Contractors and Operational Testing

A Tester's Perspective

MAJ. STEVE ELLIOTT, USA

This article is a follow-up to an article by retired Army Col. John Stoddart, from the March-April 2001 issue of PM, entitled "Contractors and Operational Testing – Some Involvement is Legal and Beneficial." Stoddart makes a number of points regarding situations where it is appropriate, and even good business, to involve contractors in operational testing (OT) of the systems they build. Without question, contractors play a central and sometimes under-appreciated role in creating the technologies, capabilities, and systems around which we mold our force.

The government cannot do it alone – every phase of our materiel development and acquisition process requires our contractors' vision, expertise, and industrial capacity. This is equally true for the Production and Deployment phase, which includes OT.

Mechanics from the Army's 4th Infantry Division split a tank power pack during operational test of the Forward Repair System at Fort Hood, Texas, February 2000.

Level of Participation

Having acknowledged this, however, we must recognize that the appropriate level of contractor involvement in operational testing is variable, and depends on the nature of the event. It can range from very high to very low, or almost no involvement at all. This is so because because different kinds of operational tests and experiments serve different purposes.

Exploratory

Some tests and experiments are exploratory in nature, and may be conducted on non-production-representative systems. The Army typically uses these early in the acquisition life cycle to evaluate concepts, identify problems, and help develop requirements or address specific, anticipated issues.

Confirmatory

Others, including Initial Operational Tests (IOT), Follow-on Operational Tests (FOT), and to some extent, Limited User Tests, are confirmatory in nature. These are typically field tests of production or production-representative systems conducted after the Milestone C decision, under realistic operational conditions, to verify a system's effectiveness, suitability, and survivability when operated and maintained by typical user personnel.

When dealing with this latter category, test officers must be extremely careful, and make tough, up-front decisions about the extent to which contractors should participate. There are strong arguments for excluding contractors from some facets of OT.

Stoddart points out that "operational test and evaluation is the field test, under realistic combat conditions, of any item ... for the purpose of determining its effectiveness and sustainability ... for use in combat by typical military users; and the evaluation of the results of such test." But there is a key assumption to be made about the expected operating environment (i.e., those "realistic combat conditions"). The assumption is that OT should replicate, as closely as possible, a combat environment in which soldiers will

***Elliott** is a member of the Army Acquisition Corps and recently served as an operational test officer. He is currently assigned to the U.S. Army Operational Test Command, Fort Hood, Texas.*

U.S. Army photos

Troops from the Army's 1st Cavalry Division inventory components during operational test of the Digital Topographic Support System at Fort Hood, Texas, July 2001.

use the system. If that environment will include contractors, then their presence on the testing "battlefield" may be appropriate. If it will not, then the opposite is true.

Controlling the Environment

Stoddart contends that "strict application of the law [that prohibits persons employed by the contractor from being involved in OT] places an unnecessary 'veil of secrecy' on the whole process." Then he goes on to suggest that "[lack of] contractor involvement in the operational test phase will hinder acquisition streamlining ... [because it forces the acquisition community] to wait until the end of test before any fixes can be applied and tested."

Except in cases where operations security is a concern, government testers and evaluators should never cast a veil of secrecy over operational testing. But testers and evaluators *do* have a primary responsibility to control their test environment, and it may sometimes be necessary to restrict the groups and individuals who have access to test plans and events in order to preserve that environment. Contractor involvement may be appropriate for some exploratory tests. But this is less likely to be the case in confirmatory tests, where the object is to determine how well soldiers can use the system on their own, in an "as fielded " condition.

On Stoddart's second note, it's helpful to remember that IOTs and FOTs are not intended as tools for system development; their goal is to demonstrate conclusively that the system, *as developed*, is operationally effective, suitable, and survivable when employed by typical user soldiers in the expected operating environment. The time for partnering on system development is *before* IOT or FOT.

The differences between operational and developmental testing are critical. Developmental testing (DT) tends to be tightly controlled and executed through strictly defined procedures. This is not surprising, since one of DT's main functions is to gauge how well systems conform to precise contract specifications. In comparison, OT is relatively uncontrolled. Soldiers or units are issued the system(s) and logistics support, and then trained and tasked to conduct missions as they would in combat.

Operational testers allow soldiers to "do what soldiers will tend to do" with the system, because the "real world" operating environment most closely replicates the expected and relatively unconstrained combat environment. Why? Because operational testers look at the system as soldiers will use it, not as it complies with contract specifications. And operational testers aren't just testing the contractor's hardware and software; they're testing the comprehensive system-of-systems, comprising all the factors of Doctrine, Training, Logistics, Organization, Materiel, and Soldier Sustainment (DTLOMS). The task of OT is to confirm that it all works, all together, and all at the same time.

> If test officers and evaluators are to verify that the system under test is really operationally effective, suitable, and survivable when operated by typical user soldiers in the expected operating environment, then they must make hard and sometimes unpopular decisions as to the appropriate degree of contractor involvement.

Joint efforts between government and industry to streamline the acquisition

process have stimulated proposals to combine DT and OT into single or concurrent events. These promise a number of benefits, especially program efficiencies in time and dollars invested to bring systems to fielding. But again, we must recognize the distinction between DT and OT.

Current wisdom calls for the acquisition community to complete most or all DT events before starting OT. When done right, this prevents us from exposing soldiers to unrecognized hazards – many of which are revealed in DT processes leading to the system safety release. It also provides contractors and program managers a chance to fix the inevitable host of developmental problems, while giving all stakeholders an acceptable level of assurance that the system works as intended, before committing operational resources (troop units) to the process. In the context of DT, developmental problems are seen for what they are – simply developmental problems to be solved. They appear in an entirely different light, however, when revealed in OT. If still present at this stage, such problems may legitimately be considered failures.

Many of us in the acquisition community fail to recognize that troop units are among the most valuable, and scarcest of testing resources. We tend to see only our individual programs, and fail to notice the cumulative burden of countless such programs on troop organizations whose primary mission is warfighting rather than testing. This might sound like an argument to remove troops from the process, and combine DT and OT, but it is not. To the contrary, it reinforces the idea that DT should generally be completed first, to make the best use of this scarce resource. The bottom line remains: OT must be performed by *real* troops, in *real* units, and in their *real* environments to produce useful results. DT, as currently conducted, cannot provide that environment.

Maintaining Test Integrity

Operational testers and evaluators face a couple of particularly hard tasks. One is assessing the degree to which the soldier-system performance *in OT* predicts its performance *in combat* (two different environments). Another is designing and executing an unbiased test while under pressure from PMs, combat developers, user units, contractors, and other stakeholders, to accommodate their unique and sometimes divergent interests. Seemingly innocuous environmental factors, like the presence or absence of contractor personnel, can contribute to big differences in performance.

> Operational Testing must be performed by *real* troops, in *real* units, and in their *real* environments to produce useful results. Developmental Testing, as currently conducted, cannot provide that environment.

The fact is that the presence of contractors, other onlookers, and participants at the test site does affect the performance of OT. Everyone present has some influence not only on how an OT is run, but also on its outcome. Part of the job of the test officer and evaluator is to minimize those influences, to ensure they don't improperly bias the test outcome. Like the Marines landing in Somalia, test user units and teams can't help but perform differently under the critical gaze of onlookers. It's challenging enough to structure operational tests so that the testers themselves don't influence the outcome. Contractors and other personnel, who aren't part of the combat scenario, only add to this difficulty.

Let's look at *why* such presence can be a problem. As noted, operational testers and evaluators are chartered to independently assess their system's performance in the full context of DTLOMS. Their job is to run the test and let the system stand on its own merits – in other words, to "let the chips fall where they may." But true independence is hard to achieve amid the array of conflicting interests held by various stakeholders in the system. Consider, for example:

- *The Program Manager (PM)*: Typical Program, Project or Product Managers balance a complex array of factors to keep their programs within acceptable limits of cost, schedule, performance, and risk. They are often forced, and expected, to trade off various factors of performance to accommodate fairly rigid cost and schedule constraints. But it's easy for those PMs to forget that operational testers have a different mandate. Any reasonably aggressive PM will try to pressure those testers to design and conduct their events in a manner that conforms to the PM's particular program constraints.

- *The Combat Developer:* The combat developer and operational tester have mandates to ensure that the system under test works for typical user-soldiers in their expected operating environment. Operational testers build most of their test around doctrine and scenarios approved by the combat developer. In other words, the combat developer is the central figure in defining that expected operating environment. But doctrine evolves over time, and changes with other factors affecting the force. To this extent, user-soldiers also define the expected operating environment.

  Operational testers must give a lot of weight to the way troops naturally tend to use their systems. Testers can find themselves at odds with combat developers in cases where doctrine doesn't match the way soldiers naturally tend to use the systems we give them.

- *The Contractor:* Despite Stoddart's assertion that "the contractor's No. 1

concern is to field the best possible piece of equipment," a typical contractor's first concern is *running a profitable business.* Fielding the best equipment isn't always the same as running a profitable business – at least to the extent that the typical user-soldier defines "best." Like our Army PMs, contractors balance cost, schedule, performance, and risk, along with a host of other requirements such as providing adequate return on their shareholders' investments. This can be particularly challenging where learning curves are involved (i.e., emerging technologies, novel applications of existing methods, or low-production rates).

The rigid constraints of many government contracts further complicate this delicate balancing act. A test officer then might reasonably expect the contractor's priorities to rank somewhat as follows: 1) make a profit, 2) meet the terms of the contract, 3) avoid actions that might threaten future business, and 4) make the best possible piece of equipment. Like the other stakeholders, it's easy for contractors to forget that operational testers have a different mandate and try to pressure both the PM and the operational testers to design and conduct test events in ways that conform to their own constraints.

- *The User* (organizations slated to receive the fielded system): Users are in a tough spot. They know the shortcomings of their existing system, and they generally have a good understanding of what it takes for a new system to do the job. Users quickly recognize the shortcomings of new systems – often as early as new equipment orientation or training, and well before the actual operational test.

  But they also know that fixing those shortcomings may be slow and costly, and that corrective efforts can delay fielding. So they weigh the costs and benefits of receiving an "imperfect" new system against the costs and benefits of the old. The manner in which users deal with this assessment (and this is a sensitive issue) goes a long way toward determining how the system will appear to perform under test. It's not unheard of for a test user unit to say, "Whatever they give us will be better than what we have now, so let's do what it takes to make this new system look good in the OT." Likewise, users who are predisposed against the system can, if unchecked, act to make that system "fail."

In the middle of all this, test officers retain their charter to gauge total-system performance unencumbered by considerations of the PM's program costs or schedule, the supplier's contract terms, or the user's predisposition.

Hard Choices

Stoddart confidently states, "The benefits of operational testing are obvious to everyone." Regrettably, this isn't always so. Many view operational testing more as an obstacle to be overcome rather than a beneficial part of the process to develop, field, and sustain our systems. This seems especially true when the test operations or methodology conflict with prevailing special interests. If test officers and evaluators are to navigate this minefield – if they are to verify that the system under test is really operationally effective, suitable, and survivable when operated by typical user soldiers in the expected operating environment – then they must accept this reality. They must make hard and sometimes unpopular decisions as to the appropriate degree of contractor involvement.

Continuing the Dialogue

Stoddart and the Industrial Committee on Operational Test and Evaluation have opened an important window for discussion of the issues affecting operational testing. It would be a mistake to think this is just about a test or series of tests called OT. Ultimately, these issues are at the heart of the all-important decision as to whether materiel systems are ready to support our troops in battle.

The question of contractor involvement in OT deserves a long and spirited dialogue among our acquisition and testing commands, acquisition schools, combat development centers, and contractor community. Undoubtedly, ways to streamline the cumulative process will or have already achieved some measure of success, especially for Acquisition Category III and IV systems, which often move from concept to production in the short span of two to five years. But in pursuing these, we must err on the side of maintaining the integrity and independence of OT.

Editor's Note: The author welcomes questions or comments on this article. Contact him at *elliottsteven@otc.army.mil.*

DEFENSE ACQUISITION HISTORY PROJECT

Initiated by the Historical Office, Office of the Secretary of Defense, the Defense (OSD) Acquisition History Project is a six-year effort that will produce a five-volume chronological history of defense acquisition from the end of World War II to the present. A sixth volume will contain documentary material. The chronological volumes will focus on OSD-level policy direction and Service-level execution of defense acquisition. The target audiences for these volumes and the project's symposia and lectures will be drawn from the ranks of the U.S. Government's defense acquisition and history communities.

The project began in October 2000 with the U.S. Army Center of Military History as its executive agent. Science Applications International Corporation (SAIC) of McLean, Va., is managing the project in its initial year. For more information on the DAH Project, visit *http://www.army.mil/cmh-pg/acquisition/acqhome.htm.*