Schrödinger's Security Partner: The Paradox of Measuring Security Force Assistance

Description

In quantum mechanics, Schrödinger's Cat is often mischaracterized as a simple problem of observation. In reality, the thought experiment illustrates something more unsettling: a system can exist in multiple states simultaneously, and the act of observation collapses that complexity into a single, misleading outcome. Security force assistance (SFA) suffers from an analogous problem. Partner forces are complex institutional systems shaped by politics, incentives, institutions, leadership, and threats. When the United States chooses metrics to observe its security partners, complexity collapses into a simplified performance snapshot that obscures more than it reveals.

This dynamic is related to Goodhart's Law, but more severe. When a measure becomes a target, it is a "self-defeating exercise despite its necessity." When a measure reshapes behavior, it ceases to describe reality at all. As with Heisenberg's uncertainty principle, precision in one dimension comes at the expense of understanding another. A training metric may reveal marksmanship proficiency while simultaneously masking leadership failure, political interference, or institutional fragility. What is measured is not merely recorded: it is produced.

These dynamics distort U.S. advising missions. Once a higher headquarters, such as a combatant command, designates a metric, advisors teach to that metric and partners rehearse for it. Training becomes performative. Assessments become affirmation. Many advisors we interviewed report teaching to the test, making sure partner forces "perform" so the advisor can report up-the-chain their own instructional effectiveness. Unfortunately, this does not translate into partner absorptive capacity, resilience, or will to fight. Instead, advisors are left with a manufactured, curated version of reality aligned with headquarters' reporting requirements and funding cycles. This is the McNamara Fallacy in action, where decisions are based on quantitative metrics like body counts while ignoring qualitative observations.

This long-standing issue now meets the Trump Administration's 2025 *National Security Strategy*. The administration's skepticism of foreign aid and demand for burden sharing intensify pressure for returns on investment as some policies reduce security assistance. The pattern of prioritizing easily quantifiable outputs over genuine institutional growth is a systemic challenge in places like Afghanistan, Iraq, and Somalia. Given this assessment paradox, we offer improved approaches to assessing

partnerships *before* a crisis opens the box and reveals the dead cat inside.

Measurement Distorts Advising: How a Linear System Fails in a Complex World

The security assistance enterprise is built to measureâ??and is required to do so under the 2017 NDAA. The Defense Security Cooperation Agency manages programs, combatant commands select indicators, and advisors teach to the requirements. In turn, partner forces alter their behavior, and staffs optimize â??logic modelsâ?• and spreadsheets for reporting systems. The result is measurement bias being hardwired into advising itself.

Per Risa Brooks, military effectiveness is tied to political legitimacy, elite cohesion, and the stateâ??s ability to monopolize violence and authorityâ??variables difficult to quickly resolve in an irregular warfare context. Tactical proficiency rarely determines outcomes. Yet SFA assessments continue to privilege what is most visible and countable, even when those indicators say little about whether a partner force can sustain operations, resist elite capture, or survive political shock.

Policy, despite its best intentions, hardens this bias. The Pentagonâ??s Assessment, Monitoring, and Evaluation (AM&E) framework was created in 2017 to improve accountability. In practice, however, it doubles down on the McNamara Fallacy by pushing the system toward visible, countable, and short-term outputs. Such a framework treats SFA as a linear engineering problem where inputs like money and training should produce predictable outputs. This fails to treat the partner force as a complex adaptive system, where small changes can have unpredictable, cascading effects. Enduring results such as a partnerâ??s ability to demonstrate initiative, plan independently, or adapt without advisors, become tertiary concerns because they mature off-cycle and resist quantification.

Assessment must be redesigned. Effective evaluation in irregular warfare must shift from proof-seeking to hypothesis-testing, elevate qualitative judgment alongside quantitative indicators, and align measurement with strategic competition rather than tactical output. These changes are not cosmetic. They are prerequisites for building partner forces that endure beyond the reporting cycle. For years, U.S. officials gave optimistic assessments of rising troop numbers and readiness rates in Afghanistan and Iraq, while Inspector General Reports showed numerous failures. These metrics masked â??ghost soldiers,â?• systemic corruption, and a crippling dependence on contractors.

Across the Sahel, the U.S. spent over $3 billion on security assistance, yet the region became a â?? global epicenter of jihadist violence,â?• according to a Congressional report. In countries such as Chad and Mali, U.S. security assistance emphasized tactical counterterrorism outputs like elite unit readiness, raid proficiency, and partner participation in joint operations. While showing progress, this security assistance masked deeper institutional decay. American-trained units became competent

enclaves within militaries hollowed out by patronage, coup-proofing, and political fragmentation. Rather than stabilizing the state, assistance inadvertently empowered officers who later seized power, while insurgent groups expanded their reach and Russian advisors displaced Western partners. The metrics reflected tactical success but the reality was a strategic failure.

Even the seemingly successful case of security assistance to Ukraine highlights the measurement dilemma. The focus of Western efforts has been the *output* of aid: over $200 billion appropriated and a hodgepodge of weapon systems delivered. These are easily counted and reported. Measuring the actual *outcome* can also be asserted and tested as a false positive: Ukraine wasnâ??t conquered by Russian military forces so assistance obviously worked.

Without Western military advisors on the ground, assessments rely on remote data, ISR, intercepted Russian communications, and Ukrainian self-reporting. While Ukraineâ??s remarkable battlefield adaptation is evident, Western military organizations struggle to measure the specific effectiveness of their training versus Ukraineâ??s own impressive capacity for innovation. According to internal European Union reports one author had access to, significant questions remain concerning how well Western-led training has actually translated to the battlefield. The metrics show a firehose of support, but they cannot fully capture the nuances of battlefield impacts or long-term institutional absorption.

A Strategic Misalignment: Measuring for the Wrong War

Some argue that security cooperation deserves a fairer evaluation. They contend critics focus on high-profile failures like Iraq and Afghanistan, while ignoring quiet, incremental successes like the access and influence gained, relationships built, and conflicts deterred. From this perspective, the problem is not the tool itself, but a lack of partner political will, something beyond American control. They call for more nuanced evaluation, not abandonment.

This argument, however, overlooks a more fundamental problem, identified not by outside critics, but by the Pentagon. A 2021 Pentagon evaluation of security cooperation revealed profound strategic misalignment. The entire security cooperation enterprise, including its assessment architecture, remained overwhelmingly focused on counterterrorism, not great power competition.

The evaluation highlighted a fundamental disconnect between the National Defense Strategyâ??s goals and the actual implementation and measurement of security cooperation programs. The metrics for success in the counterterrorism era are tactical and unit-focused: the number of raids conducted, the proficiency of a partnerâ??s special operations forces, or their ability to clear insurgents from a village. These same metrics proved misleading in the Sahel, where tactical proficiency failed to translate into strategic stability.

Success in strategic competition requires a different set of indicators. Tactical proficiency should matter less, as SFA should seek political and strategic alignment. Key questions include whether the partner chooses U.S. equipment over Chinese alternatives, which may enhance interoperability. Does their officer corps attend U.S. professional military education, internalizing democratic norms? Do they resist political and economic coercion from strategic competitors? Measuring tactical proficiency while ignoring these strategic choices creates false confidence, encouraging complacency.

Conclusion: Opening the Box

Security force assistance practitioners struggle to produce desired results because the United States measures the wrong things. With the wrong strategic *ends*, the *ways* distort everything that follows. Washington keeps trying to collapse partner development into clean indicators, forgetting that the act of measuring shapes the reality it claims to reveal. Partner forces learn to perform; advisors learn to report; headquarters learn to believe a fabricated picture. The system is self-reinforcing, confident in its progress until a crisis reveals a choreographic army.

This measurement obsession was highlighted in a 2025 DoD Inspector General report on the efficacy of Security Force Assistance Brigades. This report, like many others, demands quantifiable proof that SFA is working. But this mentality fundamentally misunderstands advising because there is no rigid checklist for effective military advising. In terms of metrics, the advisor's personality is critical to effectively working with a foreign military because advising is deeply relational, culturally embedded, and politically sensitive.

Fixing measurement obsessions means redesigning evaluations with logic and strategic focus.

First, models must shift from "proof-seeking" to "hypothesis-testing." A hypothesis-testing approach reframes assessment around institutional behavior rather than performance theater. Consider a partner force engaged against an insurgency where vehicle readiness constrains patrolling abilities. Instead of counting the number of mechanics trained, advisors might test the hypothesis that mentoring a partner logistics battalion over 18 months will improve independent maintenance and parts distribution resulting in sustained vehicle availability. If readiness does not improve, the result is not program failure but diagnostic insight about corruption in the supply chain, interference with assignments, or misaligned incentives. Success in irregular warfare sometimes depends more on identifying constraints than validating SFA programs.

Second, qualitative judgments must be elevated to stand equally with quantitative indicators. The system must formally structure and empower a seasoned advisor's narrative observations. Insights into leadership quality, unit morale, and institutional culture are often the most accurate instruments

available. Reporting systems must be created to treat advisor narratives not as â??color commentary,â?• but as key evidence. SOCIUM tracks U.S. security cooperation inputs as a classified system of record, but advisor unit inputs end up in this black hole. Finding a way to convert these into meaningful measurements will ensure the next SFA mission is effective based on trends analysis. Additionally, SOCIUM inputs must be better standardized across the joint force as 2023-2024 interviews with active duty military advisors, led to perceptions that National Guard troops were overreporting foreign military advising engagements to justify their deployments.

Third, the U.S. must measure for strategic competition, not just counterterrorism. The metrics for success against a peer adversary are not tactical but political and institutional. Key questions are not about raid proficiency but a partnerâ??s strategic alignment, resistance to rival influence, and commitment to democratic norms. The AM&E framework must be reoriented to answer these questions.

Finally, advisors should privilege indicators that partners cannot fake. Valuable indicators for irregular warfare practitioners should look beyond graduation rates, scripted exercises, and staged evaluations. U.S. policymakers must accept what experienced advisors already know: real partner capability is not a tallied list but is instead something that is witnessed. Complex human systems cannot be captured by metrics designed to simplify them.

SchrÃ¶dingerâ??s Cat endures in physics because it captures a paradox: observation changes the thing observed. Security force assistance suffers from the same logic. Washingtonâ??s choice of what to measure shapes how partners learn and how advisors teach. When measuring performance, performance results. When measuring bureaucracy, bureaucracy results.

If the United States wants a return on SFA investment, partner forces that can fight, then the Pentagon must stop treating measurement as truth and start treating it as a tool. Only then can we open the box without discovering yet another partner force that looked seemingly alive only until the moment we observed that they were not at all.

Jahara â??FRANKYâ?• Matisek, *PhD, (@JaharaMatisek) is a U.S. Air Force command pilot and nonresident research fellow at the U.S. Naval War College, Payne Institute for Public Policy, Resilience Initiative Center, and Defense Analyses and Research Corporation. He has published over one hundred and fifty articles on national security issues and is a Visiting Scholar at Northwestern University and a Co-Principal Investigator for a Defense Security Cooperation University research project.â??*

***Robert Schafer** is a strategic plans analyst at the Center for Army Lessons Learned and publishes extensively on irregular warfare-related topics, such as security force assistance and civil affairs operations. He is also a PhD candidate in international relations at Salve Regina University in Newport, RI.*

Main image generated by ChatGPT using DALLÂ·E, OpenAI (January 2026).

Date Created
2026/02/03