Evaluating and Improving Test Effectiveness Using Statistical Test Optimization

Kedar M. Phadke
Phadke Associates
kedar@phadkeassociates.com

"Approximately 50% of programs entering Initial Operational Test and Evaluation (IOT&E) have not been evaluated as Operationally Effective and Operationally Suitable."

Hon. Kenneth Krieg
Under Secretary of Defense
Acquisition, Technology, and Logistics
April 30, 2007

Defense Science Board (DSB) study from 2009 had similar findings.

Industry Perspective of the T&E Challenge

PHADKE

Raytheon

The Test Optimization Challenge

"We are being challenged by our customers and by the marketplace to develop and deliver increasingly complex systems with smaller performance margins that meet the user's requirements in the shortest time, with high reliability, open and adaptable, and at the lowest cost."

Given this challenge, there is more pressure than ever on Integration, Verification & Validation activities to deliver performance results on time and within budget.

Industry studies have estimated test and rework to represent between 30 and 50% of product development costs. Given this investment, test represents fertile ground for optimization techniques. Typical benefits of statistically-based test optimization include:

- **Increased Mission Assurance**
- **Optimized performance**
- **Improved cycle time**
- **Increased Productivity**
- **Reduced cost**

Raytheon IDS Quality & Process Performance Objectives established for:

- On-time Deliverables
- Cost and Schedule Performance
- Engineering Productivity
- Reducing Rework

Weapons Fire and Detection System Program (WFDS)

Note: Information has been changed for this presentation

Program Summary

- Program was upgrading a system that has capability to detect type and direction of enemy fire under many operating scenarios. Operating scenarios include different platforms, arms types, times of day, weather conditions, single/multiple threats, mounting methods, etc.
- The program plans to operationally evaluate prototypes to demonstrate performance for the next phase of acquisition.
- T&E Questions
  - Does the test plan provide high confidence that the delivered prototype will provide required performance (Is the test plan truly validating performance of the system under all intended operating conditions)? What is the risk?
  - What corrective actions can technical staff take to improve efficacy of the test plan to reduce risk?
  - Is there a more optimal way to test the weapons targeting system to reduce test cost and schedule?

Program Status

- **Already developed a suite of tests (71 test scenarios)**
  - **Used best practices, expert guidance, and legacy test scenarios**
- **Systematic review of the tests uncovered multiple operational test gaps**
  - **750+ test gaps identified – all deemed important**

Utilizing Statistical Test Optimization

PHADKE

- **Identify the key test parameters and their values**
- **Generate a factor table and test design**
- **Ensure that test design covers conditions of the original 71 test and eliminates test gaps**
- **Do all this while optimizing risk, cost, and schedule!**
  - **It should be easy for program staff to implement**

Test Factors and Values

PHADKE

Factor Name	No. Lvls	Dep. On	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7
Ambient Temperature, degrees F	**3**		-51	54.5	160				
Time of Day	**2**		Daylight	Night					
Weather Conditions	**4**		Clear Day	Cloudy	Rain	Sand/Dust			
Sensor position	**2**		on the move	stationary					
Platforms/Mounting Type	**7**		Ground - Tripod	Ground - UGV	Air - Fixed wing	Ground - Yoke	Ground - Gimble	Air - UAV	Air - Rotary Wing
Weapons System Mount	**2**		Not attached to Friendly Weapons System	Attached to Friendly Weapons System					
Configuration	**2**		stand alone sensor	multiple sensor with cross cueing and validation					
Power Source	**2**		Battery Power	Standard Military Power					
5.56mm threats	**4**		1 present	6 present	2 present	10 present			
5.56mm threat location	**4**		outside of range with natural barriers	within range with natural barriers	outside of range with manmade barriers	within range with manmade barriers			
7.62mm threats	**4**		1 present	2 present	6 present	10 present			
7.62mm threat location	**4**		outside of range with natural barriers	within range with manmade barriers	within range with natural barriers	outside of range with manmade barriers			
50 Cal threats	**3**		1 present	4 present	2 present				
50 Cal threat location	**4**		within range with natural barriers	outside of range with natural barriers	outside of range with manmade barriers	within range with manmade barriers			
RPG threats	**2**		1 present	2 present					
RPG threat location	**4**		within range with natural barriers	within range with manmade barriers	outside of range with manmade barriers	outside of range with natural barriers			
Mortar threats	**2**		1 present	2 present					
Mortar threat location	**4**		within range with manmade barriers	outside of range with natural barriers	within range with natural barriers	outside of range with manmade barriers			
MANPAD threats	**2**		1 present	2 present					
			outside of range	within range with	within range with	outside of range			

- **Test parameters include sensor configurations, threat profiles, interfacing systems, output types, and environmental variables**

WFDS Test Design

PHADKE

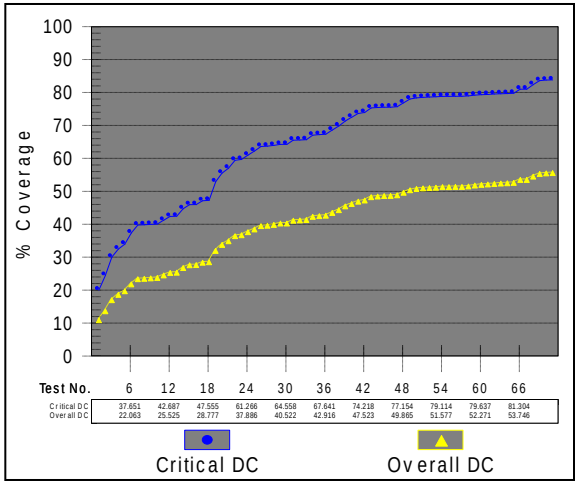
Ambient Temperature, degrees F	Time of Day	Weather Conditions	Sensor position	Platforms/Mounting Type	Weapons System Mount	Configuration	Power Source	5.56mm threats	5.56mm threat location	7.62mm threats	7.62mm threat location	50 Cal threats	50 Cal threat location	RPG threats	RPG threat location	Mortar threats	Mortar threat location	MANPAD threats	MANPAD threat location	Data Output type	Image Options
-51	Daylight	Clear Day	on the move	Ground - Tripod	Not attached to	stand alone	Battery Power	1 present	outside of range	1 present	outside of range	1 present	within range with	1 present	within range with	1 present	outside of range	1 present	outside of range	On screen image	Analog
-51	Night	Sand/Dust	stationary	Ground - Tripod	Attached to	multiple sensor	Battery Power	10 present	within range with	10 present	outside of range	1 present	within range with	1 present	outside of range	1 present	outside of range	1 present	outside of range	On screen image and UAME file	Analog
-51	Daylight	Sand/Dust	on the move	Ground - Tripod	Attached to	multiple sensor	Battery Power	10 present	within range with	6 present	within range with	4 present	outside of range	2 present	outside of range	2 present	outside of range	2 present	within range with	On screen image	Digital
-51	Night	Clear Day	stationary	Ground - Tripod	Not attached to	stand alone	Battery Power	1 present	outside of range	2 present	within range with	4 present	outside of range	2 present	outside of range	2 present	outside of range	2 present	outside of range	On screen image and UAME file	Digital
160	Daylight	Rain	stationary	Ground - Tripod	Not attached to	multiple sensor	Standard Military	2 present	outside of range	10 present	outside of range	2 present	outside of range	1 present	within range with	2 present	outside of range	2 present	outside of range	On screen image	Digital
160	Night	Cloudy	on the move	Ground - Tripod	Attached to	stand alone	Standard Military	6 present	within range with	1 present	outside of range	2 present	outside of range	1 present	outside of range	2 present	within range with	2 present	outside of range	On screen image and UAME file	Digital
54.5	Daylight	Cloudy	stationary	Ground - Tripod	Attached to	stand alone	Standard Military	6 present	within range with	2 present	within range with	1 present	within range with	2 present	outside of range	1 present	outside of range	1 present	within range with	On screen image and UAME file	Analog
54.5	Night	Rain	on the move	Ground - Tripod	Not attached to	multiple sensor	Standard Military	2 present	outside of range	6 present	within range with	1 present	within range with	2 present	within range with	1 present	within range with	1 present	within range with	On screen image and UAME file	Analog
160	Night	Sand/Dust	stationary	Air - Fixed wing	Attached to	multiple sensor	Standard Military	2 present	outside of range	2 present	within range with	1 present	outside of range	2 present	outside of range	1 present	outside of range	2 present	within range with	On screen image	Digital
160	Daylight	Clear Day	on the move	Air - Fixed wing	Not attached to	stand alone	Standard Military	6 present	within range with	6 present	within range with	1 present	outside of range	2 present	within range with	1 present	within range with	2 present	within range with	On screen image and UAME file	Digital
54.5	Night	Clear Day	stationary	Air - Fixed wing	Not attached to	stand alone	Standard Military	6 present	within range with	10 present	outside of range	4 present	within range with	1 present	within range with	2 present	outside of range	1 present	outside of range	On screen image	Analog
54.5	Daylight	Sand/Dust	on the move	Air - Fixed wing	Attached to	multiple sensor	Standard Military	2 present	outside of range	1 present	outside of range	4 present	within range with	1 present	outside of range	2 present	within range with	1 present	outside of range	On screen image and UAME file	Analog
-51	Night	Cloudy	on the move	Air - Fixed wing	Attached to	stand alone	Battery Power	1 present	outside of range	6 present	within range with	2 present	within range with	2 present	outside of range	2 present	outside of range	1 present	within range with	On screen image and UAME file	Analog
-51	Daylight	Rain	stationary	Air - Fixed wing	Not attached to	multiple sensor	Battery Power	10 present	within range with	2 present	within range with	2 present	within range with	2 present	outside of range	1 present	within range with	2 present	outside of range	On screen image and UAME file	Analog
-51	Night	Rain	on the move	Air - Fixed wing	Not attached to	multiple sensor	Battery Power	10 present	within range with	1 present	outside of range	1 present	outside of range	1 present	within range with	1 present	outside of range	2 present	outside of range	On screen image and UAME file	Digital
-51	Daylight	Cloudy	stationary	Air - Fixed wing	Not attached to	stand alone	Battery Power	1 present	outside of range	10 present	outside of range	1 present	outside of range	1 present	outside of range	2 present	outside of range	2 present	outside of range	On screen image and UAME file	Digital
-51	Daylight	Rain	on the move	Ground - UGV	Attached to	multiple sensor	Standard Military	6 present	outside of range	10 present	within range with	4 present	outside of range	1 present	outside of range	1 present	within range with	1 present	outside of range	On screen image and UAME file	Analog
-51	Night	Cloudy	stationary	Ground - UGV	Not attached to	stand alone	Standard Military	2 present	within range with	1 present	within range with	4 present	outside of range	1 present	within range with	1 present	within range with	1 present	within range with	On screen image	Analog
-51	Daylight	Cloudy	on the move	Ground - UGV	Not attached to	stand alone	Standard Military	2 present	within range with	2 present	outside of range	1 present	within range with	2 present	within range with	2 present	outside of range	2 present	outside of range	On screen image and UAME file	Digital
-51	Night	Rain	stationary	Ground - UGV	Attached to	multiple sensor	Standard Military	6 present	outside of range	6 present	outside of range	1 present	within range with	2 present	outside of range	2 present	within range with	2 present	outside of range	On screen image and UAME file	Digital
54.5	Daylight	Clear Day	stationary	Ground - UGV	Attached to	stand alone	Battery Power	10 present	outside of range	1 present	within range with	1 present	within range with	1 present	outside of range	2 present	within range with	2 present	outside of range	On screen image	Digital
54.5	Night	Sand/Dust	on the move	Ground - UGV	Not attached to	multiple sensor	Battery Power	1 present	within range with	10 present	within range with	1 present	within range with	1 present	within range with	2 present	outside of range	2 present	within range with	On screen image and UAME file	Digital
160	Daylight	Sand/Dust	stationary	Ground - UGV	Not attached to	multiple sensor	Battery Power	1 present	within range with	6 present	outside of range	2 present	outside of range	2 present	within range with	1 present	within range with	1 present	outside of range	On screen image	Analog
160	Night	Clear Day	on the move	Ground - UGV	Attached to	stand alone	Battery Power	10 present	outside of range	2 present	outside of range	2 present	outside of range	2 present	outside of range	1 present	outside of range	1 present	outside of range	On screen image	Analog
54.5	Night	Cloudy	stationary	Ground - Yoke	Not attached to	stand alone	Battery Power	10 present	outside of range	6 present	outside of range	4 present	within range with	2 present	within range with	1 present	within range with	2 present	outside of range	On screen image and UAME file	Digital
54.5	Daylight	Rain	on the move	Ground - Yoke	Attached to	multiple sensor	Battery Power	1 present	within range with	2 present	outside of range	4 present	outside of range	1 present	within range with	2 present	outside of range	1 present	outside of range	On screen image and UAME file	Digital
160	Night	Rain	stationary	Ground - Yoke	Attached to	multiple sensor	Battery Power	1 present	within range with	1 present	within range with	1 present	outside of range	1 present	outside of range	2 present	within range with	2 present	within range with	On screen image	Analog
160	Daylight	Cloudy	on the move	Ground - Yoke	Not attached to	stand alone	Battery Power	10 present	outside of range	10 present	within range with	1 present	outside of range	1 present	within range with	2 present	outside of range	1 present	within range with	On screen image and UAME file	Analog
-51	Night	Sand/Dust	on the move	Ground - Yoke	Not attached to	multiple sensor	Standard Military	6 present	outside of range	2 present	outside of range	1 present	outside of range	2 present	within range with	2 present	outside of range	1 present	outside of range	On screen image and UAME file	Analog
-51	Daylight	Clear Day	stationary	Ground - Yoke	Attached to	stand alone	Standard Military	2 present	within range with	6 present	outside of range	1 present	outside of range	2 present	within range with	2 present	outside of range	1 present	outside of range	On screen image and UAME file	Analog
-51	Night	Clear Day	on the move	Ground - Yoke	Attached to	stand alone	Standard Military	2 present	within range with	10 present	within range with	2 present	within range with	1 present	outside of range	1 present	outside of range	2 present	within range with	On screen image	Digital
-51	Daylight	Sand/Dust	stationary	Ground - Yoke	Not attached to	multiple sensor	Standard Military	6 present	outside of range	1 present	within range with	2 present	within range with	1 present	within range with	1 present	outside of range	2 present	within range with	On screen image and UAME file	Digital
-51	Daylight	Cloudy	stationary	Ground - Gimble	Not attached to	multiple sensor	Standard Military	10 present	within range with	10 present	outside of range	2 present	within range with	2 present	outside of range	2 present	within range with	1 present	within range with	On screen image and UAME file	Digital
-51	Night	Rain	on the move	Ground - Gimble	Attached to	stand alone	Standard Military	1 present	outside of range	1 present	outside of range	2 present	within range with	2 present	within range with	2 present	outside of range	1 present	within range with	On screen image	Digital
-51	Daylight	Rain	stationary	Ground - Gimble	Attached to	stand alone	Standard Military	1 present	outside of range	2 present	within range with	1 present	outside of range	1 present	within range with	1 present	within range with	2 present	outside of range	On screen image and UAME file	Analog
-51	Night	Cloudy	on the move	Ground - Gimble	Not attached to	multiple sensor	Standard Military	10 present	within range with	6 present	within range with	1 present	outside of range	1 present	outside of range	1 present	outside of range	1 present	outside of range	On screen image only	Analog
54.5	Daylight	Sand/Dust	on the move	Ground - Gimble	Not attached to	stand alone	Battery Power	6 present	within range with	1 present	outside of range	1 present	outside of range	2 present	outside of range	1 present	outside of range	2 present	within range with	On screen image and UAME file	Analog
54.5	Night	Clear Day	stationary	Ground - Gimble	Attached to	multiple sensor	Battery Power	2 present	outside of range	10 present	outside of range	1 present	outside of range	2 present	within range with	1 present	within range with	2 present	within range with	On screen image and UAME file	Analog
160	Daylight	Clear Day	on the move	Ground - Gimble	Attached to	multiple sensor	Battery Power	2 present	outside of range	6 present	within range with	4 present	within range with	1 present	within range with	2 present	outside of range	1 present	outside of range	On screen image and UAME file	Digital
160	Night	Sand/Dust	stationary	Ground - Gimble	Not attached to	stand alone	Battery Power	6 present	within range with	2 present	within range with	4 present	within range with	1 present	outside of range	2 present	within range with	1 present	outside of range	On screen image and UAME file	Digital
54.5	Night	Rain	on the move	Air - Rotary Wing	Attached to	stand alone	Battery Power	6 present	within range with	6 present	within range with	2 present	outside of range	1 present	within range with	2 present	outside of range	2 present	outside of range	On screen image and UAME file	Analog
54.5	Daylight	Cloudy	stationary	Air - Rotary Wing	Not attached to	multiple sensor	Battery Power	2 present	outside of range	2 present	within range with	2 present	outside of range	1 present	outside of range	2 present	within range with	2 present	outside of range	On screen image and UAME file	Analog
160	Night	Cloudy	on the move	Air - Rotary Wing	Not attached to	multiple sensor	Battery Power	2 present	outside of range	1 present	outside of range	1 present	within range with	2 present	outside of range	1 present	outside of range	1 present	within range with	On screen image	Digital
160	Daylight	Rain	stationary	Air - Rotary Wing	Attached to	stand alone	Battery Power	6 present	within range with	10 present	outside of range	1 present	within range with	2 present	within range with	1 present	within range with	1 present	within range with	On screen image	Digital
-51	Night	Clear Day	stationary	Air - Rotary Wing	Attached to	multiple sensor	Standard Military	10 present	within range with	2 present	within range with	1 present	within range with	1 present	within range with	1 present	within range with	1 present	outside of range	On screen image	Digital
-51	Daylight	Sand/Dust	on the move	Air - Rotary Wing	Not attached to	stand alone	Standard Military	1 present	outside of range	6 present	within range with	1 present	within range with	1 present	within range with	1 present	outside of range	1 present	outside of range	On screen image and UAME file	Digital
-51	Night	Sand/Dust	stationary	Air - Rotary Wing	Not attached to	stand alone	Standard Military	1 present	outside of range	10 present	outside of range	4 present	outside of range	2 present	outside of range	2 present	within range with	2 present	within range with	On screen image and UAME file	Analog
-51	Daylight	Clear Day	on the move	Air - Rotary Wing	Attached to	multiple sensor	Standard Military	10 present	within range with	1 present	outside of range	4 present	outside of range	2 present	outside of range	2 present	within range with	2 present	within range with	On screen image and UAME file	Analog
-51	Daylight	Sand/Dust	stationary	Air - UAV	Attached to	stand alone	Battery Power	2 present	within range with	1 present	within range with	1 present	outside of range	2 present	within range with	2 present	outside of range	2 present	outside of range	On screen image and UAME file	Digital
-51	Night	Clear Day	on the move	Air - UAV	Not attached to	multiple sensor	Battery Power	6 present	within range with	10 present	within range with	1 present	outside of range	2 present	outside of range	2 present	within range with	1 present	outside of range	On screen image and UAME file	Digital
-51	Daylight	Clear Day	stationary	Air - UAV	Not attached to	multiple sensor	Battery Power	6 present	outside of range	6 present	outside of range	2 present	within range with	1 present	outside of range	1 present	within range with	2 present	outside of range	On screen image and UAME file	Analog
-51	Night	Sand/Dust	on the move	Air - UAV	Attached to	stand alone	Battery Power	2 present	within range with	2 present	outside of range	2 present	within range with	1 present	within range with	1 present	within range with	2 present	within range with	On screen image and UAME file	Analog
160	Daylight	Cloudy	on the move	Air - UAV	Attached to	multiple sensor	Standard Military	1 present	within range with	10 present	within range with	4 present	within range with	2 present	within range with	1 present	within range with	2 present	outside of range	On screen image and UAME file	Analog
160	Night	Rain	stationary	Air - UAV	Not attached to	stand alone	Standard Military	10 present	outside of range	1 present	within range with	4 present	within range with	2 present	outside of range	1 present	outside of range	2 present	outside of range	On screen image	Analog
54.5	Daylight	Rain	on the move	Air - UAV	Not attached to	stand alone	Standard Military	10 present	outside of range	2 present	outside of range	1 present	outside of range	1 present	outside of range	2 present	within range with	1 present	within range with	On screen image	Digital
54.5	Night	Cloudy	stationary	Air - UAV	Attached to	multiple sensor	Standard Military	1 present	within range with	6 present	outside of range	1 present	outside of range	1 present	outside of range	2 present	within range with	1 present	within range with	On screen image and UAME file	Digital
160					Not attached to	multiple sensor	Standard Military		within range with		outside of range		within range with		outside of range		within range with		within range with	On screen image	
160																					
54.5																					
54.5																					
-51																					
-51																					
-51																					
-51																					

- Test design was 64 tests, Resolution III, fully orthogonal.
- All conditions from original test plan AND gaps were covered

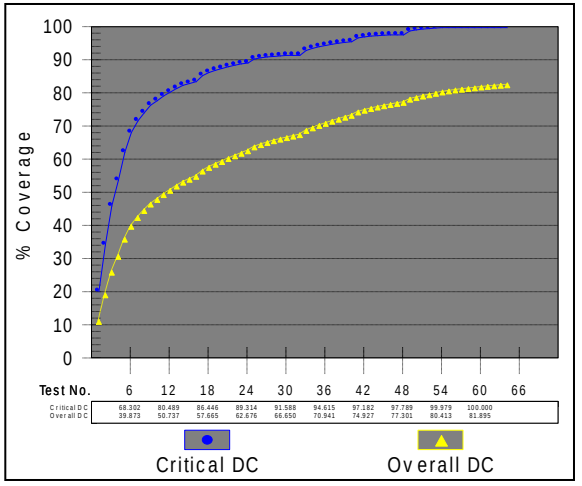
WFDS Program Benefits using Statistical Test Optimization

PHADKE

Baseline – Original 71 tests



Statistically Opt. – 64 tests



- **Improved test effectiveness (Identified and eliminated 750+ test gaps)**
- **Reduced test cost and schedule by 10%**
- **Provided a template for future deployments**

What Metrics Work for Statistical Test Optimization

Criteria for Evaluating Test Metrics

PHADKE

- **Metrics should be related to Program Objectives**
  - **Operational Effectiveness**
  - **Cost**
  - **Schedule**
- **If the Statistical Metrics are not correlated with Program Objectives, we have a problem!**

What Happens When We Use Statistical Test Power

PHADKE

- This metric is mentioned often in DOT&E literature
- Key factors that influence statistical test power
  - Acceptable error (the alpha)
  - The magnitude of the effect size you want to measure
  - Standard deviation for the population
  - Sample size
- Do we have a good handle on any of these ahead of time?
- For the WFDS example, the test power (assuming alpha = 0.10, effect size/standard deviation =0.33) was:
  - Original 71 Test Scenarios: 0.97 ← Metric looks better for original tests
  - Optimized 64 Test Scenarios: 0.95
    - Test Power calculations conducted using the JMP® Software tool from the SAS corporation

- For WFDS, statistical test power did not correlate with operational effectiveness, cost or schedule

What happens when we use Domain Coverage metric?

PHADKE

- Domain coverage is a measure of how well tests cover requirements and interop of requirements
- Commonly broken into two pieces: Critical Domain Coverage (CDC) and Overall Domain Coverage (ODC)
  - CDC Definition: How well you are covering individual requirements and interop of pairs of requirements
  - ODC Definition: How well you are covering individual requirements, interop of pairs, interop of triples, etc
- For WFDS, the Domain Coverage:
  - Original 71 Test Scenarios: CDC: 84%, ODC: 55.8%
  - Optimized 64 Test Scenarios: CDC: 100%, ODC: 82.6%

Metric looks better for optimized tests

- For WFDS, Domain Coverage correlates with operational effectiveness, cost, and schedule

Comparing Statistical Test Power and Domain Coverage

PHADKE

Test Design 1

		Threat location			
		within range with natural barriers	outside of range with natural barriers	outside of range with manmade barriers	within range with manmade barriers
50 Cal threats	1 present	XXXX	XXXX		
	2 present	XXXX	XXXX		
	4 present				

Test Design 2

		Threat location			
		within range with natural barriers	outside of range with natural barriers	outside of range with manmade barriers	within range with manmade barriers
50 Cal threats	1 present	X	X	X	X
	2 present	X	X	X	X
	4 present	X	X	X	X

What happens when we calculate test power?

Test Power of Design 1 = Test Power of Design 2

Two completely different test designs, same test power

What happens when we calculate Domain Coverage?

DC of Design 1 << DC of Design 2

DC discriminates between the test designs and favors greater test coverage and effectiveness

Defense Industry has seen significant benefits using Domain Coverage metric

PHADKE

Raytheon

Deployment Results Summary

Test	Original Test Plan	Optimized Test Plan
Subsystem Testing	28 Tests	8 Tests (71% reduction)
Systems Mission Testing	25 Missions	18 Missions (28% reduction)
Subsystem Simulation	100 Runs	40 Runs (60% reduction)
Range Testing	1036 Tests	632 tests (39% reduction)
Software Subsystem Testing	90 Tests	63 Tests (30% reduction
System Scenario Generation	8 Missions	6 Missions (25% reduction)
System MOE Testing	1600 Tests	885 tests (45% reduction)
System Testing	246 Tests	48 tests (80% reduction)

In each case, the reduction in number of test cases was achieved while maintaining or improving upon existing test coverage.

Takeaways

- **Improved Test Effectiveness and Efficiency for WFDS using Statistical Test Optimization**
  - **71 tests reduced to 64 (10% reduction in cost and schedule)**
  - **Eliminated 750+ test gaps**
- **Use statistical metrics that correlate with program metrics: test effectiveness, cost, and schedule**
  - **If they don't correlate or discriminate properly, they risk driving your programs in the wrong direction.**

References

- **Statistical Test Power calculations conducted using SAS Institute's JMP® Software. JMP is a registered trademark of SAS.**
- **Domain Coverage calculations conducted using Phadke Associates rdExpert™ Test Suite Software. rdExpert is a trademark of Phadke Associates, Inc.**
- **Reference material and articles available on http://education.phadkeassociates.net**