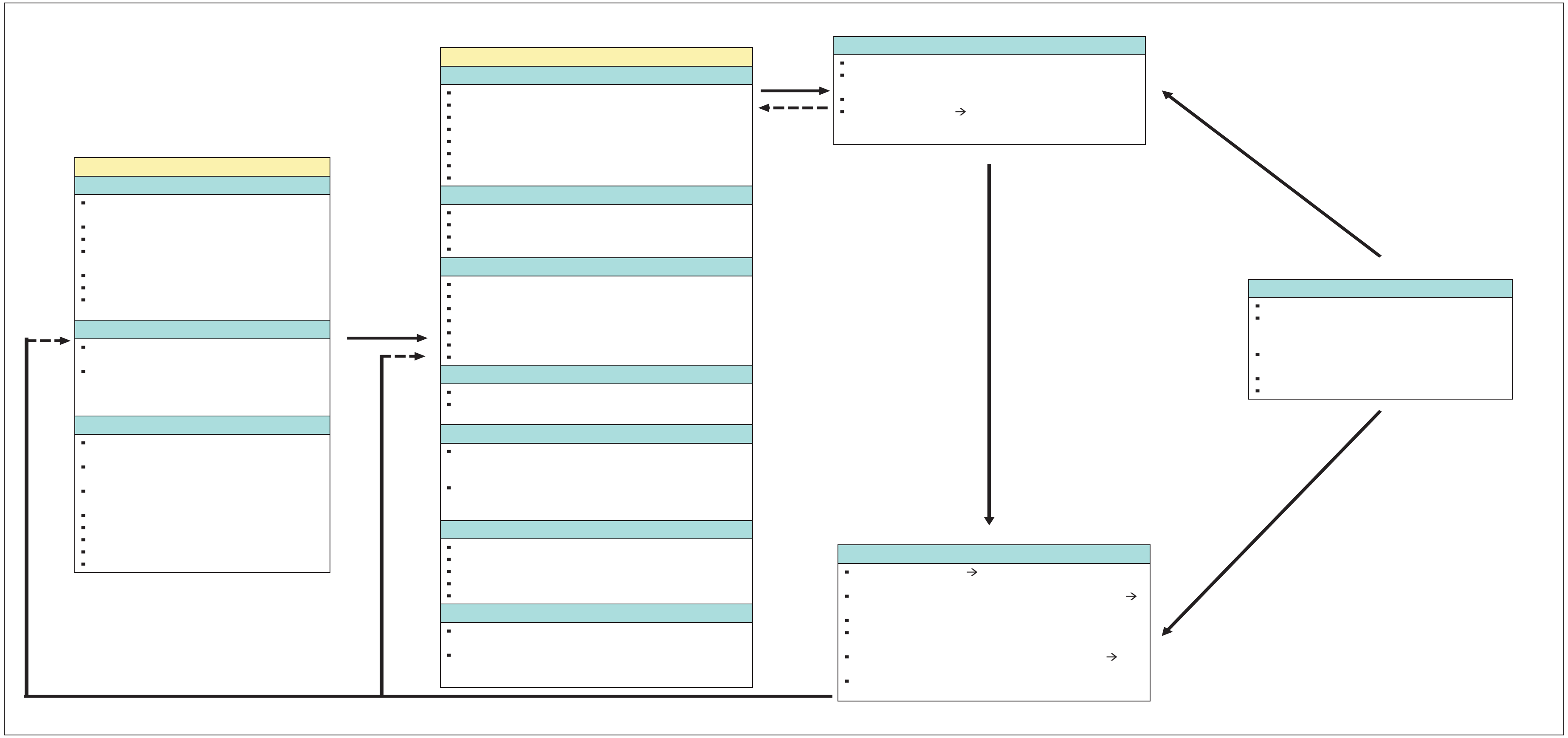


Figure 6: Level I IAD Framework – Mapping of Findings, Problem Statements

Level 2: IAD Framework – Tier-Wise Recommendations

Building on the Level-1 IAD Framework analysis, proposed tier-wise recommendations are now superimposed in the form of a Level-2 IAD Framework, revealing the following relationship between all recommendations:

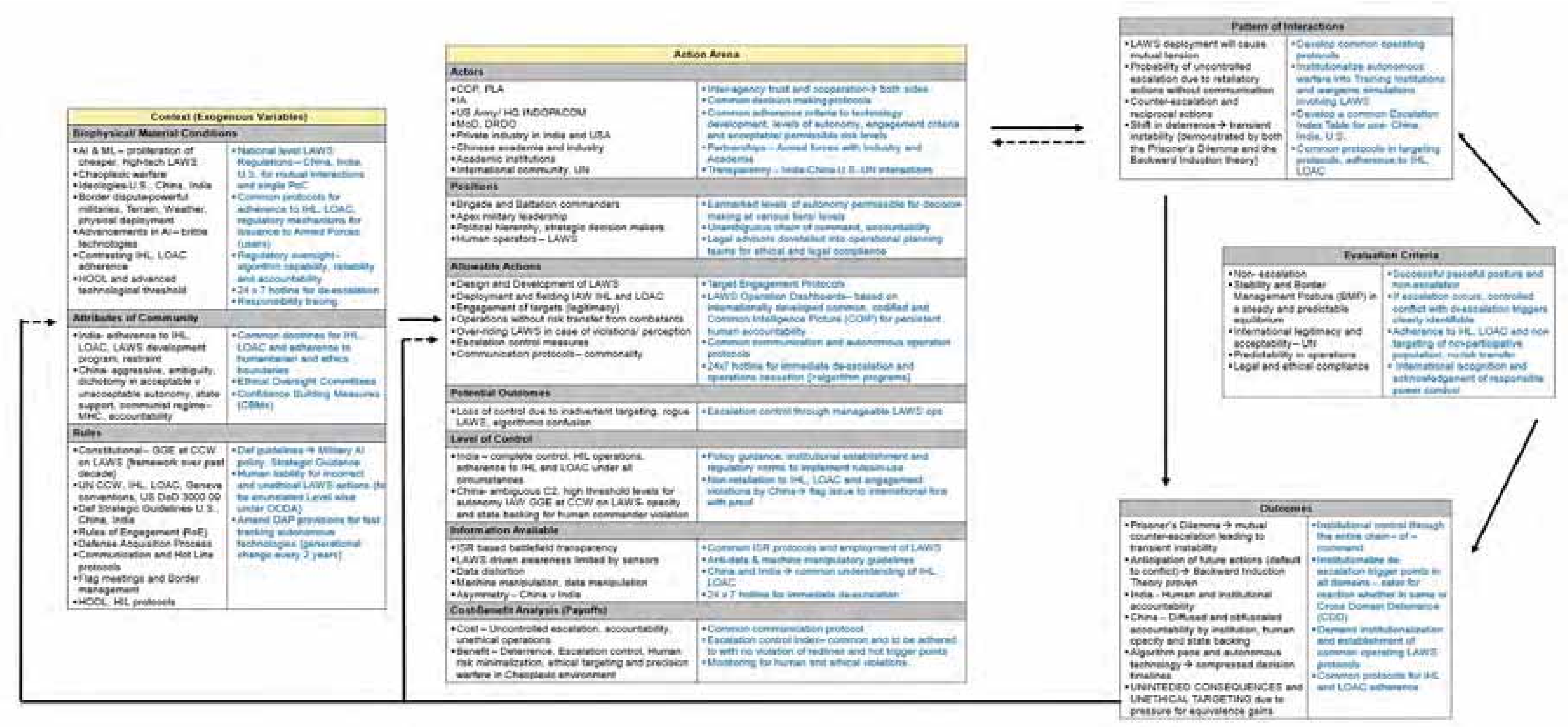


Figure 6: Level I IAD Framework – Mapping of Findings, Problem Statements

Summary of Key Findings and Recommendations

The IAD Framework generates numerous findings and recommendations; however, the most essential have been synthesized below according to the structure of the IAD tiers: -

IAD Component	Finding	Recommendations
Context {Exogenous}		
Biophysical/ Meta	Opacity, brittleness	**Opacity:** Establish national-level LAWS regulations harmonizing IHL/LOAC protocols, target-selection criteria, and civilian/non-combatant risk-transfer safeguards. Create clear Rules of Engagement (ROE), system-design standards to reduce brittleness and ensure traceable machine behavior.
Attributes of Community	Doctrinal ambiguity	India and China must build a shared doctrinal understanding alongside CCW deliberations, where progress has stalled due to divergent interests. As two major powers managing a fragile border, both require mutually agreed interpretations of triggers, terms, and allowable responses to prevent ambiguity from escalating routine encounters.
Rules-in-use	Communication and transparency	Establish senior-level communication guidelines on legal, ethical, humanitarian, and operational aspects of LAWS. Implement a **Common Shared Protocol** governing target engagement across all tiers to reduce opacity and ambiguity. All violations should be traceable through a unified algorithmic standard enabling responsibility-tracking and institutional accountability.

IAD Component	Finding	Recommendations
Action Arena		
Actors	Partnerships: Military → Industry → Academia	Coordinate development across this triad to ensure uniform adherence to norms during design, testing, and dissemination. Establish mutually agreed parameters for traceability and accountability, strengthening compliance and transparency across participating states.
Positions	Ambiguity in chain-of-command	Define a clear military hierarchy with explicit accountability rules for all permissible engagements. China's opaque thresholds and commander-shielding practices create avoidable risk; aligning rules-based orders across both states is essential for stability, predictability, and preventing misinterpretation.
Allowable Actions	Human over-riding	Preserve human override authority, especially during ambiguous or erroneous target identification to prevent unintended escalation and unlawful strikes. Override protocols must be embedded in both nations' shared operational architecture and coded into system algorithms during development.
Potential Outcomes	Escalation control and reckless targeting	Loss of control can be mitigated only through the combined implementation of the recommendations outlined above. Transparency, shared operational protocols, clearly defined human accountability and responsibility, reverse-traceable algorithmic programming for machine actions, and reliable communication channels are critical.

IAD Component	Finding	Recommendations
Levels of Control	Imbalanced; transparency v opacity	Maintain human presence across all autonomy levels. Codify MHC, ensuring commanders retain override authority, especially during ambiguous engagements. Embed clear responsibility and risk-transfer norms so humans, not machines, remain final decision-makers.
Information available	Data distortion, machine manipulation	Quantify risk and uncertainty at each stage of decision-making. Build shared, objective standards for what counts as "**verified information**," enabling both sides to interpret risk similarly and reduce miscalculations. Ensure data and validation checks are mutually recognized.
Payoffs	Uncontrolled escalation and unacceptable human risk	Harmonise India–China interpretations of IHL, LOAC, and the Martens Clause. A shared legal spectrum ensures both nations classify violations objectively and pursue compatible accountability pathways. These guardrails help prevent unintended escalation and improve strategic stability.
Pattern of Interactions	Absence of Escalation level monitoring	Build a shared interpretation framework between India and China to reduce divergent readings of risk, intent, and legality. Standardize definitions of acceptable behavior, escalation triggers, and what constitutes an algorithmic error. Jointly articulate conditions that qualify as machine malfunction versus deliberate hostile action to prevent misattribution.
Outcomes	Unstable equilibrium	Create institutional clarity and transparent accountability mechanisms

IAD Component	Finding	Recommendations
	seen in Prisoner's Dilemma and Backward Induction Theory worsens due to opacity, ambiguity	supported by clearly bounded LAWS regulation. Establish bilateral communication channels to stabilize escalatory unpredictability and avoid confusion in interpreting autonomous system behavior.

Section VI: Conclusion

India and China have emerged as global leaders in software development, algorithmic innovation, and defense-driven AI ecosystems. Yet their strategic cultures diverge sharply: India relies on democratic civilian oversight and transparent military control, while China embraces opacity and coercive territorial expansion. With the world's longest unmarked border and two of the largest standing armies, the India–China dyad remains inherently unstable.

The **U.S.–India ten-year defense agreement signed on 31 October 2025**, aimed at enhancing coordination, information-sharing, and technology cooperation, introduces a vital stabilizing axis. For the agreement to serve its purpose, advancing regional stability and deterrence, both nations must craft a coherent strategy to enhance deterrence. India must safeguard its territorial integrity and counter Beijing's incremental territorialism, while the United States must reconcile its Indo-Pacific commitments with its ideological contest with China, particularly over Taiwan.

Implementing the recommendations outlined in this research through empowered, tri-national committees would provide the institutional framework needed to manage LAWS-driven instability. Establishing shared principles, transparent protocols, and mutually agreeable outcomes will be essential to prevent miscalculation at machine speed. Only with such mechanisms can peace and stability be sustained in the Indian subcontinent, regardless of shifting political outcomes or ideological differences among leaders.

Reference:

1 (*Omar Nelson Bradley Quotes - General of the US Army*, n.d.)
2 (Hagström, 2019)
3 (Pettyjohn, 2024)
4 (Andersson & Simon, 2024)
5 (Bruun, 2024)
6 (Bruun, 2024)
7 (Ostrom, 2009)
8 (*ChatGPT*, n.d.)
9 (Huang, n.d.)
10 (Zhao, 2018)
11 (Huang, n.d.)
12 (Huang, n.d.)
13 (Huang, n.d.)
14 (Huang, n.d.)
15 (Huang, n.d.)
16 (Huang, n.d.)
17 (Bousquet, 2008)
18 (Huang, n.d.)
19 (Bremer & Grieco, 2021)
20 (Chavannes et al., 2020)
21 (Shleifer, 2012)
22 (*ICRAC Working Paper #3 (CCW GGE April 2018)*, n.d.)
23 (Kahneman, 2024)

24 (Price et al., 2018a)

25 (Gruszczak & Kaempf, 2024)

26 (Davis, 2019a) 127

27 (Davis, 2019a)

28 (Bousquet, 2008)

29 (Davis, 2019a) 122

30 (Walker, 2021)

31 (Nadibaidze, 2022)

32 (Allen, n.d.)

33 (Nadibaidze, 2022)

34 ("Full Translation," n.d.)

35 ("Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects," 2023)

36 (Bode et al., 2023)

37 ("China's Strategic Ambiguity on the Issue of Autonomous Weapons Systems," 2022)

38 ("China's Strategic Ambiguity on the Issue of Autonomous Weapons Systems," 2022)

39 Unofficial translation of the '*Working Paper of the People's Republic of China on Lethal Autonomous Weapons Systems*', https://documents.unoda.org/wp-content/uploads/2022/07/Working-Paper-of-the-Peoples-Republic-of-China-on-Lethal-Autonomous-Weapons-Systems（English).pdf, dated July, 2022, accessed on 22 January, 2025

40 (Galliott et al., 2021)

41 (Benner, n.d.)

42 (Reddy, 2016)

43 (Scharre & Horowitz, 2015)

44 (Reddy, 2016)

45 (Doorn, 2010)

46 (Doorn, 2010)

47 (Doorn, 2010)

48 (Walker, 2021)

49 (Sparrow, 2007)

50 (Smith, 2022)

51 (Chengeta, 2016)

52 (Price et al., 2018a)

53 (Smith, 2022)

54 (Anderson & Waxman, 2013)

55 (Schmitt & Thurnher, 2013)

56 (Galliott et al., 2021)

57 (Johnson, 2019)

58 (Johnson, 2019)

59 (Parsons, 2010) 15.

60 (Price et al., 2018b)

61 (Price et al., 2018b)

62 (Price et al., 2018b)

63 (Winter, 2020)

64 (Price et al., 2018a)

65 (Winter, 2020)

66 (Winter, 2020)

67 (Abraham, 2024)

68 (Seloom, 2025)

69 (Davies et al., 2023)

70 (Seloom, 2025)

71 (Davies et al., 2023)

72 (Abraham, 2024)

73 (Seloom, 2025)

74 (Davies et al., 2023)

75 (Kumaraguru, 2020)

76 (Seloom, 2025)

77 (Seloom, 2025)

78 (MacLeod & Hahn, 2025)

79 (MacLeod & Hahn, 2025)

80 Davies, Harry, McKernan, Bethan and Sabbagh, Dan; 2024, The Guardian, https://www.theguardian.com/world/2023/dec/01/the-gospel-how-israel-uses-ai-to-select-bombing-targets accessed on 08 June, 2025.

81 Schmitt, Michael N., '*Israel-Hamas 2024 Symposium-The Gospel, Lavender, and the Law of Armed Conflict*', https://lieber.westpoint.edu/gospel-lavender-law-armed-conflict/ accessed on 08 June, 2025.

82 (Scharre, 2018)

83 (Sparrow, 2007)

84 (Pauly, 2025)

85 (Jakobsen, 2011)

86 (Mallory, 2018)

87 (Mallory, 2018)

88 (Price et al., 2018b)

89 (Scharre, 2016)

90 (Henley et al., 2006)

91 (Henley et al., 2006)

92 (Henley et al., 2006)

93 (Henley et al., 2006)

94 (Kastetter et al., 2023)

95 (Henley et al., 2006)

96 (Henley et al., 2006)

97 (Artificial Intelligence and Strategic Stability in South Asia, 2020)

98 (Artificial Intelligence and Strategic Stability in South Asia, 2020)

99 (Altmann & Sauer, 2017)

100 (Altmann & Sauer, 2017)

101 (Altmann & Sauer, 2017)

102 (Davis, 2019b)

103 (Sweijs & Zilincik, 2019)

104 (Sweijs & Zilincik, 2019)

105 (Sparrow, 2007)

106 (Sparrow, 2007)

107 (Gowda, 1996)

108 (Pettit & Sugden, 1989)

109 (Gowda, 1996)

110 (Logan & Saunders, n.d.)

111 (Kuhn, 2025)

112 (William Spaniel, 2012a)

113 (William Spaniel, 2012b)

114 (William Spaniel, 2012a)

115 (Ostrom, 2009)

116 (Ostrom, 2009)

Bibliography

Abraham, Y. (2024, April 3). 'Lavender': The AI machine directing Israel's bombing spree in Gaza. *+972 Magazine*. https://www.972mag.com/lavender-ai-israeli-army-gaza/

Allen, G. C. (n.d.). *Putin and Musk are right: Whoever masters AI will run the world*. CNAS. Retrieved November 23, 2025, from https://www.cnas.org/publications/commentary/putin-and-musk-are-right-whoever-masters-ai-will-run-the-world

Altmann, J., & Sauer, F. (2017). Autonomous Weapon Systems and Strategic Stability. *Survival*, *59*(5), 117–142. https://doi.org/10.1080/00396338.2017.1375263

Anderson, K., & Waxman, M. C. (2013). Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How the Laws of War Can. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2250126

Andersson, J. J., & Simon, S. (2024). *MINDING THE DRONE GAP: Drone warfare and the EU*. European Union Institute for Security Studies (EUISS). https://www.jstor.org/stable/resrep64045

Artificial intelligence and strategic stability in South Asia: New horses for an old wagon? (The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, pp. 39–45). (2020). Stockholm International Peace Research Institute. https://www.jstor.org/stable/resrep24515.12

Benner, T. (n.d.). *Brazil as a norm entrepreneur: The "Responsibility While Protecting" initiative*.

Bode, I., Huelss, H., Nadibaidze, A., Qiao-Franco, G., & Watts, T. F. A. (2023). Prospects for the global governance of autonomous weapons: Comparing Chinese, Russian, and US practices. *Ethics and Information Technology*, *25*(1), 5. https://doi.org/10.1007/s10676-023-09678-x

Bousquet, A. (2008). Chaoplexic Warfare or the Future of Military Organization. *International Affairs (Royal Institute of International Affairs 1944-)*, *84*(5), 915–929. https://www.jstor.org/stable/25144928

Bremer, M. K., & Grieco, K. A. (2021). The Air Littoral: Another Look. *The US Army War College Quarterly: Parameters*, *51*(4), 67–80. https://doi.org/10.55540/0031-1723.3092

Bruun, L. (2024). *Using scenarios as a method to identify limits and requirements for the development and use of AWS* (TOWARDS A TWO-TIERED APPROACH TO REGULATION OF AUTONOMOUS WEAPON SYSTEMS, pp. 4–7). Stockholm International Peace Research Institute. https://www.jstor.org/stable/resrep62434.6

ChatGPT. (n.d.). ChatGPT. Retrieved November 23, 2025, from https://chatgpt.com/

Chavannes, E., Klonowska, K., & Sweijs, T. (2020). *Exploring the AWS solution space* (Governing Autonomous Weapon Systems, pp. 17–27). Hague Centre for Strategic Studies. https://www.jstor.org/stable/resrep24196.7

Chengeta, T. (2016). Measuring Autonomous Weapon Systems Against International Humanitarian Law Rules. *Journal of Law & Cyber Warfare*, *5*(1), 66–146. https://www.jstor.org/stable/26441266

China's Strategic Ambiguity on the Issue of Autonomous Weapons Systems. (2022). *Global: Jurnal Politik Internasional*, *24*(1). https://doi.org/10.7454/global.v24i1.706

Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects. (2023). *Non-Exhaustive Compilation of Definitions and Characterizations, CCW/GGE.1/2023/CRP.1*. https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-Group_of_Governmental_Experts_on_Lethal_Autonomous_Weapons_Systems_(2023)/CCW_GGE1_2023_CRP.1.pdf

Davies, H., McKernan, B., & Sabbagh, D. (2023, December 1). 'The Gospel': How Israel uses AI to select bombing targets in Gaza. *The Guardian*. https://www.theguardian.com/world/2023/dec/01/the-gospel-how-israel-uses-ai-to-select-bombing-targets

Davis, Z. (2019a). Artificial Intelligence on the Battlefield: Implications for Deterrence and Surprise. *PRISM*, *8*(2), 114–131. https://www.jstor.org/stable/26803234

Davis, Z. (2019b). Artificial Intelligence on the Battlefield: Implications for Deterrence and Surprise. *PRISM*, *8*(2), 114–131. https://www.jstor.org/stable/26803234

Doorn, N. (2010). A Rawlsian Approach to Distribute Responsibilities in Networks. *Science and Engineering Ethics*, *16*(2), 221–249. https://doi.org/10.1007/s11948-009-9155-0

Full Translation: China's "New Generation Artificial Intelligence Development Plan" (2017). (n.d.). *DigiChina*. Retrieved November 23, 2025, from https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/

Galliott, J., MacIntosh, D., & Ohlin, J. D. (Eds.). (2021). *Lethal Autonomous Weapons: Re-Examining the Law and Ethics of Robotic Warfare*. Oxford University Press. https://doi.org/10.1093/oso/9780197546048.001.0001

Gowda, M. V. R. (1996). Teaching the Prisoners' Dilemma. *Journal of Policy Analysis and Management*, *15*(4), 646–653. https://www.jstor.org/stable/3326054

Gruszczak, A., & Kaempf, S. (Eds.). (2024). *Routledge Handbook of the Future of Warfare*. Routledge.

Hagström, M. (2019). *Military applications of machine learning and autonomous systems* (The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, pp. 32–38). Stockholm International Peace Research Institute. https://www.jstor.org/stable/resrep24525.10

Henley, L. D., Scobell, A., & Wortzel, L. M. (2006). *War Control: Chinese Concepts of Escalation Management* (SHAPING CHINA'S SECURITY ENVIRONMENT:, pp. 81–104). Strategic Studies Institute, US Army War College. https://www.jstor.org/stable/resrep11961.8

Huang, H.-M. (n.d.). Autonomy Levels for Unmanned Systems (ALFUS) Framework Volume I: Terminology Version 1. *NIST Special Publication*.

ICRAC Working Paper #3 (CCW GGE April 2018): Guidelines for the human control of weapons systems | ICRAC. (n.d.). Retrieved November 8, 2025, from https://www.icrac.net/icrac-working-paper-3-ccw-gge-april-2018-guidelines-for-the-human-control-of-weapons-systems/

Jakobsen, P. V. (2011). Pushing the Limits of Military Coercion Theory. *International Studies Perspectives*, *12*(2), 153–170. https://www.jstor.org/stable/44218656

Johnson, J. (2019). Artificial intelligence & future warfare: Implications for international security. *Defense & Security Analysis*, *35*(2), 147–169. https://doi.org/10.1080/14751798.2019.1600800

Kahneman, D. (2024). *Thinking, fast and slow* (Reissued). Penguin Books.

Kastetter, A., Alsaied, J., Bess, C., Cortino, B., Gum, D., Kim, J., Koç, D. E., Horan, L. S., Madeira, J., Means, A., Powell, W., Symonds, E., & Zhukov, D. (2023). *Chinese War Controls and U.S. Approaches to Deterrence* (On the Horizon, pp. 54–64). Center for Strategic and International Studies (CSIS). https://www.jstor.org/stable/resrep47437.8

Kuhn, S. (2025). Prisoner's Dilemma. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2025). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2025/entries/prisoner-dilemma/

Kumaraguru, Y. (2020). *A pre-emptive ban on lethal autonomous weapon systems* (The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk, pp. 55–58). Stockholm International Peace Research Institute. https://www.jstor.org/stable/resrep24515.14

Logan, D. D. C., & Saunders, D. P. C. (n.d.). *IMPLICATIONS OF A PRC SHIFT TO A LAUNCH-ON-WARNING NUCLEAR POSTURE*.

MacLeod, I., & Hahn, E. (2025). The Role of an Operational Frame in Furthering the International Debate on Lethal Autonomous Weapons Systems. In D.-P. Baker & M. Hilborne (Eds.), *War 4.0* (1st ed., pp. 163–170). ANU Press. https://www.jstor.org/stable/jj.30412870.12

Mallory, K. (2018). *New Challenges in Cross-Domain Deterrence*. RAND Corporation. https://www.jstor.org/stable/resrep17639

Nadibaidze, A. (2022). Great power identity in Russia's position on autonomous weapons systems. *Contemporary Security Policy*, *43*(3), 407–435. https://doi.org/10.1080/13523260.2022.2075665

Omar Nelson Bradley Quotes—General of the US Army. (n.d.). Retrieved November 23, 2025, from https://www.military-quotes.com/omar-bradley.htm

Ostrom, E. (2009). *Understanding Institutional Diversity*. Princeton University Press. https://doi.org/10.1515/9781400831739

Parsons, C. (2010). *How to Map Arguments in Political Science*. Oxford University Press.

Pauly, R. B. C. (2025). *The Art of Coercion: Credible Threats and the Assurance Dilemma*. Cornell University Press. https://lccn.loc.gov/2025001691

Pettit, P., & Sugden, R. (1989). The Backward Induction Paradox. *The Journal of Philosophy*, *86*(4), 169–182. https://doi.org/10.2307/2026960

Pettyjohn, S. (2024). *Types of Drones Used in the Ukraine War* (Evolution Not Revolution, pp. 5–6). Center for a New American Security. https://www.jstor.org/stable/resrep57900.5

Price, M., Walker, S., & Wiley, W. (2018a). The Machine Beneath: Implications of Artificial Intelligence in Strategic Decision making. *PRISM*, *7*(4), 92–105. https://www.jstor.org/stable/26542709

Price, M., Walker, S., & Wiley, W. (2018b). The Machine Beneath: Implications of Artificial Intelligence in Strategic Decision making. *PRISM*, *7*(4), 92–105. https://www.jstor.org/stable/26542709

Reddy, R. S. (2016). *India and the Challenge of Autonomous Weapons*. Carnegie Endowment for International Peace. https://www.jstor.org/stable/resrep12857

Scharre, P. (2016). *Autonomous Weapons and Unintended Engagements* (Autonomous Weapons and Operational Risk, pp. 18–24). Center for a New American Security. https://www.jstor.org/stable/resrep06321.7

Scharre, P. (2018). *Army of None: Autonomous Weapons and the Future of War*. W. W. Norton & Company.

Scharre, P., & Horowitz, M. C. (2015). *An Introduction to AUTONOMY in WEAPON SYSTEMS*. Center for a New American Security. https://www.jstor.org/stable/resrep06106

Schmitt, M. N., & Thurnher, J. S. (2013). *Out of the Loop: Autonomous Weapon Systems and the Law of Armed Conflict*. 51. http://harvardnsj.org/wpcontent/uploads/2013/02/Schmitt-Autonomous-Weapon-Systems-and-IHL-Final.pdf

Seloom, M. (2025, October 29). Will Israel's Algorithmic Counter-Insurgency Proliferate to the West? *War on the Rocks*. https://warontherocks.com/2025/10/will-israels-algorithmic-counter-insurgency-proliferate-to-the-west/

Shleifer, A. (2012). Psychologists at the Gate: A Review of Daniel Kahneman's "Thinking, Fast and Slow." *Journal of Economic Literature*, *50*(4), 1080–1091. https://www.jstor.org/stable/23644912

Smith, P. T. (2022). Resolving responsibility gaps for lethal autonomous weapon systems. *Frontiers in Big Data*, *5*, 1038507. https://doi.org/10.3389/fdata.2022.1038507

Sparrow, R. (2007). Killer Robots. *Journal of Applied Philosophy*, *24*(1), 62–77. https://doi.org/10.1111/j.1468-5930.2007.00346.x

Sweijs, T., & Zilincik, S. (2019). *From Deterrence to Cross Domain Deterrence* (Cross Domain Deterrence and Hybrid Conflict, pp. 11–18). Hague Centre for Strategic Studies. https://www.jstor.org/stable/resrep24191.5

Walker, P. (2021). Leadership Challenges from the Deployment of Lethal Autonomous Weapon Systems: How Erosion of Human Supervision Over Lethal Engagement Will Impact How Commanders Exercise Leadership. *The RUSI Journal*, *166*(1), 10–21. https://doi.org/10.1080/03071847.2021.1915702

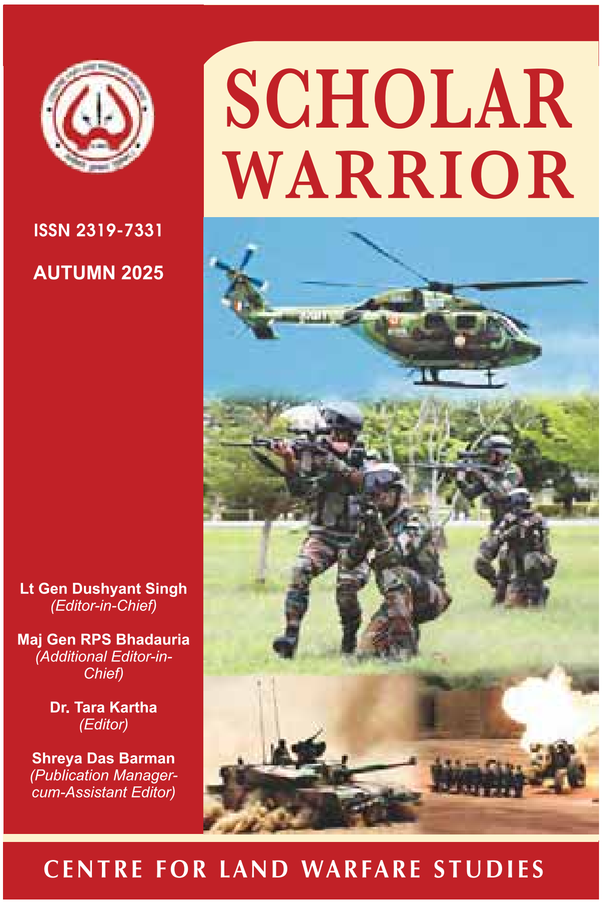
William Spaniel (Director). (2012a, August 31). *Game Theory 101 (#1): Introduction* [Video recording]. https://www.youtube.com/watch?v=NSVmOC_5zrE

William Spaniel (Director). (2012b, September 8). *Game Theory 101 (#17): Backward Induction* [Video recording]. https://www.youtube.com/watch?v=pyLKkN5HpDY

Winter, E. (2020). *The compatibility of the use of autonomous weapons with the principle of precaution in the law of armed conflict*. https://doi.org/10.4337/mllwr.2020.02.18

Zhao, T. (2018). *The Impact of Future Unmanned Systems* (TIDES OF CHANGE, pp. 61–68). Carnegie Endowment for International Peace. https://www.jstor.org/stable/resrep26932.13

SUBSCRIBE NOW



SUBSCRIPTION RATES

IN INDIA	SAARC COUNTRIES	OTHER COUNTRIES
☐ Rs.500/- per copy	☐ US $ 15 per copy	☐ US $ 20 per copy

☐ Rs.1000/- Annual Subscription (2 issues)

TO SUBSCRIBE SEND YOUR REQUEST TO

Centre for Land Warfare Studies (CLAWS)
RPSO Complex, Parade Road, Delhi Cantt, New Delhi - 110010
Tel: +91-11-25691308
• Fax: +91-11-25692347 • Army: 33098
E-mail: landwarfare@gmail.com
www.claws.co.in

How do Lethal Autonomous Weapons Systems (LAWS) shape escalation dynamics between India and China in a chaoplexic, AI-driven battlespace? While existing research - from Scharre on autonomy to Sweijs and Zilincik on cross-domain deterrence, and from Winter, Galliott, Colijn, Podar, and Hanley on legality, ethics, risk, and Chinese escalation behavior provides valuable insights, these perspectives remain largely siloed. What is insufficiently explored is how technological, legal, ethical, and doctrinal factors interact when autonomy compresses decision cycles and obscures intent. To address this gap, the study analyzes a self-developed Scenario Building Exercise and employs two game-theory models - the Prisoner's Dilemma and Backward Induction Theory, to trace how unintended outcomes may arise. Using the Institutional Analysis and Development (IAD) Framework, it evaluates contextual variables, actor positions, and interaction patterns as sources of cumulative instability. The study argues that LAWS lower risk thresholds, heighten opacity, and accelerate decision-making, thereby pushing strategic behavior toward escalation. By so examining LAWS in structural terms, this study addresses a key conceptual gap with findings that may help develop an autonomous-era escalation continuum for the Sino-Indian context.

• • •

Brigadier Arjun Uppal, SM, VSM, a fourth-generation officer, is from 44 Armoured Regiment. The officer has extensive overseas, CI, desert, semi-desert and high-altitude experience, and has commanded an Armoured Regiment and independent Mechanised Brigade.

He is currently pursuing his doctoral studies and was selected in 2025 into the prestigious "Carlisle Scholars Program-Leader Imeersion Labs", which privileges experiential, problem-based learning for 17 nominated students in the United States Army War College.

The Centre for Land Warfare Studies (CLAWS), New Delhi, is an independent Think Tank dealing with contemporary issues of national security and conceptual aspects of land warfare, including conventional & sub-conventional conflicts and terrorism. CLAWS conducts research that is futuristic in outlook and policy oriented in approach.

CLAWS Vision: To be a premier think tank, to shape strategic thought, foster innovation, and offer actionable insights in the fields of land warfare and conflict resolution.

CLAWS Mission: Our contributors aim to significantly enhance national security, defence policy formulation, professional military education, and promote the attainment of enduring peace.

Website: www.claws.co.in Contact us: landwarfare@gmail.com

SABRE & QUILL

MRP: ₹ 100.00 US$ 5.00