

# Baseline Evaluation – Guidance

---

## **Summary**

The primary objective of Defense acquisition is to acquire quality products that satisfy user needs with measurable improvements to mission capability and operational support, in a timely manner, and at a fair and reasonable price.

One way to determine “measurable improvements” is through comparative or baseline evaluation, which compares unit mission accomplishment when equipped with the new system to unit mission accomplishment when equipped with the legacy system. This comparison is in addition to assessing a new system’s achievement of its required performance characteristics.

Typically, many uncontrollable variables are present during operational testing, especially in force-on-force exercises. Areas where commonality should be sought between trials in order to enable valid comparisons include: the mission to be accomplished; the size, organization, and capability of the enemy force; the terrain (or environment) where the test is conducted; the size, organization, and capability of the Blue forces; and time available to accomplish the mission.

## **Best Practices**

Conduct a side-by-side operational test, as during the Stryker IOT&E, with a unit equipped Stryker and another unit equipped with the legacy system.

In the M2A3 Bradley IOT&E, the M2A3 Bradley unit conducted operations against a M2A1 Bradley unit for a head-to-head comparison.

In the Apache Block III IOT&E, mission performance of an Air Weapons Team (AWT) with Apache Block III was compared to mission performance of an AWT with legacy Block 2 Apache. The operational effectiveness of improved Block III flight performance was decisively demonstrated when the AWT with legacy Apache could not successfully accomplish a mission in high, hot, windy conditions that was successfully accomplished by the Block III AWT with power to spare.

The Task Force XXI Advanced Warfighting Experiment at the National Training Center used three NTC rotations to establish a baseline for normal unit performance.

Analysis of Alternatives can be helpful in determining the factors and levels to be examined, and also for estimating baseline force performance in field trials.

The Navy made effective use of hardware-in-the-loop (HWIL) M&S to support the evaluation of heavyweight torpedoes. The OT objective was to assess a form-fit-functional replacement of the weapon’s Guidance and Control section running a rehosted version of the tactical software. The HWIL simulation allowed testers to run both the legacy and upgraded systems through a series of identical scenarios and compare the results. A limited number of in-water trials were conducted to validate the model and verify system suitability. This M&S

approach provided a large, well-controlled data sample to compare the performance of the two variants in similar conditions.

**Reference**

[Test and Evaluation Policy Revisions, DOT&E, December 22, 2007](#)