RESEARCH PAPER VERSION 1.1 NOVEMBER 2026

VIRTUAL RED AIR — *for the* — NATO INTEGRATED AIR AND MISSILE DEFENCE CENTRE OF EXCELLENCE

Who Flies the Threat?

Generating credible threat mass when the live environment can no longer provide it

COUNTER-UAS PERSPECTIVE

From Red Air to the drone swarm



Nikolaos Makrygiannis · Apostolos Papadopoulos

Hellenic Air Force (ret.) · Virtual Red Air

VIRTUAL RED AIR — *for the* — NATO INTEGRATED AIR AND MISSILE DEFENCE CENTRE OF EXCELLENCE

Who Flies the Threat?

Generating credible threat mass when the live environment can no longer provide it

COUNTER-UAS PERSPECTIVE

From Red Air to the drone swarm

THE QUESTION

When a fighter force can no longer spare the aircraft to fly against itself, where does the opponent come from?

THE ANSWER

It has to be manufactured, and for the drone swarm over the home airfield there is no other source. This paper sets out what the manufactured opponent must do to be worth flying against.



AUTHORS

Nikolaos Makrygiannis
Major General (OF-7) (ret.), Hellenic Air Force

Apostolos Papadopoulos
Brigadier General (OF-6) (ret.), Hellenic Air Force

CONTRIBUTING AUTHOR

Ioannis Stathakis
Colonel (OF-5), Hellenic Army

VIRTUAL RED AIR
Founders

NATO IAMD COE
Concept Development & Experimentation Branch Head
Counter-UAS Subject Matter Expert

RESEARCH PAPER · VERSION 1.1 · NOVEMBER 2026

UNCLASSIFIED

About this paper

This paper examines how air forces can generate credible adversary presentations at the scale modern training requires, at a time when live flying programmes across the Alliance are contracting rather than growing and the threat now includes uncrewed aircraft in numbers no live programme can present. It is written for the operational, capability and acquisition communities and for those responsible for integrated air and missile defence. It stays deliberately at the level of concepts, architectures and published evidence. It describes no proprietary system, endorses no supplier, and goes no further than the open record on how uncrewed aircraft are defeated.

Passages marked ⌘ **counter-UAS perspective** extend the argument to the crews who defend against uncrewed aircraft, and to the air bases those aircraft now reach.

All sources are open. Nothing in this paper draws on classified material, and no national tactic, procedure or performance figure is disclosed.

TITLE	Who Flies the Threat? Virtual Red Air: generating credible threat mass when the live environment can no longer provide it
VERSION	1.1, public release edition
DATE	November 2026
CLASSIFICATION	Unclassified
AUTHORS	N. Makrygiannis and A. Papadopoulos
CONTRIBUTING AUTHOR	Colonel I. Stathakis, NATO IAMD COE
CONTACT	info@virtualredair.com · www.virtualredair.com

© 2026 Virtual Red Air and the authors,
N. Makrygiannis and A. Papadopoulos.

PUBLIC RELEASE

Made available by the NATO Integrated Air and Missile Defence Centre of Excellence with the permission of the rights holders.

Licence. This edition is made available under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence (CC BY-NC-ND 4.0, creativecommons.org/licenses/by-nc-nd/4.0). It may be copied and shared, in whole or in part, for non-commercial purposes, with attribution to the authors and without alteration. Any other use requires the prior written consent of the rights holders. The Centre's own use of the paper is not limited by this licence.

Suggested citation. Makrygiannis, N. and Papadopoulos, A., with Stathakis, I. (2026) Who Flies the Threat? Virtual Red Air: generating credible threat mass when the live environment can no longer provide it. Research Paper, Version 1.1. Virtual Red Air for the NATO Integrated Air and Missile Defence Centre of Excellence.

Disclaimer. The views expressed are those of the authors alone and do not represent the position of NATO, of the IAMD COE or of any nation. Making the paper available implies no endorsement by the Centre of any organisation, product or service.

Declaration of interest. N. Makrygiannis and A. Papadopoulos are founders of Virtual Red Air, which develops synthetic training. The paper describes no proprietary system and evaluates no product.

ABSTRACT

The opponent is the training

Fighter forces are being asked to hold a standard of readiness that their live flying programmes can no longer produce. When the programme contracts, the adversary function is the first thing cut, because in an organic model every adversary sortie is flown by an aircraft and a pilot who would otherwise have been training. The events survive on the schedule. What they teach has quietly changed.

This paper argues that the adversary, rather than the platform or the hour, is the scarce commodity in modern air combat training, and that it is the one input which can now be manufactured. It sets out what the adversary function actually delivers, which is pressure, surprise, dissimilarity and the raw material of assessment, and why the contracted adversary air industry that emerged over the past decade relieves the symptom without closing the gap. A commercial aggressor fleet can supply airframes and experienced aircrew. It cannot supply the radar, the electronic order of battle, the datalink behaviour or the sheer numbers of a peer force.

The paper then addresses the synthetic alternative on its own terms. It examines where the fidelity thresholds actually sit, and what happens below them, since a poorly specified device does not simply fail to teach but actively teaches the wrong thing. It distinguishes between an adversary that is merely present and one that is credible, and it treats assessment as the central problem rather than an afterthought, framed by the F2T2E2A cycle of Alliance targeting doctrine, on the grounds that a force which cannot measure the quality of its training cannot defend the decision to change it.

A counter-UAS perspective runs through the argument, because the uncrewed threat makes it sharper. Fighters have been committed against mass drone attacks over Israel and against drones over Poland, and drones have closed airports and disrupted a fighter base in Allied territory. No contracted fleet can present that threat in the numbers in which it now arrives, and outside tightly controlled trials no authority will allow it to be flown over a working airfield.

Finally the paper proposes a federated architecture that is classified by design rather than by default, a role for the NATO Integrated Air and Missile Defence Centre of Excellence (IAMD COE) as the Alliance's laboratory for the synthetic threat and for counter-UAS research, and a maturity model that makes the roadmap arguable, so that investment produces usable capability early rather than complete capability late.

KEYWORDS

Adversary air; Red Air; distributed synthetic training; LVC; training fidelity; negative training; constructive entities; behaviour modelling; readiness assessment; joint targeting; F2T2E2A; measures of effectiveness; simulation interoperability; counter-UAS; Class I UAS; integrated air and missile defence.

Contents

1 The arithmetic of readiness 7
Sortie rates, manning, availability and the compounding deficit

2 What the adversary actually delivers 9
Presentation, pressure, unpredictability and the basis of assessment

3 The contracted adversary and its ceiling ⌘ 12
An industry that solved throughput but not representation, and a threat it was never built to fly

4 Live, virtual and constructive: the working definitions ⌘ 15
Standards, correlation and latency, and the tactical gap in counter-UAS training

5 Fidelity, negative training and the limits of immersion 18
Where the thresholds sit, and what happens below them

6 Threat mass: generating an opponent worth fighting ⌘ 21
From scripted entities to reactive agents, and the drone over the home airfield

7 The assessment problem ⌘ 26
Why measurement, not immersion, is the scarce commodity, and the targeting cycle that can frame it

8 Architectures for a national and allied capability ⌘ 30
Federation, classification by design and sovereignty, and a second crew on the same hub

9 Governance, assurance and the regulatory frame 33
Human authority, the European AI regulation and the assurance of training systems

10 From threat mass to evidence: a role for the IAMD COE ⌘ 35
A synthetic laboratory for counter-UAS research, and the lessons that make it urgent

11 A maturity model and a sequenced roadmap 37
Five levels, and what to attempt in the first thirty-six months

· Conclusion ⌘ 39
The opponent is the capability, and notes and references from page 41

⌘ Includes a counter-UAS perspective

SECTION ONE

The arithmetic of readiness

Every air force expresses readiness as a rate. The rate a fighter pilot needs has barely moved in three decades. The rate the system can deliver has fallen to a fraction of it. Almost everything else in this paper follows from that single divergence, and from the fact that the adversary is the part of the rate that gets cut first.

NO SQUADRON ANNOUNCES THE MORNING IT STOPS BEING READY. There is no declaration and no red light on a wall. The programme still fills, the currencies are still signed, the monthly return still closes on time. What changes is quieter than that. The four-ship becomes a two-ship. The adversary brief gets shorter because there is less to brief. The debrief finishes early because there is less to say. Competence erodes at the edges first, in exactly those events that are hardest to arrange, most expensive to fly and quickest to perish.

That erosion is arithmetic before it is anything else. Readiness is a balance between the rate at which aircrew acquire skill and the rate at which skill decays, and decay does not negotiate. Every air force manages the balance through a continuation programme expressed in sorties per month and qualifications that expire on a date. When the count falls, the effect is not proportionate. Crews stop progressing, instructors spend their allocation requalifying colleagues rather than advancing them, and the most demanding events are deferred into a period that never quite arrives.

HOW LARGE THE GAP HAS BECOME

An average fighter pilot is generally held to need somewhere in the region of nine to ten sorties a month to be assessed as combat mission ready. Analysis presented in September 2026 by the Mitchell Institute for Aerospace Studies indicated that aircrew in non-deployed United States Air Force units were being offered roughly one sortie a month. The same analysis noted that the service stopped publishing sortie rate data after fiscal year 2023, when the figure stood at about five and a half, a little over sixty per cent of requirement.[1] Reasonable people can argue about the precise number today. Nobody is arguing about the direction, and the distance between requirement and delivery has stopped being a margin and become a multiple.

It is tempting to file this as an American problem of budget cycles and depot backlogs. It is not. The same three pressures appear in every air force in the Alliance, and each has a different remedy.

THREE DEFICITS, THREE CLOCKS

People — 5 to 10 years

US Air Force fighter aviator manning: 75% in FY2025, 69% in FY2026, a forecast 67% in FY2027, the lowest of any aircrew community, against a declared shortfall of about two thousand pilots.[2]

Aircraft — 3 to 7 years

Reported F-35 mission capable rates fell from roughly 67% in FY2021 to about 44% in FY2025, driven by depot throughput, spares and maintenance manning.[3]

Hours and airspace — 1 to 3 years

Some 1.1 million flying hours requested for FY2027, described as the maximum executable level for the total force, against an assessed need nearer 1.3 million.[4]

WHY THE ADVERSARY IS CUT FIRST

Squadron commanders facing a contracting programme protect the events that keep aircrew current, because currency is auditable and quality is not. The adversary function is structurally exposed in that calculation, for a reason that has nothing to do with its value. In an organic model, every Red Air sortie is flown by a pilot and an aircraft that could otherwise have flown a Blue Air training event. The adversary is not an addition to the flying programme. It is a subtraction from it, funded out of the same fixed pool of hours that the training is trying to consume.

The consequences accumulate in a direction that is hard to see from a monthly return. Presentations get simpler. The many-versus-few geometry that defines the modern air battle is rehearsed as two versus two, because two versus two is what the availability board allows. The opponent is flown by a squadron mate who shares the same tactics manual, the same airframe and frequently the same instructor lineage, which removes precisely the dissimilarity that gave the event its worth. A training system can lose most of its value this way while every indicator on the dashboard stays green.

> Currency is auditable and quality is not. That single asymmetry explains most of what has happened to the adversary function over the past decade.

THE MEASUREMENT NOBODY HAS

Very few air forces can state, with evidence, how much of their assessed readiness is attributable to the quality of the adversary presentation rather than to the number of sorties flown. The sortie is counted. The presentation is not. Section 7 returns to this at length, because it is the largest single pool of unrealised value in the system as it stands today.

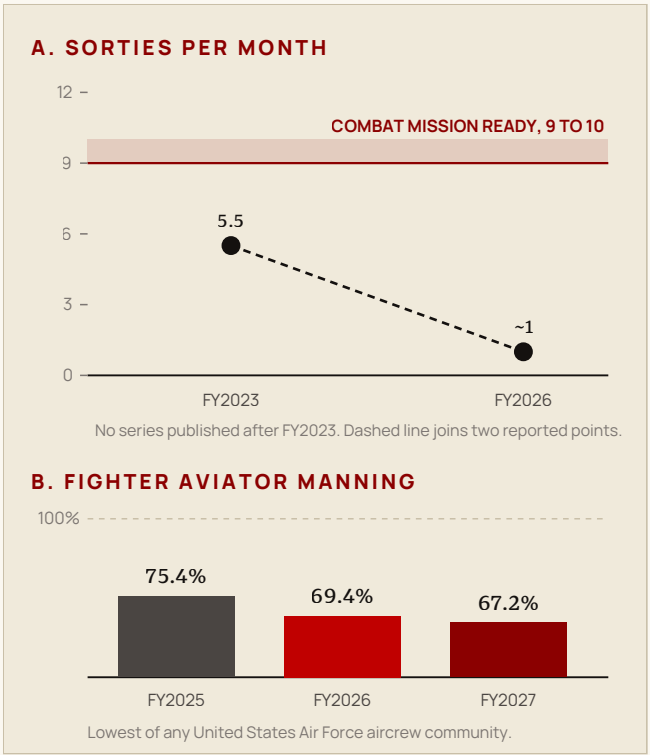


FIGURE 1.1 The readiness scissors. The required sortie rate has not moved. Delivered rate and available manning have both fallen, and because a crew needs both an aircraft and an opponent, the two effects multiply rather than add. Sources at notes 1, 2 and 4.

READINESS IS A RELATIVE MEASURE

None of this would matter if readiness were absolute. It is not. It is measured against an opponent, and the opponent has a vote. Commentary on air operations over the Middle East during 2025 noted Israeli aircrew flying at something in the region of twenty sorties a month, several times the rate of their American counterparts.[5] At the other end of the spectrum, Russian tactical aviation has been widely assessed as training its crews for between seventy and one hundred and twenty flying hours a year, a figure offered repeatedly as part of the explanation for its failure to secure air superiority over Ukraine.[6]

Read together, those comparisons say something more useful than either says alone. Flying hours correlate with tactical competence, but they are a proxy rather than the thing itself. What accumulates in a crew is the number of well assessed decisions taken under pressure against a credible opponent. The flying hour was for a long time the only available proxy, so the two were treated as interchangeable. They no longer have to be, and continuing to do so is now the more expensive error.

THREE REMEDIES, NONE SUFFICIENT ALONE

Buy back the flying hours. The correct instinct, and the slowest lever available. Hours are bounded by maintenance capacity and aircrew availability as much as by appropriation. A supplementary budget does not produce a maintainer, and a maintainer takes longer to train than an aircraft takes to break.

Contract the adversary out. This works, at scale, and over the past decade it has grown into a substantial industry. Section 3 examines it properly. Its limitation is not price. It is representation. A commercial aggressor fleet can put airframes and very experienced pilots into the air. It cannot put a peer radar, a peer electronic order of battle or peer numbers alongside them.

Move the training inside. Also correct, and the direction every serious air force is already travelling. The Royal Air Force has publicly pursued a substantially synthetic mix for Typhoon, beginning with an ambition of a fifty-fifty blend and subsequently describing a shift towards eighty per cent synthetic under its future synthetic training programme.[7] The limitation here is subtler, and it occupies the rest of this paper. A simulator with nothing credible inside it reproduces the cockpit faithfully and the fight not at all.

If the missing repetitions have to be found synthetically, and there is no longer a serious argument that they can be found anywhere else, two questions become unavoidable. The adversary has to be good enough to be worth flying against, and the force has to be able to prove that it was.

SECTION TWO

What the adversary actually delivers

Red Air is usually described as a service, measured in hours and sorties and priced accordingly. That description survives because it is convenient for accountants, and it obscures what the adversary is actually for. Before anything can be substituted for an opponent, it is worth being precise about what an opponent produces.

ASK A WEAPONS INSTRUCTOR WHAT RED AIR IS FOR AND the answer will not be about hours. It will be about what the opposing formation forces the friendly formation to do: to commit or to abort, to accept a merge or to preserve energy, to trust a contact or to doubt it. The adversary is not a target. It is the mechanism by which a crew is made to decide something difficult, quickly, with incomplete information, and then be held to account for the decision afterwards.

The adversary function supplies four distinct products. They are separable, different solutions supply them in different degrees, and the distinction matters when a force considers substitutes. A solution that supplies two of the four well and two badly is not two-thirds of an answer. It is a different training event with a different outcome.

ONE: PRESENTATION

Presentation is the geometry the opponent creates and the way it evolves. It covers formation, spacing, altitude stack, timing of the commit, the range at which weapons are employed and the manner of the subsequent manoeuvre. Presentation is what a crew learns to read, and reading it correctly is the difference between a shot taken and a shot wasted.

Crucially, presentation is threat specific. A formation flown to one doctrine looks different from a formation flown to another well before either does anything overtly hostile, and the difference is legible to a trained eye in the first thirty seconds of a run. This is the root of the dissimilar air combat training argument: an opponent flying friendly tactics in a friendly airframe teaches a crew to recognise itself.

TWO: PRESSURE

Pressure is the cognitive load the opponent imposes. It is created by numbers, by time compression, by simultaneous demands on the same operator, and by the requirement to maintain a picture while doing something else. Pressure is the reason a four versus eight is not simply two four versus fours run in parallel: at some threshold, the crew stops processing every contact and starts triaging, and it is the behaviour above that threshold that the training is trying to shape.

This is also where the live environment fails first under budget pressure, and fails in a way that is almost invisible on paper. Reducing the adversary package from eight aircraft to two does not reduce the training value by three quarters. It removes the specific regime the event existed to rehearse, while leaving the event on the schedule with its original name.

> Cutting the adversary package from eight aircraft to two does not remove three quarters of the value. It removes the regime the event existed to rehearse, and leaves the name intact.

THREE: UNPREDICTABILITY

An opponent that cannot surprise is a demonstration rather than an adversary. Unpredictability is what forces a crew to build a picture rather than recall one, and it is the property most easily destroyed by familiarity. Where Red Air is flown organically by the same squadron week after week, the presentations converge on a repertoire that both sides come to know. The event becomes a rehearsal of an agreed script, executed with professional skill and diminishing instructional return.

Unpredictability is also the property most frequently overpromised by synthetic solutions, and Section 6 examines that claim in detail. A constructive entity that executes a fixed behaviour tree is predictable in a more absolute sense than any human adversary, however good the visual model that renders it.

FOUR: THE ASSESSMENT BASELINE

The fourth product is the least discussed and arguably the most valuable. The adversary defines the standard. A crew's performance is not assessable in the abstract: it is

assessable against what the opponent did, what options that left open, and what the crew chose from among them. Every meaningful measure in a debrief, from timeline discipline to weapons employment to mutual support, is expressed relative to the adversary's actions.

This has an uncomfortable implication. If the adversary is inconsistent from event to event, then performance across events is not comparable, and a force cannot say with confidence whether a squadron is improving. Most air forces accept this silently, because in a live environment there has never been an alternative. A synthetic adversary, whatever else it does or fails to do, can be made identical on Tuesday and on Thursday, and that reproducibility is a capability in its own right.

THE FOUR PRODUCTS, AND WHO SUPPLIES THEM

Presentation — **High value**

Threat-representative geometry, timing and weapons employment. Requires doctrinal knowledge more than it requires an airframe.

Pressure — **Scarce**

Cognitive load through numbers and time compression. Scales with entity count, which is precisely what live flying cannot afford.

Unpredictability — **Fragile**

The capacity to surprise. Degrades with familiarity in organic Red Air and with determinism in poor synthetic Red Air.

Assessment baseline — **Unused**

The reference against which crew performance is judged. Requires consistency between events, which live flying cannot guarantee.

THE HONEST ADVERSARY PROBLEM

There is a structural conflict inside organic Red Air that is rarely stated plainly. The pilot flying the adversary is a colleague, usually an instructor, operating under three constraints that pull against realism.

The first is safety. Adversary flying is conducted inside training rules that constrain closure, minimum separation and manoeuvring in ways that a genuine opponent would not observe. The second is pedagogy. An instructor flying Red Air is trying to produce a specific learning outcome, which frequently means presenting a problem the student can solve rather than the problem an enemy would set. The third is reciprocity. Tomorrow the roles reverse, and the social dynamics of a squadron do not reward an adversary who makes colleagues look incompetent.

None of these is a criticism of the practice. All three are rational and largely unavoidable. Together they mean that organic Red Air systematically under-represents the opponent, in a direction nobody intends and nobody measures.

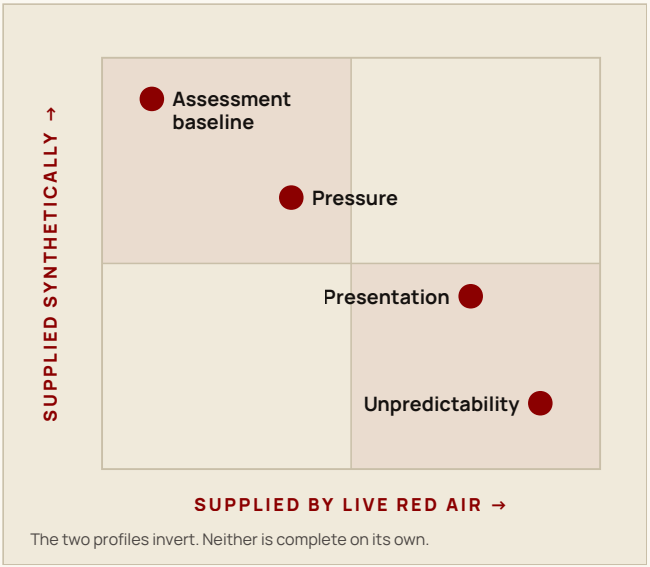


FIGURE 2.1 Presentation and pressure are what forces buy. Unpredictability and the assessment baseline are what they need, and neither appears on a contract schedule.

WHY THIS FRAMING CHANGES THE PROCUREMENT QUESTION

If Red Air is a service measured in hours, the procurement question is who supplies hours most cheaply. That question has a well developed answer, and an industry built around it, examined next.

If Red Air is the four products above, the procurement question becomes considerably more awkward. It asks which products a force is actually short of, which of them a given solution supplies, and at what fidelity. A contracted aggressor squadron supplies presentation and, at a price, pressure. It supplies unpredictability well, because the pilots are genuinely adversarial and are not flying with the friendly squadron tomorrow. It supplies almost nothing in the way of an assessment baseline, because no two sorties are the same and the data is not captured in a form that supports comparison.

A synthetic solution inverts that profile almost exactly. It supplies pressure abundantly and cheaply, because entities cost software rather than airframes. It supplies an assessment baseline natively, because everything is instrumented by construction. It supplies presentation only as well as the doctrinal knowledge encoded in it, which is a research and curation problem rather than a graphics problem. And it supplies unpredictability only if someone has done difficult work that most implementations skip.

Neither profile is complete. That is the argument for a blended capability rather than a replacement, and the reason this paper concerns itself with what to build rather than what to buy.



PLATE 1 · THE DISSIMILARITY PROBLEM

An opponent flying friendly tactics in a friendly airframe teaches a crew to recognise itself. Dissimilarity is not a refinement of adversary training. It is the reason the training exists.

SECTION THREE

The contracted adversary and its ceiling

Over a single decade, contracted adversary air went from an experiment at one American base to a multi-billion dollar industry underwriting fighter training across a continent. It is the most successful intervention in air combat training in a generation. It is also, on its own, structurally incapable of closing the gap it was created to address, and against the uncrewed threat it was never designed to try.

THE MODERN INDUSTRY HAS A TRACEABLE STARTING POINT. In late 2015 the United States Air Force, having lost aggressor capacity at Nellis, awarded a short proof of concept contract for contractor-supplied adversary support to its Weapons School.[8] The experiment worked, in the narrow and important sense that it freed organic fighter squadrons from generating Red Air for each other and returned those sorties to Blue Air training.

What followed was rapid. By 2019 the requirement had been consolidated into a Combat Air Force contracted air support programme with a ceiling reported at 6.4 billion dollars across multiple suppliers, covering tens of thousands of adversary hours a year at a dozen fighter bases.[9] A second iteration followed in September 2024, reported at 5.4 billion dollars with eight companies selected to support Air Combat Command.[10] The United States Navy, which had used contracted adversary support longer and more modestly, expanded in parallel.

Alongside the volume providers, the market produced a premium tier. Top Aces became the first operator of contractor-owned, contractor-operated F-16s, fielding an advanced aggressor configuration explicitly to present a more representative threat to fifth generation fighters under a five year contract reported at up to 175 million dollars.[11] The trajectory of the industry is legible in that single procurement: from cheap airframes generating cheap hours, towards expensive airframes attempting to generate representative threats.

WHAT THE MODEL SOLVED

Three things, and it solved them properly.

Throughput. Adversary hours became available in quantity without consuming organic fighter capacity. For a force whose binding constraint is aircraft availability, converting a Red Air sortie into a contracted one is a straight gain of a Blue Air sortie.

Experience. Contracted aggressor cadres are staffed largely by former aggressor and weapons school pilots. The problem identified at the end of the last section, the colleague who does not wish to embarrass a colleague, largely disappears when the adversary is a different organisation with a different incentive.

Cost per hour. A light attack aircraft or a regenerated third-generation fighter flown by a civilian operator costs a fraction of a front-line fighter hour, and the arbitrage is real. It is worth noting that this arbitrage is the entire commercial logic of the sector, and that it narrows sharply as the threat representation demanded goes up.

> The commercial logic of contracted Red Air is arbitrage between a cheap airframe and an expensive one. That arbitrage narrows towards zero as the threat you need represented becomes modern.

WHERE THE CEILING IS

The limitation is not price and it is not professionalism. It is representation, and it has four components that compound.

Sensors and emissions. A credible modern adversary is defined less by its airframe than by what it radiates and what it can see. Presenting a peer radar, a peer electronic attack capability and a peer emitter signature requires either the actual equipment, which is generally not exportable to a commercial operator, or a pod-based surrogate whose fidelity is bounded by what a pod can do. Aircraft performance is the easiest part of the threat to replicate and the least important.

Networked behaviour. Peer air forces do not fight as individual aircraft. They fight as a system, with ground-controlled intercept, integrated air defences, datalinked pictures and airborne early warning. A contracted adversary flight can be briefed to imitate the tactics that system produces. It cannot reproduce the system, and it is the system that generates most of the difficulty.

Numbers. Pressure scales with entity count, and entity count in a contracted model scales with money at a punishing rate. Presenting twenty adversaries against four friendlies is affordable perhaps once a year inside a flagship exercise. The requirement is weekly.

Classification. The most representative threat presentations depend on intelligence that cannot be handed to a commercial operator without a security architecture that largely defeats the commercial premise. The United States Air Force has explored classified addenda to adversary air solicitations for exactly this reason.[12] Every step towards realism adds clearance, facility and handling cost.

WHAT A PEER PRESENTATION REQUIRES, AND WHO CAN SUPPLY IT

Airframe performance Energy, turn rate, acceleration. Contracted fleets supply this adequately and improving.	**Solved**
Sensor and emitter behaviour Radar modes, electronic attack, signature. Bounded by exportability and by what a pod can represent.	**Partial**
Networked system behaviour Ground control, integrated air defence, datalinked picture, airborne early warning acting as one.	**Not feasible**
Numbers at weekly cadence Twenty-plus entities against a four-ship, every week rather than once a year.	**Not affordable**

THE EUROPEAN POSITION IS DIFFERENT, AND WORSE

It is important not to read the American experience as a template. The contracted adversary market developed in the United States because three conditions held simultaneously: a very large and predictable demand signal, a deep pool of surplus tactical airframes with exportable configurations, and ranges large enough to fly them in. None of the three holds in Europe in the same degree.

European demand is fragmented across national programmes of a size that struggles to sustain a dedicated commercial aggressor fleet. Surplus airframes are fewer and the regulatory path to operating them commercially is heavier. Most decisively, the training airspace available to most European operators cannot accommodate the geometry of a modern beyond visual range engagement at all, let alone one with meaningful adversary numbers in it. A European air force that decides to solve its adversary problem by contracting is solving it on someone else's terrain, at someone else's ranges, on a deployment basis, a handful of times a year.

This is not an argument against contracted adversary air in Europe, which has a clear role. It is an argument that the model cannot be the backbone of European adversary generation in the way it has become the backbone of American adversary generation, and that a European force planning on that basis is planning on an assumption that does not hold.

COUNTER-UAS PERSPECTIVE

THE THREAT NO AGGRESSOR FLEET WAS BUILT TO FLY

Everything above concerns an adversary that flies like an aircraft. The threat that has grown fastest since 2022 does not. Launches of long-range one-way attack drones and decoys against Ukraine peaked at 8,161 in a single month, May 2026.[13] On the night of 13 to 14 April 2024 Iran sent some 170 drones towards Israel together with cruise and ballistic missiles; Israeli fighters shot down dozens of them and United States aircraft between seventy and eighty,[14] and Royal Air Force Typhoons destroyed a number of others.[15] In September 2025 Dutch F-35s shot down drones over Poland, which SHAPE described as the first time NATO aircraft had engaged potential threats in Allied airspace.[16]

8,161 One-way attack drones and decoys launched against Ukraine in May 2026, the record month	**170** Iranian drones launched towards Israel in a single night in April 2024, with cruise and ballistic missiles
1,000+ Drones launched from Iran towards Israel in June 2025, 99 per cent of them intercepted	**1 bn+** Dollars, one estimate of what Israel's defence against the whole attack, missiles included, cost that night

The uncrewed threat in four numbers. Sources: notes 13, 14, 17 and 18.

The fighter has therefore acquired a counter-drone mission, and with it a cost problem. Open estimates put a Shahed-type drone at between twenty and fifty thousand dollars.[19] In 2025 the United States Air Force described a laser-guided rocket, whose guidance section costs fifteen to twenty thousand dollars, as its primary weapon against drones in the Middle East, rather than air-to-air missiles costing around 450,000 dollars and more. Its commanders noted that the rocket has to hit the drone directly, and that carrying it costs fuel.[20] The weapon changed. So did the tactics, the identification problem and the fuel plan, and each of those has to be trained.

Here the ceiling described above stops being economic and becomes absolute. Contracted aggressors fly aircraft; they do not fly drone salvos. Live surrogate drones exist, and the National Training Center in the United States has a threat-replication system able to swarm up to 150 drones from a single ground station,[21] but in the authors' assessment such events are rare and scheduled. In

Europe, civil rules confine open-category flight to visual line of sight and 120 metres, with anything beyond requiring a specific operational authorisation,[22] while military drone regulation remains a national competency. An article in the IAMD COE's 2025 conference journal, published in 2026, identifies airspace integration as the primary barrier to using uncrewed aircraft as aggressors for air defence training in Europe.[23] Threat-representative decoys of the kind used against Ukraine are not available to fly at a range, and outside tightly controlled trials no authority will allow a hostile profile to be flown over a working airfield.

WHAT A DRONE PRESENTATION REQUIRES, AND WHO CAN SUPPLY IT

Numbers Dozens to hundreds of contacts in one picture, at the cadence of the war that defines the problem.	**Rare live**
Representation The signature, speed and behaviour of the real threat, and decoys built to deceive. A surrogate drone is not the threat.	**Partial**
Airspace Flight beyond line of sight, at the heights and on the routes the threat uses, close to defended sites.	**Restricted**
The home airfield Hostile profiles over a working base in normal operations, with civil and military traffic in the picture.	**Not permitted**

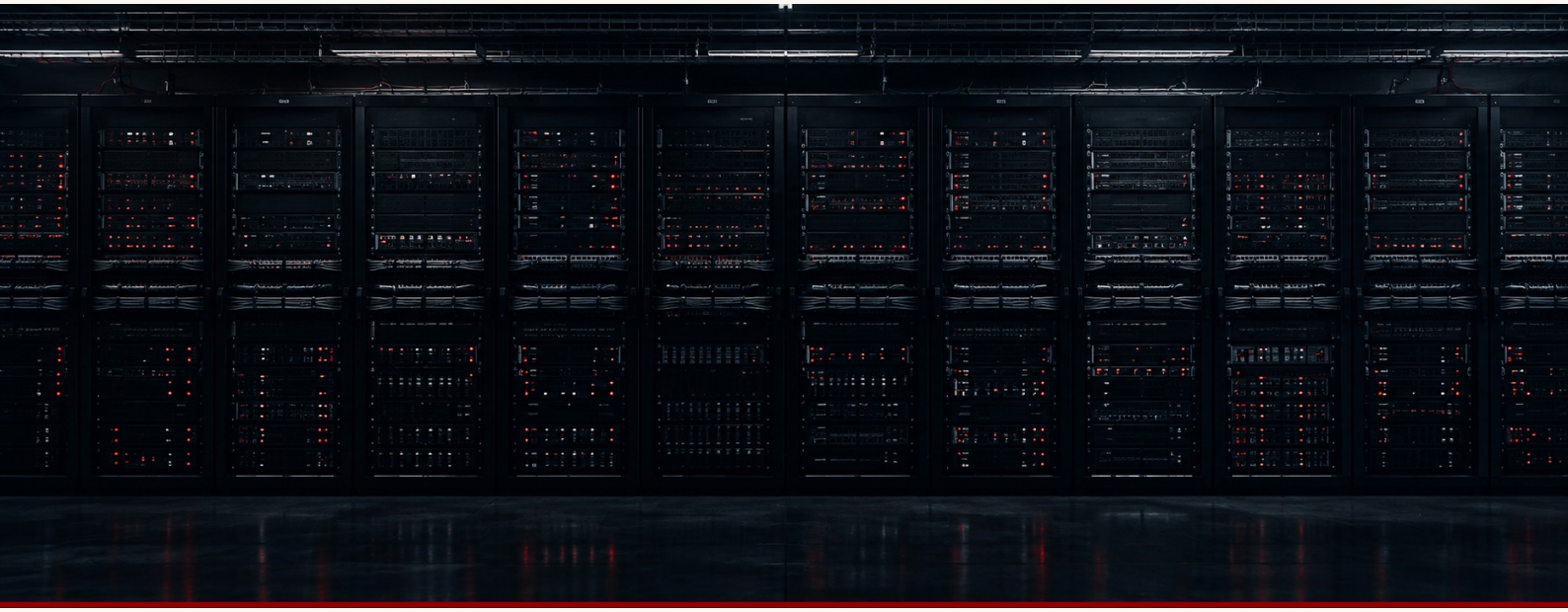
For the drone, then, the synthetic environment is not the cheaper option. For much of what matters it is the only one.

THE LIVE AUGMENTATION ROUTE

A third position has emerged between the contracted aircraft and the simulator, which is to leave the crew in the real aircraft and insert the adversary optically. Airborne augmented reality systems now place manoeuvring virtual opponents into a pilot's field of view during live flight. The approach moved from demonstrator to procurement in 2025, when a contract was awarded through Air Combat Command and the Air Force Research Laboratory to integrate such a system into the F-16, having previously been fielded on the T-38, the MC-130 and the Hawk T2.[24] A subsequent agreement in April 2026 extended the approach to a widely operated European lead-in fighter trainer.[25]

The significance is not the technology. It is what the procurement concedes. An air force that buys synthetic adversaries for its live aircraft has accepted, as a matter of funded policy, that the opponent no longer has to be a physical object. Once that concession is made, the question is no longer whether the adversary should be generated. It is where the generation should happen, how good the generated adversary has to be, and who owns the models that define it.

Those are the questions the rest of this paper addresses. They are harder than the procurement question, and they have the advantage of being answerable by a mid-sized air force rather than only by a superpower.



SECTION FOUR

Live, virtual and constructive: the working definitions

The vocabulary of synthetic training is used loosely enough that two officers can agree on a requirement and mean different capabilities. This section fixes the terms, sets out the standards that govern them, and identifies where the real cost of an LVC capability sits. It is not where most programmes assume.

THE TAXONOMY IS OLDER THAN THE TECHNOLOGY AND HAS survived because it discriminates usefully. Live means real people operating real systems. Virtual means real people operating simulated systems. Constructive means simulated people operating simulated systems. The distinction is about where the human sits, not about how convincing the graphics are.

LVC, the blended case, means an exercise in which entities of all three kinds occupy one shared synthetic battlespace and can perceive and affect one another. A pilot airborne in a real aircraft, a crew in a simulator three hundred kilometres away and a package of computer-generated adversaries all appear in the same picture, on the same timeline, with consistent positions and consistent weapon effects. When that works, the training economics change fundamentally, because constructive entities are the only component whose marginal cost approaches zero.

THE STANDARDS THAT MAKE FEDERATION POSSIBLE

Federation is a solved problem in principle and a difficult one in practice. Four families of standard carry most of the load, and a force procuring synthetic training capability should understand which of them its suppliers actually implement.

Distributed Interactive Simulation, standardised as IEEE 1278, defines a set of protocol data units broadcast between simulators. It is old, it is simple, and it remains the most widely implemented method of getting heterogeneous simulators to see each other. Its simplicity is also its limitation: it broadcasts state rather than managing it, and it scales awkwardly as entity counts rise.

The High Level Architecture, standardised as IEEE 1516, replaces broadcast with a managed federation, where a run-time infrastructure mediates what each participant publishes and subscribes to. It scales better and costs more to implement, and the federation object model has to be agreed before anything interoperates.

Test and Training Enabling Architecture emerged from the range community and addresses the live side, where truth data from instrumented aircraft has to enter the synthetic picture with useful latency.

Tactical datalinks, principally Link 16 and its internet protocol transport variants, are what allow a synthetic picture to reach a real cockpit or a real command and con-

trol node. This is the bridge that converts a simulation into LVC, and it is also where classification bites hardest.

WORKING DEFINITIONS

Live — L

Real people, real systems. An aircraft in the air, instrumented and contributing its truth position to the exercise.

Virtual — V

Real people, simulated systems. Aircrew in a simulator, making the same decisions under the same timeline.

Constructive — C

Simulated people, simulated systems. Entities generated and behaved by software, supervised rather than flown.

Distributed — D

Any of the above, at separate physical locations, federated over a network into a single exercise.

CORRELATION: THE FAILURE NOBODY BUDGETS FOR

Two simulators can exchange perfectly valid protocol data units and still produce a worthless exercise. The reason is correlation, and it is the most consistently underestimated risk in distributed training.

> Simulators are procured, and federations are negotiated. The expensive part of an LVC capability is the agreement, not the equipment.

Correlation means that all participants share the same world. The same terrain elevation at the same coordinates, so that an aircraft masked by a ridge in one simulator is masked in all of them. The same atmospheric and illumination conditions, so that a detection range computed in one participant is defensible in another. The same weapon flyout and lethality models, so that a shot declared by one federate is accepted by the entity receiving it. The same time reference, to a tolerance that matters.

Where correlation fails, the failure is rarely visible as an error message. It presents as an argument in the debrief, in which two crews remember the engagement differently and both are looking at valid data. That is more corrosive than an outright technical failure, because it systematically destroys trust in the synthetic environment among exactly the experienced aircrew whose endorsement the capability needs.

LATENCY AND WHAT IT BREAKS

Distributed training has a latency budget, and different training objectives sit at very different points in it. Command and control training, where the tempo is measured in tens of seconds, tolerates network delay that would make close-in manoeuvring training actively harmful. Beyond visual range work sits in between. Within visual range manoeuvring against a distributed opponent is the hardest case in the entire discipline, and it is the one most often promised in early programme documentation and quietly dropped later.

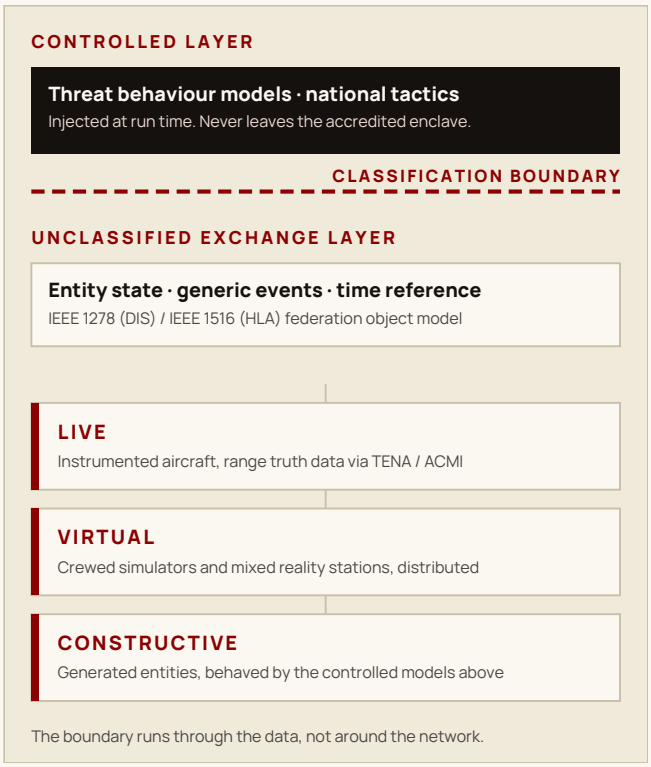


FIGURE 4.1 The classification boundary, not the network boundary, determines what a federation can do. Architectures that place it late are expensive to correct.

The practical consequence is that a force should specify its training objectives before its architecture, and should be explicit about which objectives it is not attempting to meet distributively. A capability that delivers excellent distributed large force employment training and makes no claim about distributed basic fighter manoeuvres is a coherent capability. A capability that claims both usually delivers the first and damages the second.

WHAT WORKING SYSTEMS LOOK LIKE

Two national examples illustrate the range of what is achievable, and both are documented openly.

The United Kingdom's distributed synthetic training capability entered service in 2022 at a purpose-built facility, linking simulators across the air, land and maritime environments into a single secure hub, with Typhoon, F-35, the Rivet Joint fleet and a joint fires trainer among the early integrations.[26] It has since been used for virtual large force employment exercises and for weapons instructor courses conducted ahead of live large force events, at a scale its operators state is difficult to achieve in the live environment.[27] The associated Typhoon synthetic training programme began with an ambition of a fifty-fifty live and synthetic blend and was subsequently described as moving towards eighty per cent synthetic.[28]

At Alliance level, the direction is the same. NATO's principal air exercise in 2026 combined live flying with synthetic training across two joint operations areas, involving nineteen nations and more than one hundred and fifty aircraft, at a sortie rate no single nation could generate alone.[29] Allied Command Transformation's current training initiative includes an explicit line of effort on new technologies, with development and testing of digitally and artificially intelligent enabled exercise design scheduled through 2026.[30]

Two features are common to both, and they are the features worth copying. Each is standards based rather than vendor specific, so that new platforms join a federation rather than triggering a new procurement. And each treats the synthetic environment as a persistent capability that is operated continuously, rather than as a facility that is stood up for an exercise and stood down afterwards. The second point is the more important one, because it determines whether the capability accumulates anything.

⌘ COUNTER-UAS PERSPECTIVE

THE TACTICAL GAP IN COUNTER-UAS TRAINING

Equivalent systems are harder to find for the crews who defend against drones. The Alliance tests counter-UAS equipment every year: the 2024 Technical Interoperability Exercise brought together more than 450 participants from nineteen Allies and three partners,[31] and the 2026 edition tested more than sixty systems.[32] Exercises have begun to blend the domains, and in March 2025 one used distributed interactive simulation and hardware in the loop to integrate counter-UAS into air and missile defence training.[33] Individuals are trained in growing numbers: the Joint Air Power Competence Centre's counter-UAS course reported nearly 700 students in 2025 and capacity for 1,500 in 2026.[34]

What is missing is the crew in its tactical environment: a team on its own consoles, rehearsing detection, identification and decision against a realistic picture, often enough to build proficiency and measured well enough to know whether it has. Observers at the United States Army's National Training Center reported in 2025 that some soldiers fail to notice drones at all while others see them and hesitate, and that counter-drone training had been relegated to a secondary role, thrown in a corner and forgotten about.[21] An officer of the Joint Air Power Competence Centre, writing in its journal in 2025, concluded that traditional training programmes have not sufficiently prepared personnel to identify and classify low, slow and small drones.[35]

The pattern is consistent. The equipment is tested every year, concepts are validated in occasional large exercises, and individuals are taught in courses of days or weeks. The crew, in the tactical environment where the decision is actually made, is rehearsed almost nowhere. That is the gap a persistent synthetic environment exists to close, and Section 10 returns to who might close it.

WHAT THIS SECTION HAS NOT YET ADDRESSED

Everything above concerns plumbing. A perfectly federated, perfectly correlated, low latency distributed environment with nothing credible inside it is an expensive way to fly formation against nothing. Two questions remain, and they are the ones on which the value of the whole enterprise turns.

The first is how good the individual training device has to be before a crew can transfer what it learns to the aircraft, and what happens when it falls below that threshold. The second is how an adversary can be generated inside that environment which is worth the crew's attention. They are addressed in the two sections that follow.

SECTION FIVE

Fidelity, negative training and the limits of immersion

Fidelity is discussed as though it were a dial that can be turned up until the training becomes real. It is not a dial. It is a bundle of separable properties, each with a threshold below which the device stops teaching and starts teaching the wrong thing. Those thresholds are now measurable, and most of them are now met.

THE HISTORICAL OBJECTION TO HEAD-MOUNTED TRAINING IS WELL FOUNDED and should not be waved away. For most of three decades, attempts to train pilots in head-mounted displays produced disappointing and sometimes harmful results. Users became nauseated. Text could not be read. Reaction times did not transfer. The instructional community drew a reasonable conclusion from repeated evidence, and a generation of senior aircrew formed a view that has not been revisited since they last tried a device.

What has changed is not the concept but the specification. Each of the historical failures traces to a measurable parameter, and each of those parameters has crossed a threshold within the past five years. That makes the objection answerable with numbers rather than enthusiasm, which is the only way it is ever going to be answered.

ANGULAR RESOLUTION, AND WHY IT IS THE FIRST THRESHOLD

Human vision at the commonly cited 20/20 standard resolves detail at approximately sixty pixels per degree of field of view. Below that figure, a virtual object is being presented at lower acuity than the eye could achieve in reality, and the consequences are specific rather than aesthetic. An aircraft at tactical range disappears or pixelates when it would have been visible. Instrument labels and head-up display symbology become unreadable at the distance they would actually be read.

Professional devices now sit at or above that figure in the region of the visual field where it matters. The previous generation of professional headset delivered roughly seventy pixels per degree across a foveal region of about twenty-seven degrees, with the peripheral field at above thirty.[36] The current generation delivers up to fifty-one pixels per degree in video passthrough with gaze-driven autofocus, with fixed-focus passthrough at around thirty-three.[37] The distinction matters for mixed reality work in a physical cockpit, and the supplier's own guidance is unusually candid about the consequence of getting it wrong: at lower passthrough acuity, trainees adopt compensating behaviours, moving closer, leaning in and squinting to read instrument markings, which is precisely the behaviour that must not be learned.[38]

> A trainee who leans in to read an instrument has learned something. It is not what the syllabus intended, and nothing in the training record will show it.

LATENCY, AND THE TWO BUDGETS THAT ARE NOT THE SAME

Two latency figures govern a mixed reality training device and they are routinely conflated. Motion-to-photon latency is the delay between head movement and the corresponding change in the rendered scene, and it is the dominant cause of simulator sickness. Passthrough latency is the delay between the real world and its video reproduction, and it governs whether interaction with physical controls feels natural. Professional devices now report passthrough figures well inside the twenty millisecond range at which the discrepancy becomes perceptible.

Interpupillary distance is the third parameter and the least glamorous. Evidence circulated within the sector attributes a substantial share of simulator sickness cases, commonly cited at around forty per cent, to incorrectly set separation between the optical centres. Current professional devices adjust the setting automatically on donning, across a range of roughly fifty-six to seventy-two millimetres.[39] A problem that consumed years of research effort has been largely engineered out of existence by a motor and an eye tracker.

THE THREE WAYS TRAINING FAILS

It is worth being precise about failure modes, because they are not equivalent and only one of them is dangerous.

Omission occurs when a skill is simply not taught, or is taught but fails to transfer to the aircraft. The cost is an opportunity cost. The crew is no worse than before.

Negative transfer occurs when the training environment slows subsequent learning in the real environment, because something has to be unlearned first. The cost is time, and it is recoverable.

Negative training occurs when the crew acquires a behaviour that is actively wrong, and carries it into the aircraft without knowing it. This is the failure that matters. Its distinguishing feature is that it is invisible: the syllabus was completed, the event was assessed satisfactory, and the defect only surfaces under conditions nobody has yet flown.

PARAMETERS, THRESHOLDS, AND THE FAILURE EACH PREVENTS

Angular resolution	At or above roughly 60 PPD in the region of interest. Below it: missed detections at realistic range, unreadable symbology, compensating head movement.	Met
Motion-to-photon latency	Low enough that head movement and scene change are not separable. Below it: simulator sickness and shortened sessions.	Met
Passthrough latency	Inside about 20 ms. Below it: unnatural reaction times learned against physical controls.	Met
Interpupillary setting	Automatic, per user, on donning. Below it: a large fraction of avoidable sickness cases.	Met
Control correlation	Virtual switches must match real ones in position and feel, or be physical. Below it: negative training, the failure that does not announce itself.	Design choice

The final row is the one that has not been solved by hardware, and it will not be. Whether a crew can reach the correct switch without looking depends on whether the switch is where the aircraft puts it and feels like the aircraft's switch. That is a decision about the cockpit shell, not about the headset, and it is the point at which mixed reality earns its additional complexity: the trainee sees real hands operating real controls, with the synthetic world beyond the canopy rail.

THE MEASUREMENT DIVIDEND

One consequence of the current generation of devices is routinely underestimated. Eye tracking was introduced to enable foveated rendering and automatic optical adjustment. It now runs at two hundred hertz with accuracy of around one degree as a standard feature.[40] That is an instructional instrument that has no equivalent in the live environment.

A fundamental airmanship requirement is the visual scan: clearing the airspace at intervals, checking specific indications at specific points, completing a checklist by looking at the thing being checked rather than reciting it. In an aircraft, and in a conventional simulator, an instructor can only infer whether this happened. With gaze data, it is observable, reviewable and comparable between events and between crews. Section 7 develops the point, but it belongs here as well, because it changes the argument about fidelity. The synthetic environment is no longer competing with the aircraft on realism alone. On some instructional dimensions it is now strictly superior, and pretending otherwise concedes ground that does not need conceding.

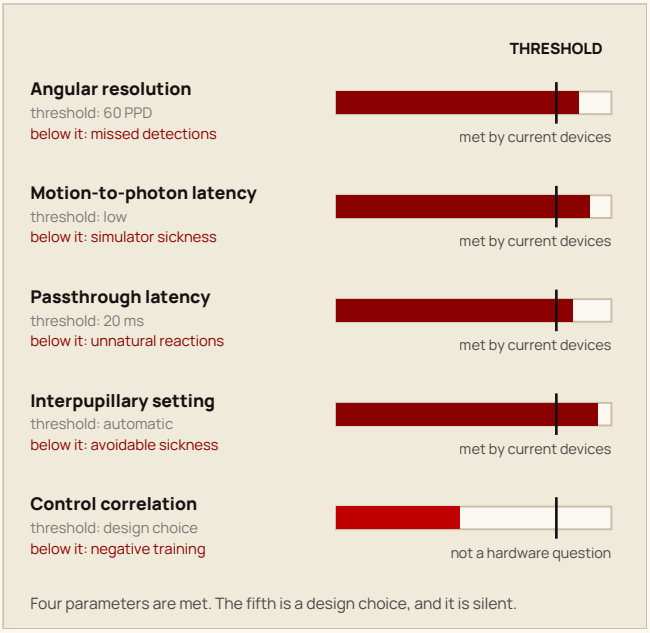


FIGURE 5.1 Fidelity is not a single dial. Each parameter has an independent threshold and a distinct failure mode below it, and a device can be excellent on four and disqualifying on the fifth.

THE REGULATORY PRECEDENT, AND THE GAP IT EXPOSES

Civil aviation has already crossed this ground, which is instructive because it had to justify the crossing formally. In April 2021 the European Union Aviation Safety Agency granted the first qualification of a virtual reality based flight simulation training device, at Flight and Navigation Procedures Trainer level II for a light helicopter, having applied special conditions to existing certification specifications to account for the display, cockpit and motion characteristics of the new technology.[41] The qualification allowed time in the device to be credited against licence requirements. A subsequent device for a single-engine tur-

bine helicopter was qualified at flight training device level and approved for proficiency checks and type ratings.[42]

Two features of that process deserve attention from military training authorities. The first is that the regulator did not lower the standard. It held the required training outcome constant and adapted the compliance method, which is the only intellectually defensible way to qualify a new technology. The second is that the evaluation was conducted with operational pilots and instructors rather than with engineers, and the qualification rested on demonstrated training effect.

WHAT THIS MEANS FOR A SPECIFICATION

The practical consequence for anyone writing a requirement is that fidelity should never appear in it as an adjective. Words such as high fidelity, immersive and realistic are unfalsifiable, and a supplier that meets them has met nothing. Each parameter above is measurable, and each should be stated as a number tied to the training objective it protects.

A force that wants crews to identify aircraft type at tactical range should specify angular resolution in the relevant part of the field of view and then test it with aircrew rather than with a chart. A force that wants procedural training against physical controls should specify passthrough latency and hand occlusion, and should decide explicitly whether the controls are virtual or physical. A force that intends to measure visual scan should specify gaze sampling rate and accuracy, and should require the data in an exportable form rather than only inside a vendor's review tool. None of these is exotic, and all of them are routinely omitted.

FIVE QUESTIONS THAT SEPARATE A SPECIFICATION FROM A WISH

Which training objective does each fidelity parameter protect, and what happens to that objective if the parameter is missed by ten per cent?

Who conducts acceptance, and are they instructors rather than engineers?

Are the controls the crew touches physical, virtual, or a mixture, and has anyone flown the mixture?

In what format does the device export what it observed, and who owns that data?

What is the evidence that a crew trained on this device performs differently in the aircraft, and when will it be collected?

The last of those questions is the one that is almost never asked before contract award, and it is the only one whose answer would settle the argument. Transfer of training studies are expensive, slow and methodologically awkward, which is why forces prefer to argue from device specification instead. That preference is understandable and it is the reason the debate has run for twenty years without resolving.

Military synthetic training has no equivalent framework. There is no accepted means by which a national air force can state that a given device, with a given adversary model inside it, is qualified to deliver a given training objective to a given standard. In the absence of such a framework, every adoption decision is argued on the personal judgement of whoever is currently in post, and every adoption decision is therefore reversible by their successor. Section 9 returns to this, because it is a governance problem that the technical community cannot solve on its own and that an alliance-level body is unusually well placed to address.

SECTION SIX

Threat mass: generating an opponent worth fighting

Putting forty adversaries into a synthetic exercise is trivial and costs almost nothing. Making any one of them behave like a competent enemy is the hardest problem in the discipline, and it is the problem on which the value of the entire synthetic training enterprise depends.

ENTITY COUNT IS NOT THREAT MASS. This is the single most common category error in synthetic training procurement, and it survives because entity count is easy to demonstrate and easy to write into a requirement. A vendor can populate a scenario with two hundred aircraft in an afternoon. Whether they impose anything on the crew beyond screen clutter is a different question, and counting them does not answer it.

Threat mass, properly defined, is the cognitive and tactical load a crew must absorb. It is generated by opponents that make choices the crew did not anticipate, at times the crew did not choose, in a manner consistent with a doctrine the crew is expected to recognise. Entity count contributes to it. Entity count alone does not constitute it.

THE LADDER OF ADVERSARY GENERATION

Constructive adversaries fall into three broad classes, and most fielded systems sit on the first rung while describing themselves as being on the third.

Scripted entities follow a predetermined route and execute predetermined actions at predetermined trigger points. They are entirely deterministic. They have a legitimate and permanent role: they populate a scenario, they generate radar returns, they occupy the crew's attention, and they are ideal for procedural training where the point is the crew's process rather than the opponent's ingenuity. They cannot supply unpredictability, and a training system built on them exclusively will be defeated by familiarity within a few months.

Reactive entities select behaviours from a repertoire according to the tactical situation, typically through behaviour trees, finite state machines or rule sets encoding doctrinal tactics. They will commit, abort, defend, re-attack or drag according to geometry and threat state. This is where the serious engineering work sits, and where doctrinal knowledge rather than software skill becomes the binding constraint. A reactive adversary is only as good as the tactical understanding encoded in it, which means the critical resource is not a developer but a subject matter expert with current threat knowledge and the patience to formalise it.

Learning agents derive behaviour from training against an objective rather than from explicitly authored rules. The public evidence here is now substantial. The Air Combat Evolution programme demonstrated machine learning agents controlling a modified F-16 in within visual range engagements against a human-piloted F-16, with demonstrations conducted through 2023 and continuing into 2024.[43] Reporting on the programme describes something in excess of one hundred thousand lines of flight-critical software change across twenty-one test flights in the transition from installation to demonstration.[44] The capability has since been extended, with the test aircraft taking on an autonomous intercept role in 2026.[45]

THREE CLASSES OF CONSTRUCTIVE ADVERSARY

Scripted	Deterministic route and actions. Populates scenarios and generates workload. Defeated by familiarity within months.	Mature
Reactive	Doctrine encoded as behaviour rules. Commits, aborts, defends, re-attacks on geometry. Constrained by subject matter expertise, not by software.	The real work
Learning	Behaviour derived from training against an objective. Demonstrated in flight against human pilots. Hard to explain, hard to constrain, hard to assure.	Emerging

THE AGENT THAT WINS IS NOT THE AGENT THAT TEACHES

There is a seductive assumption in the learning agent literature that an adversary optimised to win is the ideal training opponent. It is not, and the distinction is fundamental to anyone designing a training system.

An agent optimised purely to defeat its opponent will converge on whatever exploits the simulation permits.

Some of those exploits are tactically meaningful. Many are artefacts of the model, and a crew that learns to counter them has learned nothing transferable. Worse, a relentless and unbeatable opponent produces demoralisation rather than instruction: the crew stops experimenting, reverts to conservative play, and the training value collapses in a way that is well documented in the wider pedagogical literature.

A training adversary has to be optimised against a different objective, which is to fight to a defined standard. It should be beatable by correct execution and unforgiving of incorrect execution. It should behave consistently enough that two crews facing it can be compared, and vary enough that neither can predict it. It should be capable of being dialled to a proficiency level, because a crew on its third event and a weapons instructor candidate need different opponents. None of these is a natural product of optimisation. All of them have to be designed in deliberately.

> An adversary optimised to win converges on whatever the simulation permits. An adversary optimised to teach fights to a standard, and loses when it should.
>
> SECTION 6 · THE DESIGN OBJECTIVE

WHO CERTIFIES THAT THE MODEL IS RIGHT

This is the question that separates a serious capability from a procurement. A behaviour model asserts how a particular opponent fights. That assertion is either correct, in which case it is a valuable and sensitive national asset, or it is incorrect, in which case an air force is systematically training its crews against a fiction.

Validating such a model is genuinely hard, because there is no ground truth available. Nobody can fly the actual opponent and compare. What is available is a disciplined substitute, made of four parts.

Provenance: every behaviour traces to an identified source, doctrinal, intelligence or expert, recorded alongside it. **Expert adjudication**: a standing panel reviews behaviours against current assessment, with disagreements recorded rather than resolved silently. **Configuration control**: the model is versioned, so a force can state which release its crews trained against. **Behavioural envelopes**: the model is tested to confirm it stays within stated performance bounds and does not, under unusual geometries, do something no opponent would do.

Taken together these constitute a threat model authority, and it is fundamentally an institutional function rather than a technical one. It sits naturally with a body that already holds doctrinal and threat expertise across several nations, that is trusted to arbitrate between them, and that has no commercial interest in the outcome. Alliance Centres of Excellence are an obvious candidate, and the absence of such an authority is currently a more serious obstacle to adoption than any remaining technical limitation. Section 10 considers the one best placed to take it on for air and missile defence.

TRAINING TO THE MODEL

A final risk deserves naming because it is the one most likely to be discovered late. Any consistent adversary can be learned. Crews exposed repeatedly to the same behaviour model will, without intending to, develop tactics optimised against the model rather than against the threat it represents. This is the familiarity problem of organic Red Air in synthetic form, and it is worse, because the synthetic opponent is more consistent.

Three countermeasures exist and all three should be used: stochastic variation within doctrinally valid bounds; model rotation, so crews meet several variants rather than one canonical opponent; and periodic adversarial review, in which experienced aircrew try to exploit the model and their successes are logged as defects rather than admired as clever flying.

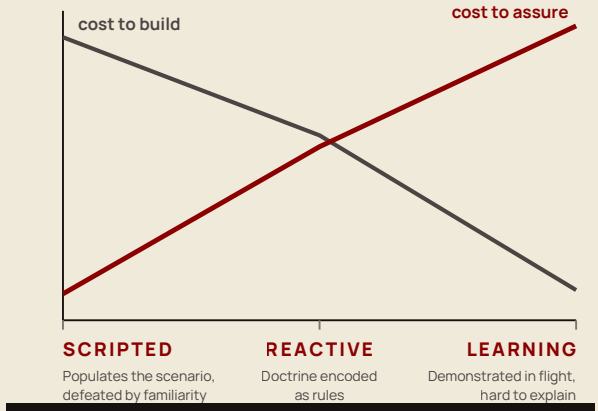


FIGURE 6.1 Cost of construction falls along the spectrum while cost of assurance rises. Most programmes budget for the first curve and are surprised by the second.

⌘ COUNTER-UAS PERSPECTIVE

THREAT MASS DOES NOT END AT THE MERGE

Threat mass has so far been described as what a crew meets in the engagement. The uncrewed threat adds a second kind, which arrives where no crew is looking for an opponent: over its own airfield. In December 2023 drones

flew over Langley Air Force Base, home to F-22 fighters, for seventeen days; the base cancelled night training and moved its aircraft elsewhere.[46] In September 2025 drones closed Copenhagen airport for around four hours and, two nights later, Aalborg, an airport shared by civil and military traffic, where they stayed overhead for about three hours; drones were also reported near Skrydstrup, home of the Danish fighter fleet.[47] In June 2025 small drones launched from trucks deep inside Russia struck military aircraft parked at four air bases, among them long-range bombers. Ukrainian authorities claimed forty-one aircraft hit; United States officials later put the figure at up to twenty, about ten of them destroyed.[48]

None of this can be rehearsed live at operational realism. Outside tightly controlled trials, no authority will allow drones to be flown over a working airfield to exercise its recovery procedures, and few air forces would close their own base in order to practise closing it. A synthetic environment can do both, presenting the drones to the base defence crew, the closure to the controllers and the consequences to the formation on recovery, in one exercise and on one timeline. Threat mass, properly generated, includes the disruption as well as the opponent.

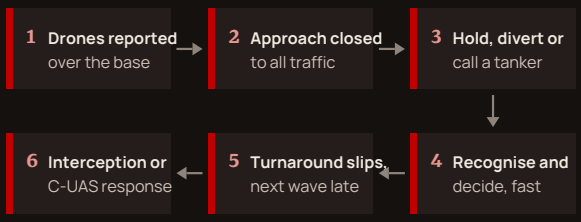
The conclusion is narrower than the technology headlines suggest, and more actionable. A force does not need learning agents to obtain most of the available value. It needs curated reactive adversaries with traceable provenance, stochastic variation, version control, and a body competent to adjudicate them, for the drone as much as for the fighter. That is achievable within a few years.

STATEMENT — THE HOME AIRFIELD

A drone does not have to reach the fight **to shape it.**

A few of them over a recovery airfield can close the approach to fighters and airliners alike. Aircraft hold on falling fuel, divert, or call for a tanker that was planned for something else. Controllers, base defenders and commanders must recognise what they are seeing and decide what to do, under time pressure and with civil traffic in the picture. The turnaround slips, the next mission launches late or not at all, and an interception may have to be mounted against an object far cheaper than anything used to stop it.

The cost is real even when not a single sortie is lost. At almost no expense to the side that launched them, drones impose a heavy one on the side that must respond.



LOW COST TO LAUNCH — HIGH COST TO ABSORB





PLATE 2 · THE OPPONENT AS AN ARTEFACT

Once the adversary is software, it becomes a versioned, auditable object with a provenance and a release history, whether it flies as a fighter or as a stream of drones. That is an unfamiliar idea in air combat training, and it is the point at which the discipline acquires something it has never had.



SECTION SEVEN

The assessment problem

Air forces have spent thirty years improving how training is delivered and almost nothing improving how it is judged. The debrief remains the most valuable hour in the syllabus and the least instrumented. This is the part of the problem where the synthetic environment offers something the aircraft cannot, and it is the part most programmes treat as a reporting requirement.

ASK WHAT A SQUADRON KNOWS ABOUT ITS OWN PROFICIENCY and the honest answer is a set of counts. Sorties flown, events completed, qualifications current, hours by category. Each of those is a measure of activity. None of them is a measure of competence, and the gap between the two is where readiness assessment currently lives.

The substantive judgement happens in the debrief, and it is made by a human being from memory, from a tape, and from a reconstruction of what probably occurred. That judgement is frequently excellent. It is also unrecorded in any form that survives the room, incomparable between instructors, and invisible to anyone trying to understand whether a squadron is getting better or worse.

WHAT THE DEBRIEF ACTUALLY LACKS

Three things, and only one of them is a technology problem.

A record of what was perceived. The debrief can establish what a crew did. It cannot establish what the crew saw, when they saw it, or what they were looking at when they missed something. That distinction is the difference between a decision error and a perception error, and the two require entirely different remediation. Instructors infer it constantly and have never been able to verify it.

> The debrief can establish what a crew did. It cannot establish what they saw. That single gap separates a decision error from a perception error, and the two need opposite remediies.
>
> SECTION 7 · WHAT THE DEBRIEF LACKS

A stable reference. Performance is judged against what the adversary did, and in the live environment the adversary does something different every time. Two crews cannot be compared, and the same crew cannot be compared with itself three months later, because the problem set has changed underneath the measurement.

A measurable performance assessment framework. Most air forces lack the capability to measure and assess tactical performance objectively through automated systems. The challenge is to translate established doctrine and tactics, techniques and procedures (TTPs) into machine-measurable criteria, so that assessment becomes

consistent, repeatable and evidence-based. This is an operational requirement enabled by technology, not defined by it.

WHAT THE SYNTHETIC ENVIRONMENT SUPPLIES BY CONSTRUCTION

Everything inside a simulation is, in principle, observable. The position and state of every entity at every timestep, every switch actuation, every sensor mode change, every radio transmission, every weapon employment and its outcome. None of this requires additional instrumentation because none of it exists outside the instrumentation.

Added to that, as discussed in Section 5, is gaze. Where a crew looked, for how long, in what order, and what was in the field of view when a contact was not called. An instructor reviewing an event can now answer the question that has been unanswerable for the history of the discipline: was that a failure of judgement or a failure of scan?

THE DISTINCTION THAT MAKES THE DATA USABLE

Raw telemetry is not assessment. A recording of every entity position at fifty hertz tells an instructor nothing that the instructor did not already suspect, and drowns the useful signal in volume.

What converts telemetry into assessment is an agreed set of tactically meaningful events, with machine-measurable criteria for judging their execution, applied the same way in every debrief. Without that layer, more data makes the problem worse rather than better.

The consequence is that assessment stops being an opinion formed under time pressure and becomes an argument supported by evidence. This is not a proposal to replace the instructor. It is a proposal to give the instructor the same evidentiary basis that every other high-consequence profession takes for granted.

THE MEASURABLE PERFORMANCE ASSESSMENT PROBLEM

This deserves greater attention because the critical limitation in modern air combat training is not the lack of tactical terminology, but the inability to objectively measure, assess and compare tactical performance through automated systems.

Consider a seemingly simple tactical action: a formation commits on a contact. The challenge is not defining a commit (experienced fighter instructors understand its tactical meaning) but determining whether it was executed correctly, at the right time, under the prevailing tactical conditions and in accordance with established TTPs.

Traditional instructor assessments rely heavily on experience and professional judgement. While essential, they cannot consistently quantify every tactical decision or aggregate performance across hundreds of engagements.

The real challenge is translating established fighter doctrine and TTPs into machine-measurable performance indicators, enabling automated, repeatable and objective assessment of tactical execution, decision-making and mission effectiveness.

This is fundamentally an operational and instructional requirement, enabled by technology, not a software problem to be defined by developers.

FROM TELEMETRY TO JUDGEMENT, FOUR LAYERS

Raw state — Free

Entity positions, switch actuations, sensor modes, transmissions, gaze. Generated automatically, useful to nobody in this form.

Events — Codable

Tactically meaningful occurrences, defined once and detected consistently. Doctrinal coding and software work.

Measures — The bottleneck

Timing, discipline, mutual support, employment within parameters, expressed relative to what the adversary did.

Judgement — Human

The instructor's assessment, now supported by evidence rather than recollection. Still human, and should remain so.

COUNTER-UAS PERSPECTIVE

A DOCTRINAL SPINE: THE F2T2E2A CYCLE

The assessment framework does not have to be invented from nothing. Alliance doctrine already offers a way of dividing an engagement into named steps. AJP-3.9, the Allied joint doctrine for joint targeting, offers component commands three engagement processes. The first is a cycle of seven steps (find, fix, track, target, engage, exploit and assess, abbreviated F2T2E2A), which they may adopt or adapt as their targeting method and which is also the process used for dynamic targeting.[49] The Alliance's doctrine for air and space operations, reissued in February 2026, describes the advantage of air power in the same terms, as finding, fixing, tracking, targeting, engaging and assessing adversary forces, and states that active air and missile defence is achieved through integrated detection, identification, assessment, interception

and engagement.[50] For a training system the value of the cycle is not that crews should recite it. It is that every step closes on something observable, and a step that closes on something observable can be trained, measured and compared. Figure 7.2 reads each step from a counter-UAS console.



FIGURE 7.2 The F2T2E2A cycle of AJP-3.9, read from a counter-UAS console. The summary of each step follows AJP-3.9, Figure 5.1 and Section 5.3. Each step closes on an observable event, and each event yields a measure of performance.

The cycle was written for targeting, and a crew defending an airfield against drones works at the far end of what the same publication calls the engagement continuum, where planning time runs out and joint targeting no longer applies. AJP-3.9 states that its procedures do not limit the right of self-defence. It defines a combat engagement, one too immediate for deliberate or dynamic targeting, as action outside self-defence under NAC-approved rules of engagement and the law of armed conflict, says plainly that it is not a joint targeting activity, and leaves the on-scene commander a basic, unwritten estimate of collateral damage.[49] Air and missile defence engagements are conducted as defensive counter-air, under weapons control procedures and measures set in advance by the air and missile defence commander.[50] At the console, then, the targeting process does not run. The functions its steps name still have to be performed, in seconds or minutes rather than hours, against identification criteria and rules set beforehand, and often by a single operator.

What changes is the time, not the substance. Identification, deconfliction with friendly and civil aircraft, the judgement of proportionality and the choice of effector all still happen; they happen faster, on less information and without the chance to consult. Over a home base in peacetime the law of armed conflict does not frame the decision at all: national law and policy do, within the State's international obligations, which makes the judgement harder, not easier. This is the strongest argument for rehearsal anywhere in the cycle. Judgement that a targeting staff would exercise with time in hand has to become a crew skill exercised without it, and for the reasons set out in Section 3 the only place it can be practised at operational density is the synthetic environment.

MEASURES, IN THE TERMS OF AJP-3.9

Measure of performance — **Crew**

Doing things right. One or more for each step, from detection to acknowledgement to the cost of the effector chosen. Used in the crew's debrief.

Indicator — **Link**

Tactical evidence that feeds a judgement of effect made higher up: contacts defeated before a set line, leakers, re-engagements, friendly or civil aircraft wrongly engaged.

Measure of effectiveness — **Force**

Doing the right things. Tied to the defended asset: aircraft launched and recovered on time, the disruption described in Section 6 denied, the cost per contact defeated.

The steps do not transfer equally. The first four are where the synthetic environment is strongest and live training weakest: the density of contacts, the mix of drones, decoys, birds and airliners, and the pressure of a decision taken over a working airfield can be presented at will and scored against known ground truth. Engagement transfers in part. The procedure, the sequence and the timing train well, but the physical effect of a missile, a gun or a jammer is a model, and live firing remains necessary for weapon handling and to keep the model honest. Assessment transfers well, because in a simulation the true outcome of every engagement is known and the crew's judgement of it can be checked, which a range can rarely offer.

AJP-3.9 also supplies the distinction the measurement layer needs. A measure of performance assesses friendly actions and is tied to task accomplishment; a measure of effectiveness assesses changes in behaviour, capability or the operating environment and is tied to an objective or an effect.[51] In the doctrine's own words, the first asks whether a force is doing things right, the second whether it is doing the right things. The doctrine adds that whether a measure of effectiveness has been met is hard to

judge at the tactical level and is better treated as an operational function, informed by tactical indicators. That is the right discipline for a training system too: the crew is debriefed on performance, the force is judged on effect, and the two are kept apart.

The same spine serves the fighter crew. An engagement against Red Air is also found, fixed, tracked, targeted, engaged, exploited and assessed, and the commit discussed above is one of the events that marks the passage from one step to the next. An assessment framework built on the cycle therefore lets a federated event (Section 8) assess the fighter crew, the air defence crew and the counter-UAS crew in one grammar and against one record. It also answers the question of who defines the requirement: definitions anchored in a doctrine whose use the nations have agreed, as recorded in STANAG 2524, belong to the Alliance and its members, not to whoever wrote the software. Defense News reported in October 2023 that NATO's first counter-drone doctrine, then about to go for ratification, would recommend common training standards for operators.[52] The F2T2E2A cycle is where such standards would become measurable.

WHAT BECOMES POSSIBLE ONCE THE LAYERS EXIST

Three things that no air force can currently do.

Longitudinal tracking. A crew's performance against a fixed adversary model, measured on the same events, three months apart. This is the first genuine measure of whether training is working, as distinct from whether training is happening.

Diagnosis rather than description. Aggregating events across a squadron reveals systematic weaknesses that no individual debrief would surface, because each instructor sees only their own events. A pattern of late commits across eleven crews is a training system problem. Seen one crew at a time, it looks like eleven individual lapses.

Evidence for resource decisions. A force that can demonstrate a measurable relationship between a category of training and a category of performance is in a fundamentally different position when defending its programme. Every air force currently argues for flying hours on the grounds that flying hours are necessary, which is true and unevidenced. The argument would be considerably stronger with numbers attached.

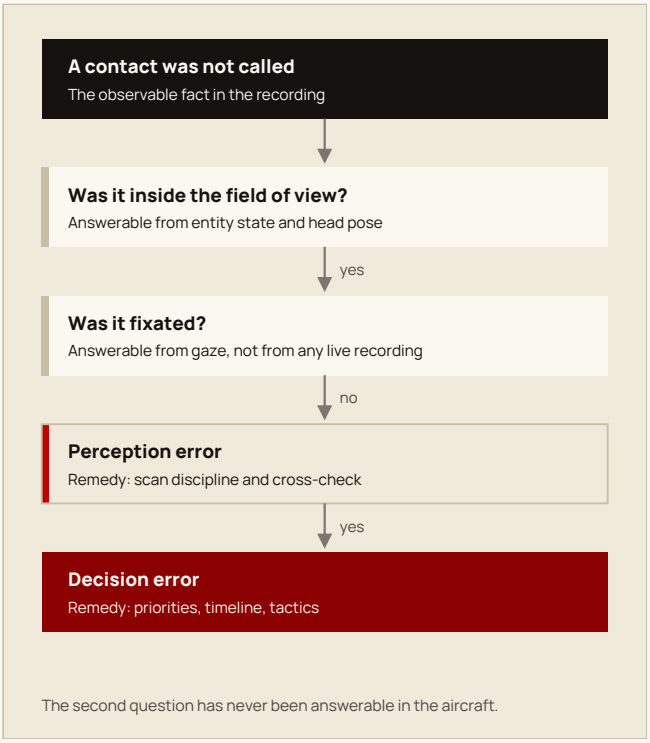


FIGURE 7.1 The question the debrief has never been able to answer. Gaze evidence separates a failure of scan from a failure of judgement, and the two need opposite remedies.

SECTION EIGHT

Architectures for a national and allied capability

The architectural choices made in the first eighteen months of a synthetic training programme determine what it can do in year ten. Three of those choices are effectively irreversible, and all three are routinely made by default rather than by decision.

MOST NATIONAL SYNTHETIC TRAINING CAPABILITY BEGINS THE SAME WAY. A platform is procured with simulators attached. Those simulators are excellent at what they were bought for, which is training individual crews on that platform, and they are incapable of talking to anything else because nobody asked them to. Repeat across three or four fleets over a decade and a force arrives at a collection of expensive isolated devices with no path to connecting them that does not involve paying each original supplier to open their own product.

This is not a failure of foresight by the officers involved. It is the predictable outcome of buying simulators as accessories to aircraft rather than as elements of a training system. Avoiding it requires three decisions to be taken deliberately and early.

DECISION ONE: WHERE THE CLASSIFICATION BOUNDARY SITS

This is the most consequential and the most frequently deferred. Every synthetic training capability contains data at several sensitivities: entity positions and generic events at one end, threat parameters and national tactics at the other. The architectural question is whether the boundary between them is designed in from the start or discovered later.

An architecture that treats the whole environment as classified is simple to reason about and expensive to live with. Every simulator, every network segment, every developer and every visiting ally is inside the accreditation boundary. Federation with partners becomes a negotiation each time, development slows to the pace of the security process, and commercial tooling is difficult to use.

An architecture that separates the layers permits something much more useful. The exercise environment, the entity exchange and the basic event stream operate at a low classification, while threat behaviour models, performance parameters and national tactical data sit in a controlled layer injected at run time. The formulation that matters is not that data never leaves, which is rarely true and rarely necessary. It is that classified data never leaves, while the unclassified machinery of the exercise is free to federate.

> The right formulation is not that data never leaves. It is that classified data never leaves, while the unclassified machinery of the exercise federates freely.
>
> SECTION 8 · THE CLASSIFICATION BOUNDARY

Getting this right has a second benefit that is not obvious until the first deployment. Hardware exists that is built for this problem: current professional headsets are offered in variants that can operate entirely offline, configured without radio components, for use in accredited facilities.[53] A force that has designed its data layers properly can use such equipment inside its secure environment and ordinary equipment outside it, running the same scenarios.

DECISION TWO: FEDERATION OVER INTEGRATION

Integration means making systems work together by modifying them. Federation means making them work together by agreeing an interface. The first produces a capability faster and a dead end sooner.

The practical test is what happens when the fifth platform joins. In an integrated architecture, each new participant requires bilateral work with every existing one, and the effort grows with the square of the number of participants. In a federated architecture with a published object model, a new participant implements the interface once. The published examples examined in Section 4 are explicit on this point: the systems that have scaled are described by their operators as standards based and commercially sourced rather than bespoke.[54]

Federation also determines whether a national capability can join an allied one. A force whose simulators speak a documented standard can participate in a coalition exercise by agreeing a federation agreement. A force whose

simulators speak a proprietary protocol can participate by buying a gateway from the vendor who sold them the problem.

THREE DECISIONS AND THEIR TEN-YEAR CONSEQUENCE

Classification boundary — Hardest to undo

Designed in layers from the outset, or discovered later. Determines whether federation with allies is a configuration or a negotiation.

Federation over integration — Compounding

Published object model versus bilateral modification. Determines whether the fifth platform costs the same as the second.

Ownership of models and data — Sovereignty

Whether the threat models, event definitions and training records belong to the force or to its suppliers.

DECISION THREE: WHO OWNS THE MODELS AND THE DATA

A synthetic training capability accumulates three assets over its life that are worth more than the equipment: the threat behaviour models, the event definitions that make assessment possible, and the accumulated training record of the force.

If those assets sit inside a supplier's product in a proprietary format, the force does not own its own training history and cannot change supplier without losing it. This is not a theoretical risk. It is the ordinary outcome of a contract that specifies a capability and is silent on data.

The mitigation is unglamorous and effective. Specify the export format for training data before award. Specify that event definitions are documented and owned by the force. Specify that threat models are delivered as data with a documented schema rather than compiled into an application. None of this prevents a supplier from adding value, and all of it preserves the option to change one's mind in year six.

THE HUB, THE SPOKES AND THE DEPLOYABLE NODE

A workable national topology has three elements, and their relative priority is often inverted in planning.

The **hub** is the persistent exercise environment: scenario management, entity generation, data capture and the after-action infrastructure. It is operated continuously rather than stood up for exercises, because a capability that is switched on twice a year accumulates nothing and is re-learned each time by whoever is currently posted to it.

The **spokes** are the training devices at operating bases, from full mission simulators through to lightweight mixed reality stations. The important design point is that the spokes should vary in fidelity deliberately, because a force needs a small number of high fidelity devices and a large number of adequate ones, and buying only the former limits repetition to those who can travel.

The **deployable node** is a containerised or otherwise transportable instance, capable of operating disconnected or, when needed, connected to the hub. Its justification is not expeditionary romance. It is that the alternative to training at a dispersed location is not training, and that a capability which only works at two main operating bases has assumed away the conditions under which it would matter most.

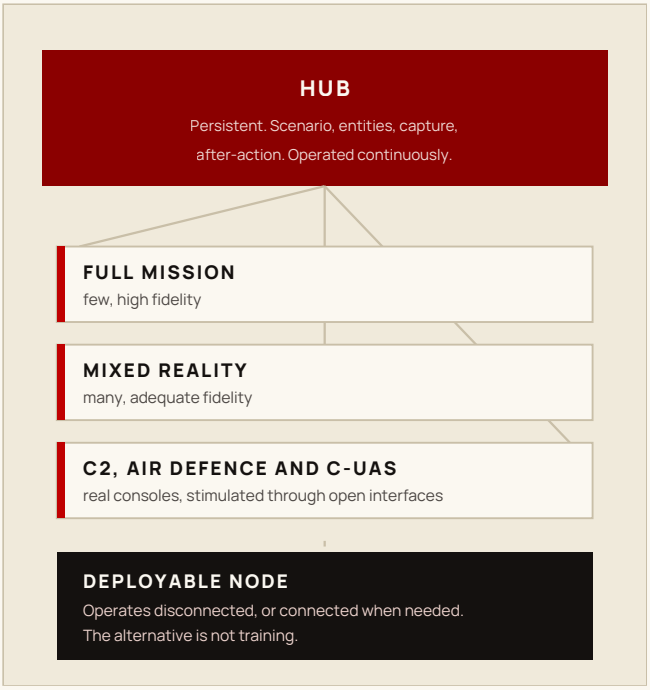


FIGURE 8.1 A national topology: a persistent hub, spokes of deliberately varied fidelity, and a deployable node that can operate disconnected or connect to the hub when needed. The same hub can serve fighter, air defence and counter-UAS crews.

COUNTER-UAS PERSPECTIVE

THE SAME HUB, A SECOND CREW

Nothing in this architecture is specific to fighters, and the crews who defend against drones are the clearest case for building it once. NATO's integrated air and missile defence policy, updated in 2024, states that Class I uncrewed systems must form an integral part of the Alliance's approach to air and missile defence, and sets the Alliance the aim of aligning NATO and national air and missile defence training and exercises.[55] A hub that already generates threat mass for fighter crews can present the same drones to ground-based air defence and counter-UAS crews, and the same exercise can rehearse the hand-off between them.

Three adjustments make that work. The crew should train on its own command and control system, stimulated with simulated sensor data through the open interface used in NATO's counter-UAS interoperability exercises, which NATO announced in 2023 it would adopt as a counter-UAS standard.[56] The classification boundary moves, because in a counter-UAS federation the most sensitive data is often not the threat but the defender: how far a nation's sensors really see and how well its effectors really work. That belongs in the national layer. And the civil world, air traffic control, the airport operator and the police, should be represented at each national node rather than connected to it, because those systems are civil, national and safety critical.

THE ALLIED DIMENSION

Nothing above requires an alliance to be useful, and all of it is improved by one. The reason is arithmetic again: the scarce input in adversary generation is not computing but doctrinal expertise, and no single mid-sized air force holds enough of it to curate a comprehensive threat library alone.

A federation of national capabilities sharing a common assessment framework and a jointly adjudicated threat model library would produce something none of the participants could build alone, at a cost each could bear. The precedent exists in the way alliances already maintain shared doctrine and standards. What does not exist is the equivalent arrangement for the models that define the simulated enemy, and Section 9 argues that building it is now the binding constraint.

SECTION NINE

Governance, assurance and the regulatory frame

A training system that generates adversaries with machine-learned behaviour, measures aircrew performance and informs readiness decisions is three regulated things at once. None of the three is currently governed by an accepted framework, and the absence is a larger obstacle to adoption than any remaining technical limitation.

THE GOVERNANCE CONVERSATION USUALLY BEGINS IN THE WRONG PLACE, with autonomous weapons. Policy on autonomy in weapon systems, most explicitly the United States directive on the subject, governs systems that select and engage targets, and requires appropriate levels of human judgement over the use of force.[57] A synthetic adversary selects and engages nothing. It exists inside a simulation and its weapons are numbers.

That distinction is real and it should be stated plainly, because it removes an objection that would otherwise stall the discussion. It does not, however, leave the field ungoverned. Three separate frameworks apply, and they apply for different reasons.

THE TRAINING SYSTEM AS A REGULATED ARTEFACT

Civil aviation has a mature answer to the question of whether a device may be credited for training, described in Section 5. A device is qualified at a level, against a specification, by an authority, following an evaluation with operational pilots, and the qualification states what training may be credited against it.[58]

Military synthetic training has nothing equivalent. There is no accepted means by which an air force can state that a given device, with a given adversary model, is qualified to deliver a given training objective. The result is that every adoption decision rests on the personal credibility of the officer proposing it and is reversible by their successor, which is precisely why capability in this field advances and retreats with postings.

Building such a framework is unglamorous work of exactly the kind that alliance bodies exist to do. It requires an agreed taxonomy of training objectives, an agreed method of demonstrating that a device meets one, and a body willing to make and stand behind the determination. The civil precedent shows that the standard does not have to be lowered to accommodate new technology. What changes is the compliance method, not the required outcome.

Every adoption decision currently rests on the personal credibility of the officer proposing it, and is reversible by their successor. That is why capability in this field advances and retreats with postings.

SECTION 9 · THE ASSURANCE GAP

THE ADVERSARY MODEL AS A CONTROLLED ASSET

Section 6 set out why a behaviour model needs provenance, adjudication, version control and tested envelopes. The governance point is who performs those functions and on what authority.

Three properties make a body suitable. It must hold current doctrinal and threat expertise, which in practice means access to national assessments from more than one nation. It must be trusted to arbitrate between conflicting national views without a national position of its own. And it must have no commercial interest in the outcome, because a supplier certifying the accuracy of its own threat models is an arrangement whose weakness needs no elaboration.

Alliance Centres of Excellence satisfy all three by construction. They are multinationally staffed, they exist to develop and validate concepts rather than to sell them, and they already perform an adjudicating role in adjacent domains. The specific function proposed here is narrow: to maintain a reference library of adversary behaviour models with documented provenance, to convene the expert panel that reviews them, and to publish version-controlled releases that member nations may adopt, extend nationally or decline. Section 10 considers the Centre best placed to perform that role for air and missile defence.

WHAT A THREAT MODEL AUTHORITY WOULD AND WOULD NOT DO

It would maintain provenance, convene expert adjudication, hold configuration control, and publish tested releases with stated behavioural envelopes.

It would not hold national classified tactics, mandate adoption, certify commercial products, or replace national threat assessment. The value is in the shared baseline, not in central control.

THE ARTIFICIAL INTELLIGENCE FRAME

European regulation on artificial intelligence takes a risk-based approach, applies obligations according to the use to which a system is put, and excludes systems used exclusively for military, defence or national security purposes from its scope.[59] The framework is nonetheless a useful reference, and two of its features bear directly on synthetic training.

The first is that systems used in education and vocational training to evaluate learning outcomes, including where those outcomes steer a learner's progress, are classed as high risk. They carry obligations on transparency and human oversight, and a person significantly affected by a decision based on one has a right to an explanation. An assessment system that contributes to a judgement about whether a pilot progresses is in that territory in substance, whether or not the regulation formally binds a defence ministry.

The second is the requirement for meaningful human oversight, which aligns with Section 7, where judgement stays with the instructor. Measures inform the instructor and do not replace the instructor. That is good pedagogy independently of regulation, and it happens also to be the position that will survive scrutiny.

The prudent posture for a European force is therefore to design to the spirit of the framework from the beginning rather than to argue about applicability. The cost of doing so is modest. The cost of retrofitting explainability into a system that was not built for it is not.

EXPLAINABILITY, AND WHY IT CONSTRAINS THE TECHNOLOGY CHOICE

This is where the governance argument feeds back into the engineering one, and it is the reason Section 6 concluded in favour of curated reactive adversaries rather than learning agents for near-term adoption.

An instructor debriefing a crew will be asked why the adversary did what it did. With a rule-based model, the answer exists: a behaviour was selected because stated conditions were met, and the rule traces to a documented source. With a learned policy, the answer is that the network produced that output, which is true, unhelpful, and corrosive of the crew's confidence in the whole environment.

This is not an argument that learning agents have no place. It is an argument about sequence. A force that has established provenance, adjudication, version control and behavioural envelopes for its rule-based models has built exactly the apparatus needed to assure a learned one later. A force that adopts learned adversaries first has no apparatus at all, and will find that the first serious disagreement in a debrief is unanswerable.

SECTION TEN

From threat mass to evidence: a role for the IAMD COE

The functions this paper has described, a curated threat library, an agreed assessment framework, a federated environment and the evidence that training works, need an institutional home that no nation and no supplier can provide alone. For integrated air and missile defence, and above all for counter-UAS, a candidate already exists.

⌘ COUNTER-UAS PERSPECTIVE

SECTIONS 6, 8 AND 9 ARRIVED AT THE SAME REQUIREMENT by different routes. Each needs a body that holds doctrinal and threat expertise from several nations, can arbitrate between national views without holding one of its own, and has no commercial interest in the answer. Alliance Centres of Excellence meet those conditions by construction. For the part of the problem that is growing fastest, the integration of the uncrewed threat into air and missile defence, the IAMD COE is the natural candidate, and its recent work, and proposals published in its conference journal, point the same way.

A POLICY BASIS THAT ALREADY EXISTS

NATO's integrated air and missile defence policy treats the threat as a single spectrum, running from small uncrewed aircraft to cruise, ballistic and hypersonic missiles, and sets the Alliance the aim of aligning NATO and national air and missile defence training and exercises.[55] Counter-UAS is one of the IAMD COE's focus areas.[60] The Centre has published two studies of the uncrewed threat: one drew lessons from the war in Ukraine and described the unfavourable cost-interception curve that cheap, attritable drones create,[61] and the other examined their asymmetric use.[62] It hosted, in Chania in June 2024, a meeting of the NATO research task group working on mission-level modelling and simulation for counter-UAS, whose intended uses include live, virtual and constructive training,[63] and it took part in the Alliance's layered counter-UAS experiment in Romania in April 2026.[64]

The journal of the Centre's 2025 annual conference, published in 2026, goes further. In it, one of the present authors, writing as a former Director of the IAMD COE, proposed an Alliance synthetic training network across services and domains, supported by a modelling and simulation facility at the IAMD COE with remote connectivity over the Combined Federated Battle Laboratories Network, and treating simulation data as a strategic asset.[65] Another article in the same issue argues for using uncrewed aircraft as aggressors in air defence training and identifies airspace integration as the barrier.[23] Read alongside this paper, the two proposals describe a single capability: live aggressor drones where only live flight will do, and a persistent synthetic laboratory for the density, the mix and the measurement that airspace will never permit.

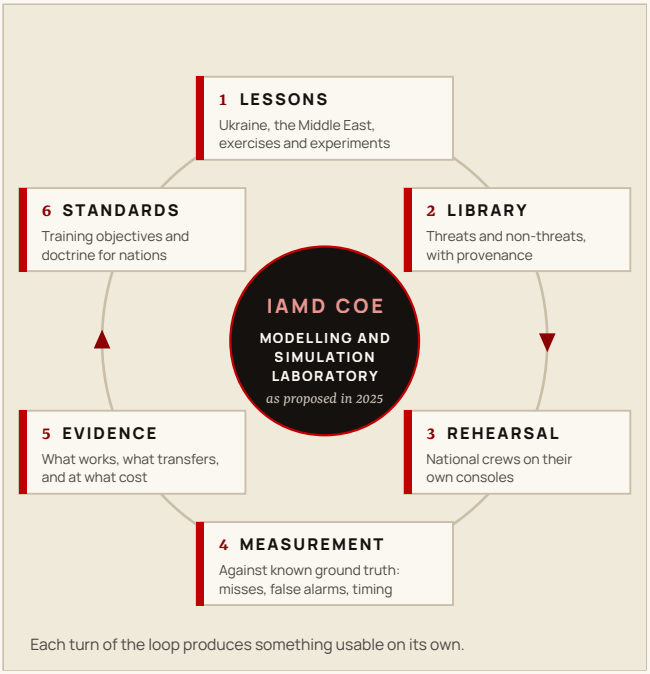


FIGURE 10.1 The laboratory loop. Lessons enter a curated library, crews rehearse against it on their own consoles, the environment measures them against known ground truth, and the evidence returns as training objectives and doctrine. Nations keep their national layer throughout.

WHAT A MODELLING AND SIMULATION LABORATORY WOULD ADD

A laboratory of that kind, connected to national nodes, could take counter-UAS research and development further than any single national programme or annual exercise, because it would bring together five things that are rarely found in one place, each of them an extension of something this paper has already argued for.

Threat mass on demand is Section 6 applied to a new opponent. Known ground truth is the assessment layer of

Section 7, which a synthetic environment supplies by construction. The recorded sky and the stimulated console belong to the hub and the national layer of Section 8. The curated library is the threat model authority of Section 9, extended to the birds, civil aircraft and friendly drones that a counter-UAS crew must learn not to engage. Lessons from Ukraine already reach the Alliance's experimentation through the Joint Analysis, Training and Education Centre,[66] and would be among the library's most valuable inputs.

FIVE CAPABILITIES, ONE LABORATORY

Threat mass on demand — Scale

Hundreds of drones and decoys in one picture, alongside cruise missiles and aircraft, at the cadence of the war that defines the problem rather than of a range booking.

Known ground truth — Evidence

Every contact presented is known, so crews and systems can be measured against what was actually there: misses, false alarms and the time each step takes.

The real sky, a synthetic threat — Honesty

Sensor recordings from real sites, with their birds, civil traffic and clutter, replayed with synthetic threats injected into them.

Real consoles, open interfaces — Transfer

National command and control systems stimulated through the open sensor interface already used in the Alliance's interoperability exercises.

A short loop from lessons to training — Speed

New threat behaviour seen in operations entered into a curated library, reviewed and released to nations in weeks rather than procurement cycles.

THE LESSONS THAT MAKE IT URGENT

The case is not theoretical. Over Ukraine the defence has moved its weight from the munition to the operator: interceptor drones costing from under 1,500 dollars to several thousand accounted for around forty per cent of Shahed-type kills by April 2026, flown by roughly a thousand crews.[67] Where the Ukrainian Air Force counted its own interceptor crews, sixty-six of more than three hundred had destroyed more than ten targets each, while one hundred and seventy had destroyed none.[68] Over Israel, more than a thousand drones were launched from Iran in June 2025 and 99 per cent were intercepted.[17] In a single night in April 2024, Israeli and American fighters shot down a large share of some 170 drones,[14] British Typhoons a number of others,[15] and one former Israeli military financial adviser put the cost of Israel's defence that night, against missiles and drones together, at one to 1.3 billion dollars.[18] Over the Red Sea, the United States Navy fired around 220 missiles in some fifteen months of operations.[69]

Three lessons recur across these theatres. The crew matters at least as much as the equipment. The cost of the response, not only its success rate, decides whether a defence can be sustained. And the threat reaches the air base as readily as the front line. Each is a research question as much as a training one, and none can be studied live at the scale the evidence demands.

QUESTIONS THE ALLIANCE CANNOT YET ANSWER WITH EVIDENCE

How long the human steps between first detection and engagement take for real crews, and how training changes them.

What false alarm rate a training picture must carry before it stops teaching crews to trust every track.

Whether synthetic counter-UAS training transfers to live performance, and by how much.

How fighter, ground-based air defence and counter-UAS crews should share one picture and one decision when the threat is mixed.

What a drone over the home airfield costs an air operation, and how the response is best rehearsed.

None of these questions can be answered by procurement, and none by a single nation. Each could be answered by a focused, multinational research effort built on the kind of synthetic environment this paper describes, hosted where the relevant expertise already sits, and attached to exercises and experiments that already gather the equipment. Funding routes have begun to open: the Alliance's Science for Peace and Security programme ran a special call on counter-UAS from April to June 2026 that made advanced training courses eligible.[70] None of this would require a new institution, and it could be taken forward within existing arrangements if the Centre's sponsoring nations chose to. Every step would produce something usable on its own.

The return would be tangible: training objectives the Alliance can share, a library it can trust, and the first measured evidence of how counter-UAS crews perform and what improves them. The lessons have already been paid for, in Ukraine and over the Middle East. Research is how the Alliance turns them into readiness, and the IAMD COE is well placed to show the way.

SECTION ELEVEN

A maturity model and a sequenced roadmap

Capability arguments fail when they cannot be made specific. A maturity model exists to make the argument arguable: to say what a force should be able to do at each level, so that investment can be sequenced to produce usable capability early rather than complete capability late, and so that progress can be contested with evidence.

THE LEVELS BELOW DESCRIBE CAPABILITY, NOT EQUIPMENT. A force at level two with excellent hardware is still at level two. The distinguishing question at each level is what the force can demonstrate, not what it has purchased, and the transitions that matter are the ones that change what a squadron does on an ordinary Tuesday.

LEVEL 1: INSTRUMENTED INDIVIDUAL TRAINING

Devices exist at operating bases and are used for procedural and emergency training. Data is captured but not analysed beyond the individual session. Adversaries are scripted. The training value is real and confined to the individual crew.

The test: can a crew complete a meaningful proportion of procedural currency without flying? Most forces are here or approaching it.

LEVEL 2: VIRTUAL AND CONSTRUCTIVE ADVERSARIES WITH DOCTRINAL PROVENANCE

Scripted entities are replaced by reactive behaviour models whose tactics trace to documented sources, with stochastic variation, version control and expert review, and experienced aircrew fly virtual adversaries from simulators where a thinking opponent is needed. Training events now contain an opponent that can surprise a crew.

The test: can the force state which version of which threat model its crews trained against last month, and who approved it? This is the level at which synthetic training stops being a procedures trainer and starts being adversary training. It is achievable without federation and without new aircraft, and it is where artificial intelligence can first enter with a human in the loop, as the instructor's assistant, selecting, varying and adapting adversary behaviour within approved envelopes while the instructor decides.

LEVEL 3: FEDERATED COLLECTIVE TRAINING

Devices at separate locations participate in a single exercise over a documented standard interface, with correlated terrain, environment and weapon models. A national hub operates continuously. Large force employment can be rehearsed at a scale the live environment cannot supply, and the same hub can present the drone threat to air defence and counter-UAS crews.

The test: can four crews at three locations fly one event, and can they all defend the same account of what happened in the debrief?

LEVEL 4: INSTRUMENTED ASSESSMENT

An agreed assessment framework exists. Measures are derived consistently, compared across events and crews, and used to inform instructor judgement rather than to replace it. Longitudinal tracking becomes possible.

The test: can the force demonstrate, with evidence, that a squadron is better at something than it was six months ago? Almost no force can currently do this, and it is the level at which the investment starts defending itself in budget discussions.

LEVEL 5: LIVE-VIRTUAL-CONSTRUCTIVE INTEGRATION

Airborne aircraft, simulators and constructive entities occupy one battlespace. Aircrew in the air see synthetic opponents; aircrew in simulators see live participants; the assessment layer spans both.

The test: can a live four-ship and a virtual four-ship fly the same event against a common adversary package, and be debriefed together?

THE FIRST THIRTY-SIX MONTHS

Four things, in this order, for a mid-sized air force starting from level one.

Establish the adversary before the architecture. Commission the doctrinal work to define threat behaviour models with documented provenance. This is the long-lead item, it is cheap in capital and expensive in expert time, and every subsequent step depends on it. It can be started now with people the force already employs.

Agree the assessment framework in parallel. Define the tactically meaningful occurrences, and the machine-measurable criteria, that assessment will be built on. This is a doctrinal task of similar character and similar duration, and it has the useful property that it improves live debriefs immediately, before any system exists to exploit it. The F2T2E2A cycle of Alliance targeting doctrine already supplies a skeleton for it (Section 7).

SEQUENCING, AND THE COMMON ERROR AT EACH STEP

Levels 1 to 2 — 0 to 12 months

The error is buying fidelity instead of building adversaries. A better headset does not produce a better opponent.

Levels 2 to 3 — 12 to 30 months

The error is federating before the adversary is worth sharing. Connecting three sites to fight scripted entities multiplies the disappointment.

Levels 3 to 4 — 24 to 48 months

The error is buying analytics before agreeing the assessment framework. The doctrinal work is very difficult to procure.

Levels 4 to 5 — 36 to 72 months

The error is attempting live integration to impress rather than to train. LVC is a consequence of the first four levels, not a substitute for them.

Specify data ownership at the first procurement. Export formats, documented schemas, force-owned event definitions and threat models delivered as data. This costs nothing at the point of writing and is unpurchasable afterwards.

Buy a small number of connectable devices, not a large number of isolated ones. Fidelity at the point of need, standards at every point. Resist the procurement logic that treats simulators as accessories to aircraft, because that logic is what produced the isolated estate the force is trying to escape.

> Establish the adversary before the architecture. It is the long-lead item, it is cheap in capital and expensive in expert time, and everything else depends on it.
>
> SECTION 11 · THE FIRST RECOMMENDATION

WHAT THIS COSTS, AND WHAT IT IS COMPARED AGAINST

This paper has deliberately avoided figures, because national costs vary by an order of magnitude with existing infrastructure and because a number quoted out of context does more harm than good. The comparison that matters is not the absolute cost of the capability. It is the cost of the alternative.

The alternative is a continuing decline in adversary quality, unmeasured because nobody measures adversary quality, inside a flying programme that every published indicator suggests will contract further. A force choosing not to build synthetic adversary capability is not choosing to save money. It is choosing to accept a reduction in training quality whose size it has no means of estimating.

CONCLUSION

The opponent is the capability

Air combat training has always been constrained by the availability of a credible enemy. For most of its history that constraint was unavoidable, because the only way to produce an opponent was to fly one.

THE ARGUMENT OF THIS PAPER IS NARROW. It is not that simulation should replace flying, which nobody serious proposes and no evidence supports. It is that the adversary function has become separable from the airframe that used to provide it, and that a force which recognises this early acquires an advantage that compounds.

Four claims support that argument, and each is answerable with evidence rather than assertion.

The first is that the live environment cannot supply the adversary at the rate the training requires, and the trend is going the wrong way. Sortie rates, manning and availability are all documented in the open record, and none of the three has a remedy available inside a decade.

The second is that the fidelity objection to synthetic training, which was well founded for thirty years, has been answered parameter by parameter. The thresholds are now published numbers rather than matters of opinion, and a force that declines synthetic adversary training on fidelity grounds should be asked which parameter it means.

The third is that the hard problem was never the simulator. It is the opponent inside it, and the opponent is a doctrinal artefact rather than a software product. That is inconvenient for anyone hoping to solve the problem by procurement alone, and encouraging for any air force with experienced aircrew and access to threat expertise, because the scarce input is one it already holds. It also means the work needs people who have flown and taught the tactics, working alongside the engineers who encode them.

> The hard problem was never the simulator. It is the opponent inside it, and the opponent is a doctrinal artefact rather than a software product.
>
> CONCLUSION

The fourth is that the drone has made the argument urgent. It brings threat mass in numbers no aggressor fleet can present, into airspace where hostile profiles can be flown only in tightly controlled trials, and it reaches the air base as well as the fight. For the crews who must defend against it, the synthetic environment is more than a supplement to live training. For much of what matters, it is the only place the training can happen.

Two consequences follow for anyone deciding what to do next.

For a national force, whichever service owns the mission (for ground-based air and missile defence, often the army, as in the United States and the United Kingdom), the sequence matters more than the ambition. Curating adversary behaviour models and agreeing an assessment framework are cheap in capital, expensive in expert time, and prerequisite to everything else. They can begin without a procurement and they improve live training immediately. The architectural decisions taken alongside them, particularly where the classification boundary sits and who owns the models and the data, will determine what remains possible in year ten.

For an alliance, the gap is institutional rather than technical. Nobody currently adjudicates whether a simulated enemy is a fair representation of a real one, and nobody qualifies a training device against a training objective. Adjudicating threat models needs a multinational body with doctrinal expertise and no commercial interest, and for air and missile defence the IAMD COE is such a body; it is also well placed to host the research that would show whether synthetic counter-UAS training works. Qualifying devices against training objectives would need an Alliance or national authority, which such a Centre could advise. Industry cannot give itself that authority, but it can make both functions practical by delivering threat models as open, versioned data and instrumenting devices for qualification; the partners best placed to do so combine doctrinal expertise with engineering.

The question in the title has a short answer. When a fighter force can no longer spare the aircraft to fly against itself, the opponent comes from the doctrinal knowledge the force already holds, formalised into something a machine can execute, adjudicated by people qualified to judge it, and flown against every week instead of twice a year. For the drone over the home airfield, there is no other way to fly it at all. The technology to do this is available. The institutional arrangements to make it trustworthy are not yet, and building them is the work that now matters.



PLATE 3 · WHAT THE ARITHMETIC LEAVES BEHIND

The shelter is built, maintained and manned. Whether anything inside it is ready depends on how often its crews have flown against an opponent worth the name, and whether the sky above it is still its own.

REFERENCES

Notes and references

All sources are open. Numbered notes correspond to the superscript markers in the text; entries are given in Harvard form and were accessed between 15 September and 14 October 2026 unless otherwise stated. Where a claim is contested or a figure is an estimate, the text says so at the point of use.

SECTION 1

1 Stassis, C. (2026) 'Air Force fighter pilots fall below sortie minimum to be deemed combat ready', Military Times, 11 September. Available at: https://www.militarytimes.com/news/your-military/2026/09/11/

2 Simple Flying (2026) 'Air Force budget documents project fighter aviator manning will fall to 67% in 2027'. Available at: https://simpleflying.com/air-force-fighter-aviator-manning-2027/

3 Simple Flying (2026), as note 2, citing United States Air Force budget documentation on mission capable rates.

4 The National Interest (2026) 'US Air Force pilots aren't flying enough, internal report says', 28 May. Available at: https://nationalinterest.org/blog/buzz/

5 Common Defense (2026) 'Air Force fighter pilots plunge below sortie minimums for combat readiness', 13 September. Available at: https://www.commondef.com/2026/09/

6 The National Interest (2026), as note 4, on comparative Russian tactical aviation training hours.

7 Janes (2022) 'UK Typhoon fleet to get new networked simulators in 2022'. Available at: https://www.janes.com/defence-intelligence-insights/defence-news/c4isr/

SECTION 3

8 Skies Magazine (n.d.) 'Top Aces makes a bold move bringing F-16s into contract Red Air'. Available at: https://skiesmag.com/news/top-aces-bold-move-f-16-contract-red-air/

9 Key.Aero (n.d.) 'USAF sees red with eagerly awaited aggressor training contract'. Available at: https://www.key.aero/article/usaf-sees-red-eagerly-awaited-aggressor-training-contract

10 European Security & Defence (2025) 'Countering the aerial threat', 16 June. Available at: https://euro-sd.com/2025/06/articles/44809/countering-the-aerial-threat/

11 Key.Aero (2022) 'Top Aces selected to provide adversary air training at two USAF bases', 26 September. Available at: https://www.key.aero/article/top-aces-selected-provide-adversary-air-training-two-usaf-bases

12 Aerospace America (AIAA) (n.d.) 'Red Air reset'. Available at: https://aerospaceamerica.aiaa.org/features/red-air-reset/

13 Institute for Science and International Security (2026) Monthly analysis of Russian Shahed-136 deployment against Ukraine, updated 7 July, compiling Ukrainian Air Force data. Available at: https://isis-online.org/isis-reports/monthly-analysis-of-russian-shahed-136-deployment-against-ukraine

14 Ott, H. (2024) 'What Iran launched at Israel in its unprecedented attack, and what made it through the air defenses', CBS News, updated 16 April. Available at: https://www.cbsnews.com/news/iran-israel-attack-what-weapons-launched-how-air-defenses-worked/

15 Osborne, T. (2024) 'RAF Eurofighters shot down Iranian drones, UK Prime Minister says', Aviation Week, 15 April. Available at: https://aviationweek.com/defense-space/budget-policy-operations/raf-eurofighters-shot-down-iranian-drones-uk-prime-minister

16 Stars and Stripes (2025) 'Poland, NATO respond to Russian drone incursion', 10 September. Available at: https://www.stripes.com/theaters/europe/2025-09-10/poland-nato-russian-drone-incursion-19041045.html

17 Alma Research and Education Center (2025) 'Summary of the Iran-Israel war on the Israeli front: Operation Rising Lion (June 2025)', 30 June. Available at: https://israel-alma.org/summary-of-iran-israel-war-on-the-israeli-front-operation-rising-lion-june-2025/

18 Middle East Eye (2024) 'Israel's defence against Iran attack overnight likely cost over $1bn', 14 April, reporting an estimate by Brigadier General (ret.) Reem Aminoach, a former financial adviser to the Israeli military, given to Ynet. Available at: https://www.middleeasteye.net/news/israels-defence-against-iran-attack-overnight-likely-cost-over-1bn

19 Hollenbeck, N., Altaf, M.H., Avila, F., Ramirez, J., Sharma, A. and Jensen, B. (2025) 'Calculating the cost-effectiveness of Russia's drone strikes', Center for Strategic and International Studies, 19 February. Available at: https://www.csis.org/analysis/calculating-cost-effectiveness-russias-drone-strikes; Jensen, B. and Atalan, Y. (2025) 'Drone saturation: Russia's Shahed campaign', Center for Strategic and International Studies, 13 May. Available at: https://www.csis.org/analysis/drone-saturation-russias-shahed-campaign

20 Trevithick, J. and Altman, H. (2025) 'Laser-guided rockets now primary anti-drone weapon for USAF jets in Middle East', The War Zone, 25 September. Available at: https://www.twz.com/air/laser-guided-rockets-now-primary-anti-drone-weapon-for-usaf-jets-in-middle-east

21 Paulsen, D. and Baker, J. (2025) 'Counter-Unmanned Aerial Systems (C-UAS) training and implementation at the National Training Center', Center for Army Lessons Learned, 2 September. Available at: https://www.army.mil/article/288207

22 European Commission (2019) Commission Implementing Regulation (EU) 2019/947 on the rules and procedures for the operation of unmanned aircraft, 24 May. Available at: https://eur-lex.europa.eu/eli/reg_impl/2019/947/oj

23 Caetano, J. (2026) 'Challenges of unmanned aircraft systems airspace integration and their potential for enhanced integrated air and missile defence', in IAMD COE Annual Conference Journal 2025. Souda: IAMD COE, published July 2026. Available at: https://iamd-coe.org/wp-content/uploads/2026/07/IAMD-magazine-2026_28-4_SITE_compressed.pdf

24 Red 6 Aerospace (2025) 'Red 6 secures US Air Force contract to revolutionize F-16 pilot training with world-first augmented reality training system', press release, 15 August. Available at: https://www.prnewswire.com/news-releases/

25 Leonardo and Red 6 (2026) 'Red 6 and Leonardo redefine live air combat training with M-346 ATARS integration', press release, 8 April. Available at: https://www.prnewswire.com/news-releases/

SECTION 4

26 Defence Equipment & Support (2022) 'World-class synthetic training system opened at RAF Waddington', 22 February. Available at: https://des.mod.uk/

27 Janes (2026) 'RAF expands use of Gladiator synthetic training environment', 5 March. Available at: https://www.janes.com/osint-insights/defence-news/c4isr/

28 Janes (2022), as note 7, on the Typhoon Future Synthetic Training programme and the synthetic to live training mix.

29 NATO Allied Air Command (2026) 'NATO Allied Air Command to conduct Exercise Ramstein Flag 2026 in June', 6 May. Available at: https://natoaircom.substack.com/

30 NATO Allied Command Transformation (2026) 'AI in Audacious Training: JWC advances a NATO Allied Command Transformation beacon project', 15 March. Available at: https://www.act.nato.int/article/ai-audacious-training/

31 NATO (2024) news release on the 2024 Counter-UAS Technical Interoperability Exercise, September. Available at: https://www.nato.int/cps/en/natohq/news_228911.htm

32 NATO Communications and Information Agency (2026) 'Allies and industry test the latest counter-drone technology during NATO exercise', May. Available at: https://www.ncia.nato.int/newsroom/news/allies-and-industry-test-the-latest-counterdrone-technology-during-nato-exercise

33 NATO Communications and Information Agency (2025) 'NCIA strengthens Allies' counter-drone defence training', March. Available at: https://www.ncia.nato.int/newsroom/news/ncia-strengthens-allies-counter-drone-defence-training

34 Joint Air Power Competence Centre (2026) Annual Report 2025. Kalkar: JAPCC. Available at: https://www.japcc.org/reports/annual-report-2025/

35 Vogt, K. (2025) article in The Journal of the JAPCC, edition 39, May. Kalkar: Joint Air Power Competence Centre. Available at: https://www.japcc.org/?p=18426

SECTION 5

36 Varjo Technologies (2021) Varjo XR-3 technical specification: focus area 70 PPD, peripheral area above 30 PPD, 115 degree horizontal field of view, automatic interpupillary adjustment 59 to 71 mm.

37 Varjo Technologies (2026) XR-4 Series product specification. Available at: https://varjo.com/products/xr-4

38 Varjo Technologies (2026) 'XR-4 or XR-4 Focal Edition? How to choose the right passthrough for your use case', 18 May. Available at: https://varjo.com/blog/

39 Varjo Technologies (2026), as note 37, on automatic interpupillary adjustment across 56 to 72 mm.

40 Varjo Technologies (2026), as note 37, on integrated gaze tracking at 200 Hz with one-degree accuracy.

41 European Union Aviation Safety Agency (2021) 'EASA approves the first Virtual Reality (VR) based Flight Simulation Training Device', press release, 26 April. Available at: https://www.easa.europa.eu/en/newsroom-and-events/press-releases/

42 Airbus (2022) 'VRM Switzerland and Airbus Helicopters work together to qualify under EASA regulations the world's first H125 virtual reality simulator', 19 May. Available at: https://www.airbus.com/en/newsroom/stories/

SECTION 6

43 Defense Advanced Research Projects Agency (n.d.) Air Combat Evolution (ACE). Available at: https://www.darpa.mil/about/innovation-timeline/ace

44 Aerotime (2024) 'USAF and DARPA conduct first AI vs human dogfights', 23 September. Available at: https://www.aerotime.aero/articles/usaf-darpa-x-62a-ai-dogfight

45 The Defense Watch (2026) 'US Air Force X-62 AI intercept test marks major step toward autonomous air combat', 5 August. Available at: https://thedefensewatch.com/aerospace-aviation/

46 Mitchell, E. (2024) 'Pentagon confirms incursions of unauthorized drones over Air Force base', The Hill, 16 October, reporting The Wall Street Journal. Available at: https://thehill.com/policy/defense/4937166-drone-flights-virginia-base/

47 Reuters (2025) 'Drones disrupt flights at second Danish airport in a week', 25 September, via Daily Maverick. Available at: https://www.dailymaverick.co.za/article/2025-09-25-drones-disrupt-flights-at-second-danish-airport-in-a-week; Agence France-Presse (2025) 'Nordic airports reopen', via Khaleej Times. Available at: https://www.khaleejtimes.com/world/europe/nordic-airports-reopen-drone-mystery

48 The Kyiv Independent (2025) 'Reuters: Ukraine struck fewer Russian bombers during Operation Spiderweb than estimated, US officials claim', 5 June. Available at: https://kyivindependent.com/reuters-ukraine-struck-fewer-russian-bombers-during-operation-spiderweb-than-estimated-us-officials-claim/

SECTION 7

49 NATO Standardization Office (2021) AJP-3.9, Allied Joint Doctrine for Joint Targeting, Edition B, Version 1, November; agreement of nations recorded in STANAG 2524. Preface (scope and application), paragraph 1.3.5 (combat engagement), Section 1.4 and Figure 1.1 with its note 2 (the engagement continuum), Sections 5.2 and 5.3 and Figure 5.1 (F2T2E2A). Available from the NATO Standardization Document Database: https://nso.nato.int/nso/

50 NATO Standardization Office (2026) AJP-3.3, Allied Joint Doctrine for Air and Space Operations, Edition C, Version 1, February, paragraphs 1.5, 2.19a and 3.5a. Available at: https://assets.publishing.service.gov.uk/media/698ef9fb7da91680ad7f4389/AJP_3_3_EdC_V1.pdf

51 AJP-3.9, as note 49: lexicon, 'measure of effectiveness' and 'measure of performance' (NATO Agreed); paragraph 1.5.1f (Phase 6, assessment); paragraph 5.4.4.

52 Gosselin-Malo, E. (2023) 'NATO to adopt first-ever counter-drone doctrine for member nations', Defense News, 20 October. Available at: https://www.defensenews.com/unmanned/2023/10/20/nato-to-adopt-first-ever-counter-drone-doctrine-for-member-nations/

SECTION 8

53 Varjo Technologies (2026) XR-4 Secure Edition: variants able to operate without online connectivity and configurable to exclude radio components for use in accredited facilities. Available at: https://varjo.com/products/xr-4-secure-edition

54 Janes (2026), as note 27, on the standards-based, commercially sourced design of the national environment described there.

55 NATO (2025) NATO Integrated Air and Missile Defence Policy, public summary, 13 February, paragraph 57. Available at: https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2025/02/13/nato-integrated-air-and-missile-defence-policy; NATO (2024) Washington Summit Declaration, 10 July, paragraph 8, on the updated policy.

56 UK Government (2024) SAPIENT autonomous sensor system, guidance updated 6 September, describing the interface published as BSI Flex 335. Available at: https://www.gov.uk/guidance/sapient-autonomous-sensor-system; Savage, O. (2023) report on NATO's adoption of SAPIENT as a counter-UAS standard, Janes, 25 September. Available at: https://www.janes.com/defence-news/c4isr-command-tech/latest/nato-to-adopt-sapient-as-c-uas-standard

SECTION 9

57 United States Department of Defense (2023) Directive 3000.09: Autonomy in Weapon Systems.

58 European Union Aviation Safety Agency (2021), as note 41, on qualification level, evaluation method and training credit.

59 European Union (2024) Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), Official Journal of the European Union, Article 2(3), Articles 13, 14 and 86, and Annex III, point 3. Available at: https://eur-lex.europa.eu/eli/reg/2024/1689/oj

SECTION 10

60 NATO Integrated Air and Missile Defence Centre of Excellence (n.d.) 'Countering Unmanned Aerial Systems (C-UAS)', focus area. Available at: https://iamd-coe.org/focus-areas/countering-unmanned-aerial-systems-c-uas/

61 NATO Integrated Air and Missile Defence Centre of Excellence (2024a) The Evolving UAS Threat: Lessons from the Russian-Ukrainian War since 2022 on Future Air Defence Challenges and Requirements. Souda: IAMD COE. Available at: https://iamd-coe.org/wp-content/uploads/2024/02/The-Evolving-UAS-Threat-Lessons-from-the-Russian-Ukrainian-War-Since-2022-on-Future-Air-Defence-Challenges-and-Requirements.pdf

62 NATO Integrated Air and Missile Defence Centre of Excellence (2024b) The Evolving Asymmetric Threat of Unmanned Aircraft Systems. Souda: IAMD COE. Available at: https://iamd-coe.org/wp-content/uploads/2024/11/The-Evolving-Asymmetric-Threat-of-Unmanned-Aircraft-Systems-final.pdf

63 NATO Integrated Air and Missile Defence Centre of Excellence (2023) briefing on research task group SCI-SET-353, 'C-UAS mission-level modelling and simulation', 28 July. Available at: https://iamd-coe.org/wp-content/uploads/2023/09/5.-Systems-Concepts-Integration-Sensors-Electronics-Technologies-353-Research-Task-Group-SCI-%e2%80%93-SET-%e2%80%93-353-RTG-C-UAS.pdf; and report of the group's meeting in Chania, 11 to 13 June 2024. Available at: https://iamd-coe.org/?p=6600

64 NATO Integrated Air and Missile Defence Centre of Excellence (2026) report on its participation in the layered counter-UAS experiment Crucible 1-26 (Romania, April 2026), 4 May. Available at: https://iamd-coe.org/?p=7848

65 Makrygiannis, N. (2026) 'Train as you fight reinvented', in IAMD COE Annual Conference Journal 2025. Souda: IAMD COE, published July 2026. Available as note 23.

66 NATO Allied Command Transformation (2025) 'NATO and Ukraine open JATEC', February. Available at: https://www.act.nato.int/article/nato-ukraine-open-jatec/; Joint Analysis, Training and Education Centre (2026) article on its contribution to the layered counter-UAS initiative. Available at: https://www.jatec.nato.int/?p=13245

67 Reuters (2026) 'Inside Ukraine's drive to defeat the dreaded Shahed drone', 29 April, via The Star. Available at: https://www.thestar.com.my/news/world/2026/04/29/inside-ukraine039s-drive-to-defeat-the-dreaded-shahed-drone

68 Ukrainska Pravda (2026) 'Ukraine's Air Force explains rate of Shahed UAV interception by drone-based air defence crews', 30 April. Available at: https://www.pravda.com.ua/eng/news/2026/04/30/8032528/

69 Military.com (2025) 'Navy's fight in Red Sea used 220 missiles, officials say', 15 January. Available at: https://www.military.com/daily-news/2025/01/15/navys-fight-red-sea-used-220-missiles-officials-say-thats-changing.html

70 NATO Science for Peace and Security Programme (2026) Special call for proposals on counter-UAS, April. Available at: https://www.nato.int/content/dam/nato/webready/documents/sps/grants/sps-special-call-for-proposals-2026-2.pdf

DOCTRINAL AND REGULATORY MATERIAL CONSULTED

Institute of Electrical and Electronics Engineers (2012) IEEE 1278: Standard for Distributed Interactive Simulation.

Institute of Electrical and Electronics Engineers (2010) IEEE 1516: Standard for Modeling and Simulation High Level Architecture.

United States Department of Defense (2023) Directive 3000.09: Autonomy in Weapon Systems.

European Union (2024) Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence.

United Kingdom Ministry of Defence (2022) Joint Doctrine Publication 0-30: UK Air and Space Power.