



NATO and the Problem of Compounded Crises

What NATO exercises test - and what remains unproven
under simultaneous hybrid, cyber and military pressure

DATE

9 October 2026

GEOGRAPHIC AREA

NATO / Europe

INTELLIGENCE DOMAIN

Readiness / Hybrid Warfare / C2 / Cyber / Critical Infrastructure / Exercises / Resilience

Executive Summary

NATO readiness is not a single alliance-wide score. The NATO Force Model organises forces in three readiness tiers: 0-10 days, 10-30 days and 30-180 days. Allies declare and report forces, while the NATO Readiness Process provides for verification against agreed standards. These measures are necessary, but they describe only part of the operational chain. A force listed at short notice must still be released nationally, assembled with its equipment, moved across borders, integrated into a joint command, supplied and kept effective after the opening deployment.

The proposition that NATO exercises generally assume orderly, single-domain crises is too broad. Recent exercises have tested hybrid civil-military scenarios, Article 5 command and control, simultaneous Allied headquarters, contested communications, dispersed threats, cyber and space coordination, large-scale reinforcement and sustainment. NATO's 2026 resilience requirements also address physical attack, sabotage, cyber disruption, supply-chain dependencies, transport, energy, communications and the simultaneous movement of civilians and forces.

The unresolved issue is whether these elements are tested together, at operational scale and for long enough to measure continuity under cascading disruption. Public descriptions identify the objectives and broad design of exercises, but rarely disclose the inject sequence, assessment criteria, force-level performance data or remediation results. Command-post exercises, cyber ranges and live-force deployments each reveal different parts of the problem. Their success cannot by itself establish that the Alliance can sustain a contested operation while civilian infrastructure, communications, transport and political decision-making are simultaneously under pressure.

The central risk is therefore a measurement and integration gap, not an absence of complex exercises. Readiness measured as force availability or successful exercise completion may overstate the capacity of the wider system to preserve command, movement and sustainment when several dependencies degrade at once. A more discriminating test would link ambiguous hybrid activity to bounded cyber and physical disruption, then track the consequences through national decisions, cross-border movement, host-nation support, logistics and recovery.

1. What NATO Means by Readiness

The NATO Force Model places nationally designated forces into readiness tiers and connects them to regional and collective defence plans. The model is intended to make forces available at shorter notice and in greater scale than the former NATO Response Force framework. NATO states that forces are subject to verification to ensure alignment with standards. This is a stronger basis for planning than a nominal order of battle, but the tiers remain notice-to-readiness categories; they do not, on their own, represent a common measure of sustained combat effectiveness across all Allies.

Operational readiness is a chain of conditions. Personnel availability and equipment serviceability matter, as do ammunition, maintenance, deployable headquarters, communications, transport, host-nation support and national release procedures. Strategic mobility then depends on border permissions, rail and port capacity, route security, fuel distribution, reception and onward movement. At the destination, forces must integrate with the designated command structure and continue operating while drawing on national and commercial support. Failure at any critical link can delay or reduce the effect of a force that met its initial readiness target.

Four distinctions therefore matter. Declared readiness is not the same as readiness sustained over time; unit availability is not deployability; deployability is not operational effectiveness; and exercise success is not proof of resilience in war. These are not reasons to discount readiness tiers or exercises. They define what those instruments can establish and where additional evidence is required.

Measurement should separate the time to make a force available from the time to deploy, integrate and sustain it. It should also identify which conditions are tested in the field, which are simulated, and which depend on civilian providers or decisions outside military command. Without that separation, an aggregate readiness label can conceal very different national and alliance-level performance.

2. How Recent Exercises Represent Crisis

NATO's exercise programme already includes distinct formats for strategic decisions, operational command, cyber defence and live force movement. CMX25 rehearsed consultation and decision-making at the political-military level through a complex civil-military scenario in a hybrid environment. It involved civilian and military staffs across Allied capitals, NATO Headquarters and the Strategic Commands, but no deployed forces. Its design is suited to testing consultation and decision procedures; it cannot directly demonstrate whether a brigade, rail network or fuel system can perform under physical attack.

Steadfast Duel 25 was a computer-assisted command-post exercise involving about 7,000 military and civilian personnel. It trained an Article 5 scenario across all 32 Allies and, for the first time at operational level, connected the three Joint Force Commands in a simulated 24-hour battlespace. Steadfast Dagger 25 trained the next Allied Reaction Force rotation across land, maritime, air, special operations, cyber, space and logistics. Public reporting describes dispersed threats, contested communications and complex civil-military interaction among its operational challenges. These examples do not fit a simple picture of linear or single-domain training.

Steadfast Defender 24 added a different form of evidence: roughly 90,000 personnel from all Allies and Sweden took part in a series of 15 national tactical exercises focused on transatlantic reinforcement, interoperability, deployment and sustainment. Follow-on logistics discussions identified work on mobility corridors, digitalisation, multinational maintenance, interoperability and delivery in a contested environment. This is evidence that large-scale movement is being exercised and that the exercise generated practical logistics issues for further work; it is not evidence that every constraint has been resolved.

Cyber training also has dedicated venues. Cyber Coalition exercises collective cyber defence, while Locked Shields uses a simulated national infrastructure; its 2026 edition included systems such as power grids, satellite management and 5G. Such specialisation supports technical realism and focused learning. The analytical question is whether the operational consequences from these cyber scenarios are carried into the same command and logistics problems being assessed in force exercises, rather than whether cyber is represented at all.

These activities should be treated as complementary instruments rather than as one consolidated test. A command-post exercise can reveal decision latency and coordination problems while simulating force movement; a cyber range can expose technical defence issues without necessarily measuring their effect on reinforcement; a live exercise can test physical transit and support but may not reproduce strategic ambiguity. A shared set of dependencies and operational outputs would make the results comparable: whether a cyber incident changes movement timelines, whether an alternative route can carry the required flow, and whether authority to reroute is clear while attribution remains uncertain.

The public record supports a more precise conclusion: NATO exercises address many of the relevant domains and have become more complex, but each format has a defined level, audience and purpose. Public exercise summaries do not provide enough detail to determine how often multiple disruptions are synchronised across all levels or how performance is scored against sustained operational outcomes.

3. The Compounded-Crisis Problem

A compounded crisis is not simply a long list of simultaneous threats. Its operational effect comes from interactions and timing. A physical attack on a transport node may slow reinforcement; a cyber incident may impair booking, signalling or cargo visibility; communications disruption may delay the attribution of the event and the rerouting decision; disinformation may increase public pressure or obstruct local cooperation. The same disruption can affect civilian movement, military traffic and emergency response, forcing authorities to prioritise scarce routes and services.

Ambiguity matters because the response is distributed across national authorities, NATO commands, civilian agencies and commercial operators. A suspicious incident may be serious enough to demand protective action while remaining too uncertain to attribute quickly. Political consultation, legal authorities and public communication then become operational variables. A force may be physically ready while the system that authorises movement, protects infrastructure or allocates transport capacity is still resolving jurisdiction and risk.

The relevant test is not whether exercise control can inject several events. It is whether the training audience must make linked decisions while information is incomplete and resources are contested. A cyber disruption that causes no change to the logistics picture tests a different problem from one that forces a commander to reroute forces while a host nation protects civilian evacuation, restores power and responds to sabotage. Conversely, an exercise that degrades every service at once may become an implausible stress test with little diagnostic value.

A credible design should use a limited number of concurrent disruptions, chosen for their dependency effects. The sequence should permit analysts to identify which decision or dependency produced delay, what redundancy worked, and whether recovery restored the required military function. The objective is to expose cascading effects without turning the scenario into a catastrophe script.

4. Command and Control Under Degradation

Command resilience depends on more than a functioning headquarters or a secure network. It requires commanders to maintain an adequate operational picture, validate reports, assign authority and transmit decisions when some communications paths, data feeds or commercial services are delayed or unavailable. In a hybrid setting, information may be technically accessible but difficult to trust. Duplicate, contradictory or manipulated reporting can consume decision time even when networks remain online.

Exercises should distinguish graceful degradation from total failure. Relevant measures include time to recognise a loss, continuity of command across alternate sites, message delivery and acknowledgement, time to establish a usable common picture, and the ability to reconstitute a degraded service. They should also test whether civil authorities and infrastructure operators receive clear priorities and whether national and Allied command arrangements remain aligned when attribution is uncertain.

The key operational question is whether the command system can continue to direct forces and enable civil support while operating with reduced information and competing demands. A successful digital demonstration is not sufficient if its assumed network, power or commercial data service is unavailable in a more demanding scenario.

5. Logistics and Strategic Mobility Under Pressure

NATO's reinforcement and sustainment model depends on national infrastructure and substantial civilian and commercial support. Forces require fuel, food, maintenance, medical support, transport services and specialised supplies. Host-nation support enables forces to enter, transit and operate across Allied territory. Rail capacity, ports, border procedures, road movement, route security and digital logistics systems therefore form part of readiness, even when they sit outside a military unit's own readiness report.

Steadfast Defender 24 demonstrated that NATO can organise a large transatlantic reinforcement exercise and move forces across Europe. The subsequent focus on mobility corridors, fuel requirements, digitalisation, multinational maintenance and performance in a contested environment indicates that movement and sustainment remain active planning concerns. These discussions should not be presented as proof of a shortfall in every corridor or Ally. They do show that exercise output must feed into practical changes in infrastructure, procedures and support arrangements.

A compounded-crisis test should follow the deployment chain from alert to arrival and onward movement. It should measure whether border clearances are completed within planned timelines; how quickly forces can switch between routes; whether logistics data remain available; how fuel and ammunition are prioritised; and whether maintenance and medical support continue when commercial providers are disrupted. It should assess competing civilian and military demand, including evacuation or emergency response, rather than treating host-nation support as an unlimited service.

Sustainment needs a time dimension. Initial movement may be successful while the flow of replacements, fuel, ammunition or repair capacity degrades after several days. Exercises should therefore record not only arrival but the rate at which the force can be supplied, maintained and regenerated under partial disruption.

6. Civilian Infrastructure and Multi-Domain Resilience

NATO's civil preparedness framework recognises three functions: continuity of government, continuity of essential services and civil support to military operations. Its 2026 resilience requirements cover continuity of government, energy, movement of people, food and water, health, communications and transport. The updated requirements address physical attack and sabotage, damaging cyber activity, supply-chain dependencies, repair and restoration, and reliance on space-based services. This policy framework already treats resilience as a whole-of-government and whole-of-society problem.

The test is whether those requirements are connected to military outcomes during exercises. Power and communications support command; transport and border systems support reinforcement; health and emergency services sustain both the population and military operations; commercial operators control or maintain many of these functions. A crisis may therefore degrade military effectiveness without attacking a military unit directly. Counter-UAS protection, infrastructure repair, cyber response and civil-military coordination must be assessed as an interdependent package where the scenario warrants it.

Testing should involve the authorities and operators whose decisions shape the outcome, within realistic legal and security limits. It should examine information-sharing thresholds, priority access to communications, restoration timelines, alternative transport capacity and the process for resolving competing civil and military needs. The exercise should also record where a capability exists on paper but cannot be activated, shared or sustained at the required speed.

7. Assessment and Indicators

The initial hypothesis is confirmed only in a qualified form. Recent NATO exercises do not generally reduce readiness to tidy, single-domain scenarios: CMX25 included a complex hybrid civil-military setting; Steadfast Duel 25 linked all 32 Allies and three Joint Force Commands in a continuous command-post environment; Steadfast Dagger 25 addressed contested communications, dispersed threats and multi-domain coordination; Steadfast Defender 24 exercised large-scale reinforcement; and dedicated cyber exercises tested critical infrastructure. NATO has also revised its resilience requirements to address the same cross-sector dependencies that a compounded crisis would exploit.

The remaining gap is one of integration, duration and evidence. Public reporting does not establish how often the same exercise combines ambiguous hybrid activity, cyber and physical disruption, civilian infrastructure loss, cross-border friction and sustained logistics pressure, then measures operational performance through recovery. Nor does it make assessment criteria or remediation outcomes consistently visible. This absence of public detail is not proof that such testing does not occur. It limits the conclusions that can be drawn from exercise participation or successful completion alone.

For readiness assessment, the decisive question is whether a force and the system around it can retain decision authority, move, integrate, sustain and recover under pressure. Exercises should be judged by demonstrated outcomes rather than the number of participants, domains named or injects delivered. Results should identify the specific dependency that failed, the operational consequence, the corrective action and whether the action was later retested.

A useful assessment separates four stages: design, execution, outcome and remediation. Design asks whether the scenario represented relevant dependencies and concurrent pressure. Execution asks whether participants had to make decisions with incomplete information and competing priorities. Outcome records whether the required military or civil function was maintained, degraded or restored, and for how long. Remediation tracks whether the identified gap produced a funded or assigned corrective action and whether a later exercise verified it. This approach distinguishes a complex-looking scenario from a test that produces evidence about operational performance.

Exercise constraints also matter. Command-post scenarios are repeatable and allow staffs to rehearse decisions; cyber ranges provide technical depth; live exercises involve real movement and national procedures. Safety, classification and training objectives limit how closely any one event can reproduce a real crisis. Public summaries are not full evaluation reports, and the absence of a described inject cannot establish that it was omitted. The assessment should therefore remain bounded: the available record demonstrates substantial adaptation, while leaving alliance-wide performance under prolonged, simultaneous disruption unproven in public.

Indicators to monitor include the complexity and sequencing of exercise injects; communications degradation and restoration; mobilisation, border crossing and onward-movement timelines; fuel, ammunition and maintenance sustainment; port and rail alternatives; host-nation support under competing civilian demand; counter-UAS integration; continuity of command; information sharing with civilian operators; and closure of lessons identified through follow-on testing.