INSTITUTE FOR DEFENSE ANALYSES

# Perspectives on Operational Testing: Guest Lecture at Naval Postgraduate School

Vincent A. Lillard

January 2017

IDA Document
D-8333-NS

Log: H 2017-000058/1

INSTITUTE FOR DEFENSE ANALYSES
4850 Mark Center Drive
Alexandria, Virginia 22311-1882

*The Institute for Defense Analyses is a non-profit corporation that operates three federally funded research and development centers to provide objective analyses of national security issues, particularly those requiring scientific and technical expertise, and conduct related research on other national challenges.*

About this Publication

This document was prepared to support Dr. Lillard's visit to the Naval Postgraduate School, where he will provide a guest lecture to students in the T&E course. The briefing covers three primary themes: 1) evaluation of military systems on the basis of requirements and KPPs alone is often insufficient to determine effectiveness and suitability in combat conditions, 2) statistical methods are essential for developing defensible and rigorous test designs, 3) operational testing is often the only means to discover critical performance shortcomings.

Acknowledgments

The IDA Technical Review was performed by Mr. Robert R. Soule.

For More Information:

Vincent A. Lillard, Assistant Director, Operational Evaluation Division
vlillard@ida.org, 703-845-2230

Robert R. Soule, Director, Operational Evaluation Division
rsoule@ida.org, 703-845-2462

INSTITUTE FOR DEFENSE ANALYSES

IDA Document D-8333-NS

# Perspectives on Operational Testing: Guest Lecture at Naval Postgraduate School

Vincent A. Lillard

# Perspectives on Operational Testing

**Guest Lecture**

Dr. V. Bram Lillard

January 25, 2017

“There’s no requirement for that.”

“We’ll accept the risk.”

# Both types of testing are essential!

Developmental Testing: Focused on verifying functionality, technical parameters of system performance; highly structured, not typically operationally representative

Operational Testing: Characterize systems' ability to enable/improve users' mission accomplishment in operational scenarios under realistic combat conditions

# Key Principles for OT, in 3 Acts

Testing merely to determine if KPPs or requirements have been satisfied often does not capture improvements in mission accomplishment, or achievement of intended capabilities in an operational environment

Statistical methods (e.g., Design of Experiments) are essential for designing rigorous and defensible operational tests, determining quantitatively why a test design is good, and allocating resources in the most efficient and powerful way

Operational testing has to be performance against expected realistic operational threats and in realistic conditions, and is often the ONLY means of identifying critical performance problems

# Act 1 – Tunnel Vision on Requirements

# Common Argument

Systems that meet their Key Performance Parameters (KPPs) are Effective

Systems that fail their KPPs are Not Effective

# Corollary…

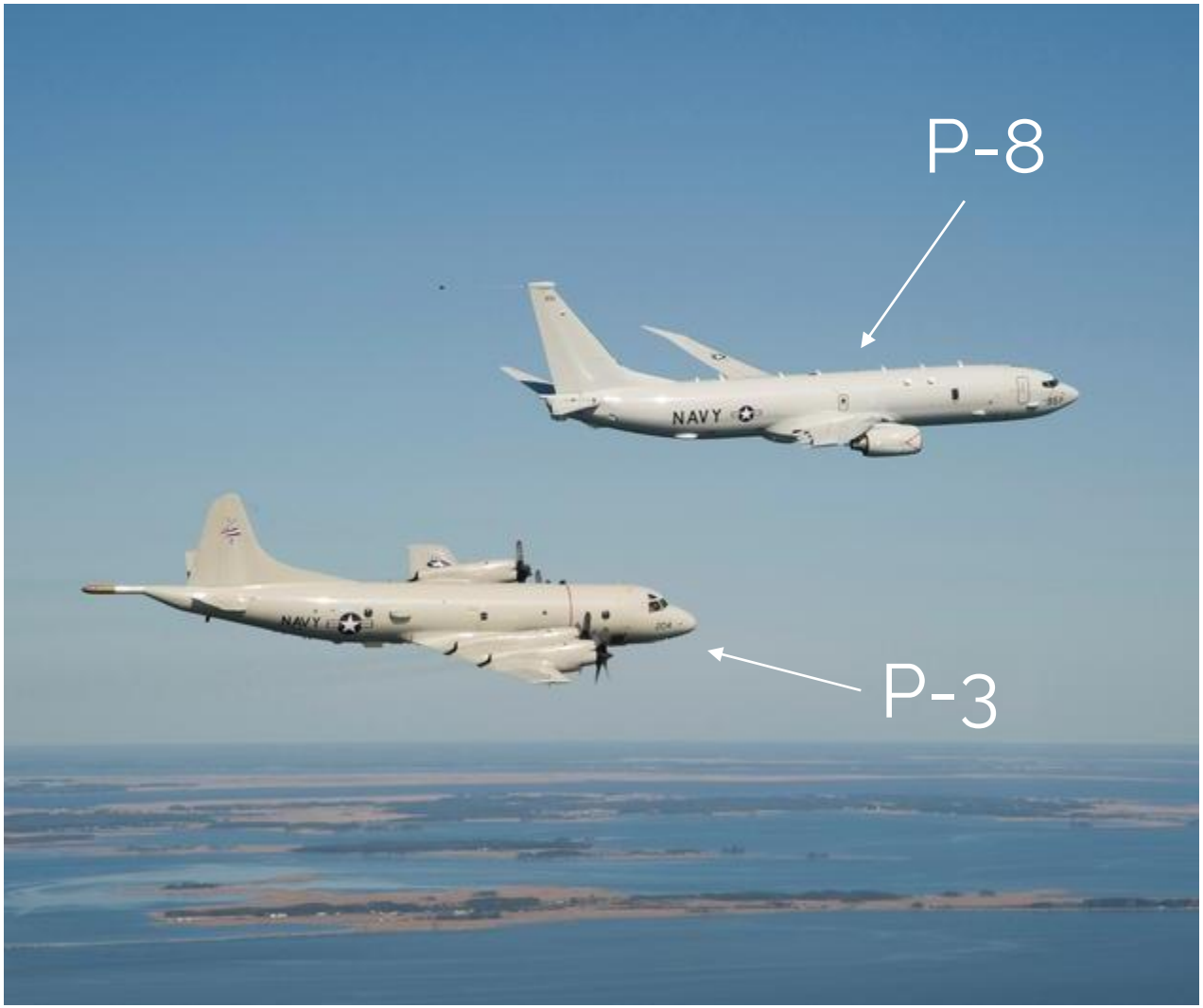
Systems that meet their Key Performance Parameters (KPPs) are Effective

Systems that fail their KPPs are Not Effective

Testers should limit their evaluation of system performance to only KPPs

# Evaluating the P-8A…

P-8A Poseidon is a maritime patrol aircraft that will replace the P-3C Orion. P-8A is based on the Boeing 737-800 airframe; its primary mission is Anti-Submarine Warfare (ASW), but also is equipped for other missions.



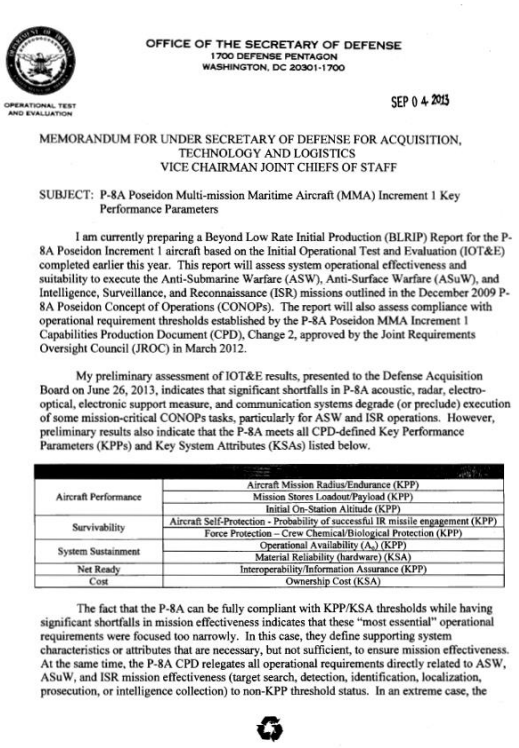
# P-8's KPPs were not mission focused

| | |
|---|---|
| Aircraft Performance | Aircraft Mission Radius/Endurance (KPP) |
| | Mission Stores Loadout/Payload (KPP) |
| | Initial On-Station Altitude (KPP) |
| Survivability | Probability of successful IR missile Engagement (KPP) |
| | Force Protection – Crew Chem/Bio Protection (KPP) |
| Sustainment | Operational Availability (KPP) |
| | Material Reliability (hardware) (KSA) |
| Net Ready | Interoperability/Information Assurance (KPP) |
| Cost | Ownership Cost (KPP) |

# KPPs could be achieved without finding and killing enemy submarines or conducting reconnaissance

<u>September 2013 DOT&E memo to Chairman of Joint Chiefs:</u>

"…could deliver an aircraft that met all the KPPs but have no mission capability whatsoever. Such an airplane would only have to be designed to be reliable, equipped with self-protection features and radios, and capable of transporting weapons and sonobuoys across the specified distances, but would not actually have to have the ability to successfully find and sink threat submarines in an Anti-Submarine Warfare mission."



OFFICE OF THE SECRETARY OF DEFENSE
1700 DEFENSE PENTAGON
WASHINGTON, DC 20301-1700

OPERATIONAL TEST AND EVALUATION

SEP 04 2013

MEMORANDUM FOR UNDER SECRETARY OF DEFENSE FOR ACQUISITION, TECHNOLOGY AND LOGISTICS
VICE CHAIRMAN JOINT CHIEFS OF STAFF

SUBJECT: P-8A Poseidon Multi-mission Maritime Aircraft (MMA) Increment 1 Key Performance Parameters

I am currently preparing a Beyond Low Rate Initial Production (BLRIP) Report for the P-8A Poseidon Increment 1 aircraft based on the Initial Operational Test and Evaluation (IOT&E) completed earlier this year. This report will assess system operational effectiveness and suitability to execute the Anti-Submarine Warfare (ASW), Anti-Surface Warfare (ASuW), and Intelligence, Surveillance, and Reconnaissance (ISR) missions outlined in the December 2009 P-8A Poseidon Concept of Operations (CONOPs). The report will also assess compliance with operational requirement thresholds established by the P-8A Poseidon MMA Increment 1 Capabilities Production Document (CPD), Change 2, approved by the Joint Requirements Oversight Council (JROC) in March 2012.

My preliminary assessment of IOT&E results, presented to the Defense Acquisition Board on June 26, 2013, indicates that significant shortfalls in P-8A acoustic, radar, electro-optical, electronic support measure, and communication systems degrade (or preclude) execution of some mission-critical CONOPs tasks, particularly for ASW and ISR operations. However, preliminary results also indicate that the P-8A meets all CPD-defined Key Performance Parameters (KPPs) and Key System Attributes (KSAs) listed below.

| | |
|---|---|
| Aircraft Performance | Aircraft Mission Radius/Endurance (KPP) |
| | Mission Stores Loadout/Payload (KPP) |
| | Initial On-Station Altitude (KPP) |
| Survivability | Aircraft Self-Protection - Probability of successful IR missile engagement (KPP) |
| | Force Protection – Crew Chemical/Biological Protection (KPP) |
| System Sustainment | Operational Availability ($A_o$) (KPP) |
| | Material Reliability (hardware) (KSA) |
| Net Ready | Interoperability/Information Assurance (KPP) |
| Cost | Ownership Cost (KSA) |

The fact that the P-8A can be fully compliant with KPP/KSA thresholds while having significant shortfalls in mission effectiveness indicates that these "most essential" operational requirements were focused too narrowly. In this case, they define supporting system characteristics or attributes that are necessary, but not sufficient, to ensure mission effectiveness. At the same time, the P-8A CPD relegates all operational requirements directly related to ASW, ASuW, and ISR mission effectiveness (target search, detection, identification, localization, prosecution, or intelligence collection) to non-KPP threshold status. In an extreme case, the

# Might seem an extreme case, but…

KPPs drive behavior

Without focus on Mission, incentive to push through OT when problems exist

Evaluation becomes little more than check in the box exercise

"The lack of KPPs/KSAs related directly to mission effectiveness will inevitably create a disconnect between the determination of operational effectiveness in test reports and the KPP/KSA compliance assessments that typically drive program reviews throughout development."

# DOT&E argued for an OT that examined whether the Navy's Concept of Employment for P-8 could be executed under realistic combat conditions

Testing went beyond simple verification of KPPs

- Navy's Operational Test Agency agreed with this approach

Navy performed realistic testing during Fleet exercises using a full set of mission systems and crew to examine their ability to find and attack submarines and perform reconnaissance using the P-8A

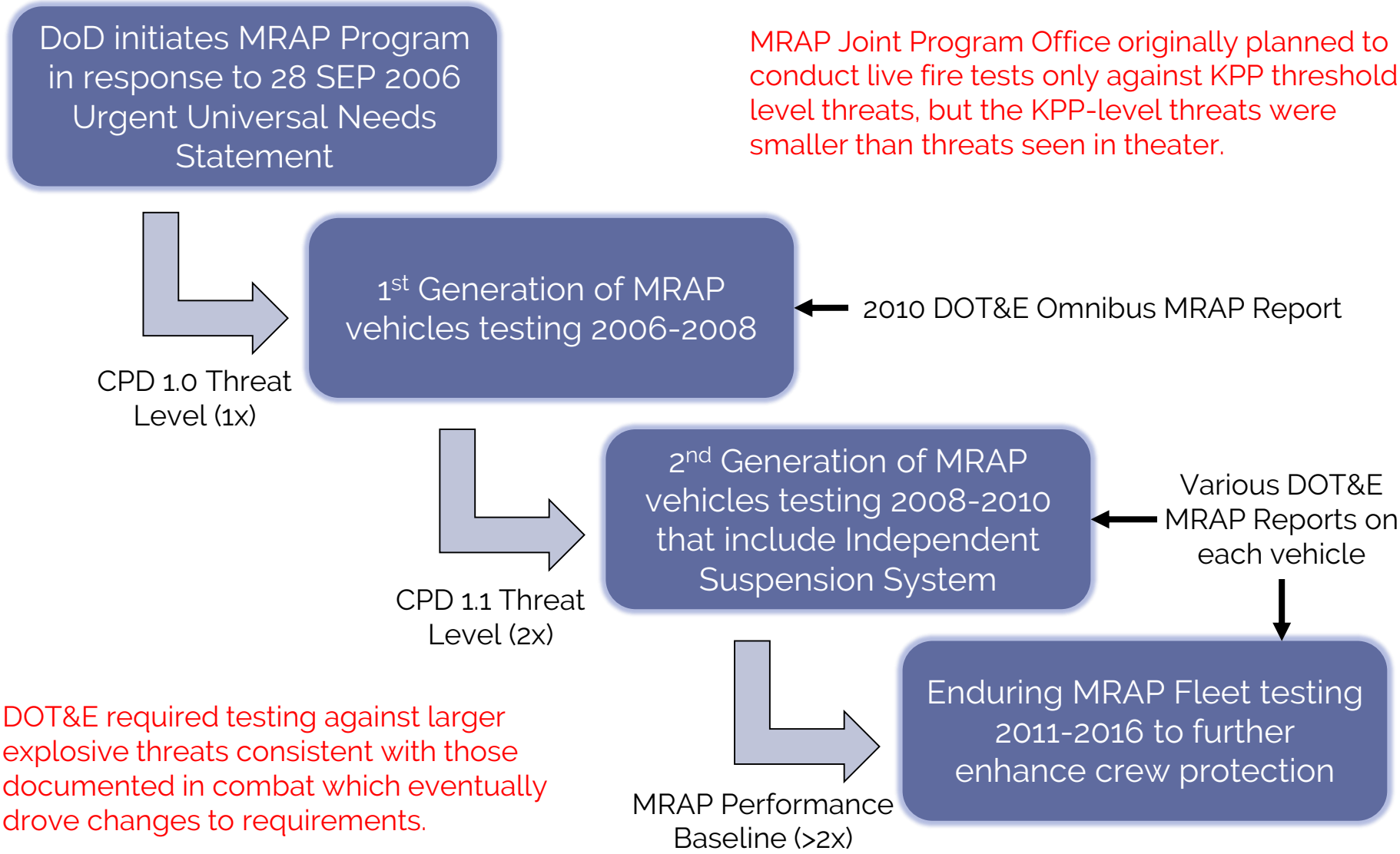
Testing revealed important deficiencies the Navy is now working to fix through improved sensors

# Mine Resistant Ambush Protected (MRAP) Vehicle Testing

Mine Resistant Ambush Protected (MRAP) vehicles are a family of vehicles designed to provide increased crew protection against battlefield threats, such as Improvised Explosive Devices (IEDs), mines, and small arms.

Because of the urgent operational need for increase crew protection against battlefield threats in Iraq and Afghanistan, multiple MRAP vehicle configurations had to be procured, tested and fielded on a highly accelerated basis.

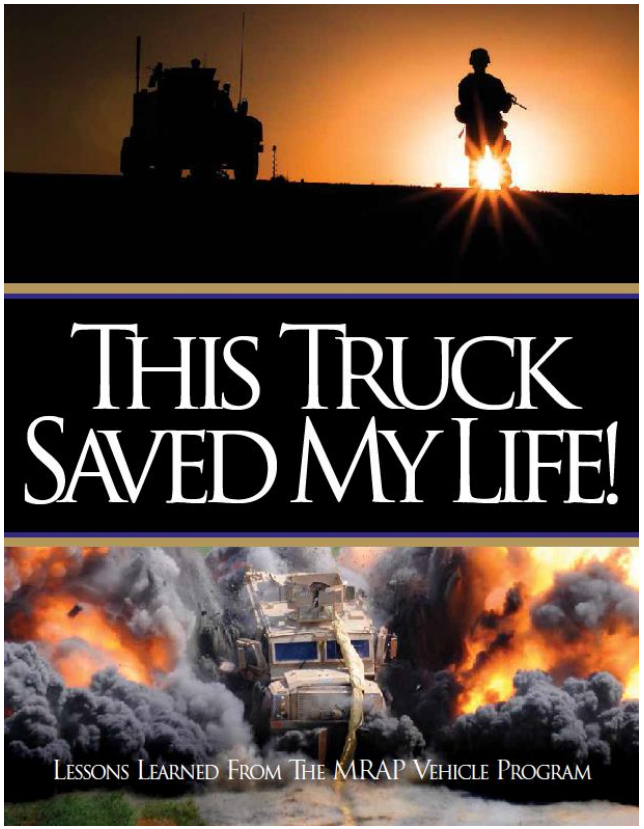
# MRAP designs evolved significantly to meet changing requirements against real world threats.



# Rapid realistic testing of MRAP vehicles improves design and saves lives.

Testing revealed:

- Significant vulnerabilities against larger, more operationally realistic threats
- Stark differences between crew protection provided by the different MRAP variants as threat sizes increased




DOT&E immediately reported these vulnerabilities and performance differences, leading the Program Office to develop, test, and implement design changes that could be retrofitted on to vehicles in theater as well as built into future production lines

The Army and the Marine Corps considered these differences when selecting the MRAP variants that would be part of the "enduring fleet"

# Act 2 – Rigorously-designed and defensible operational tests

# Key Principles for OT, in 3 Acts

Testing merely to determine if KPPs or requirements have been satisfied often does not capture improvements in mission accomplishment, or achievement of intended capabilities in an operational environment

Statistical methods (e.g., Design of Experiments) are essential for designing rigorous and defensible operational tests, determining quantitatively why a test design is good, and allocating resources in the most efficient and powerful way

Operational testing has to be performance against expected realistic operational threats and in realistic conditions, and is often the ONLY means of identifying critical performance problems

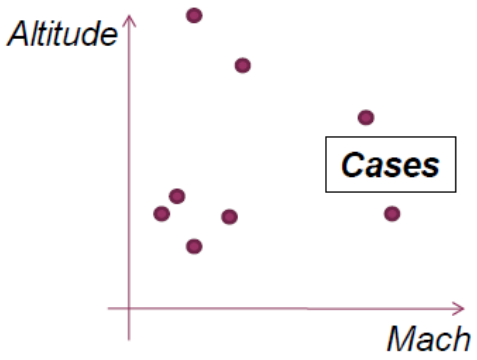
# All Tests are Designed, Some Poorly…

DWWDLT – “Do what we did last time”

# All Tests are Designed, Some Poorly…

DWWDLT – "Do what we did last time"

Special Cases / Most Critical Cases

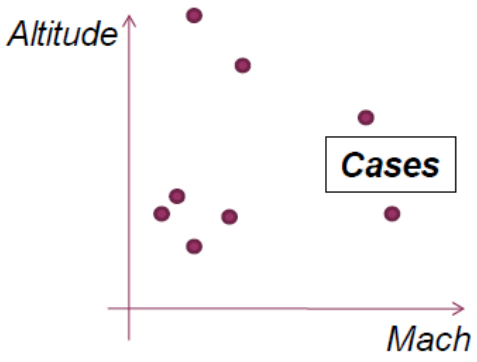


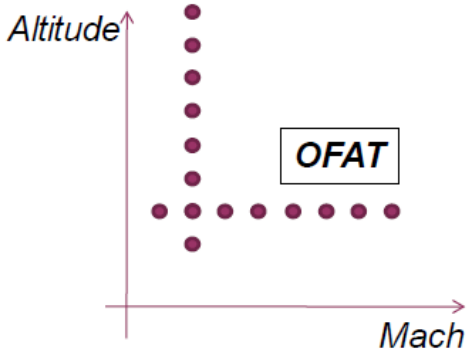
# All Tests are Designed, Some Poorly…

DWWDLT – "Do what we did last time"

Special Cases / Most Critical Cases

One-Factor-At-A-Time (OFAT)



# All Tests are Designed, Some Poorly…

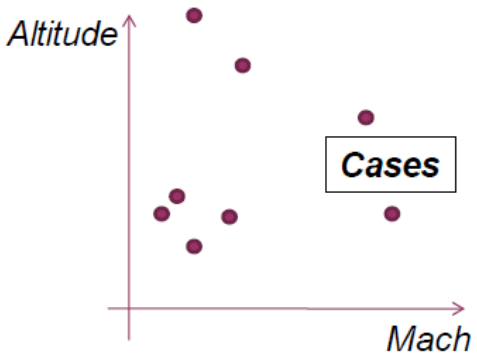
DWWDLT – "Do what we did last time"

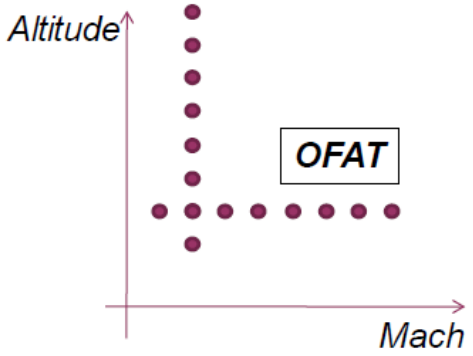
Special Cases / Most Critical Cases

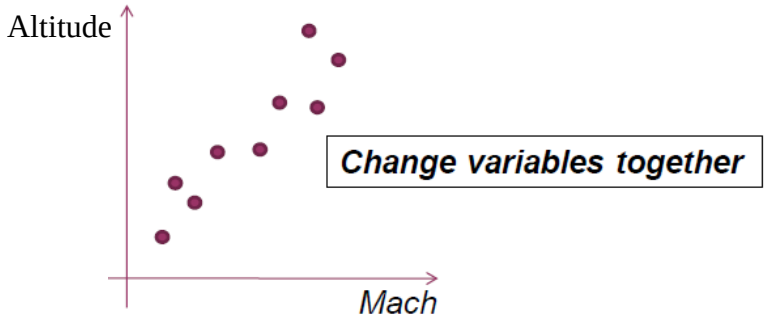
One-Factor-At-A-Time (OFAT)

Historical data – data mining

Observational studies



# All Tests are Designed, Some Poorly…

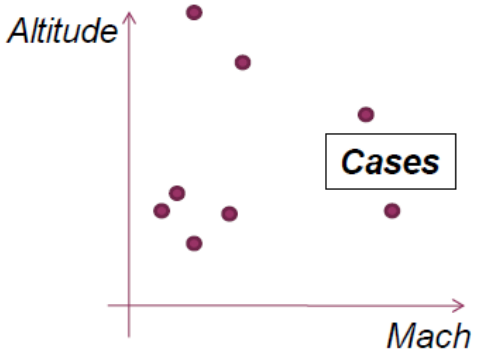
DWWDLT – "Do what we did last time"

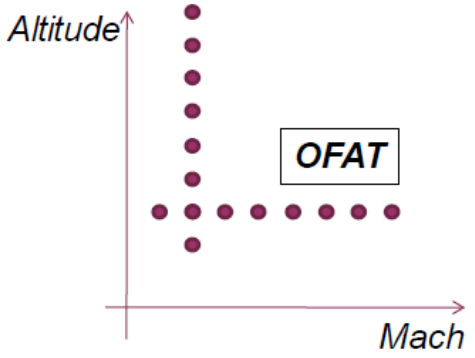
Special Cases / Most Critical Cases

One-Factor-At-A-Time (OFAT)

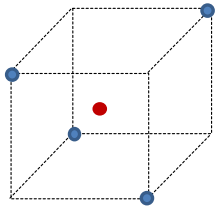
Historical data – data mining

Observational studies

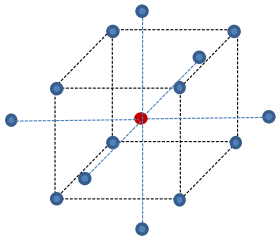


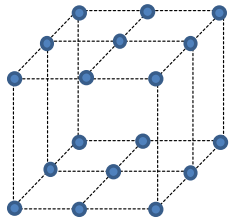

## Design of Experiments



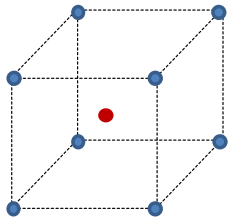
Fractional Factorial
$2^{3-1}$ design



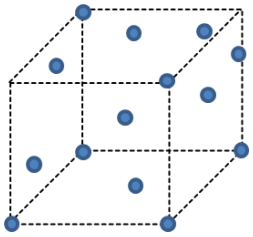
Response Surface
Central Composite design



General Factorial
3x3x2 design



2-level Factorial
$2^3$ design



Optimal Design
IV-optimal

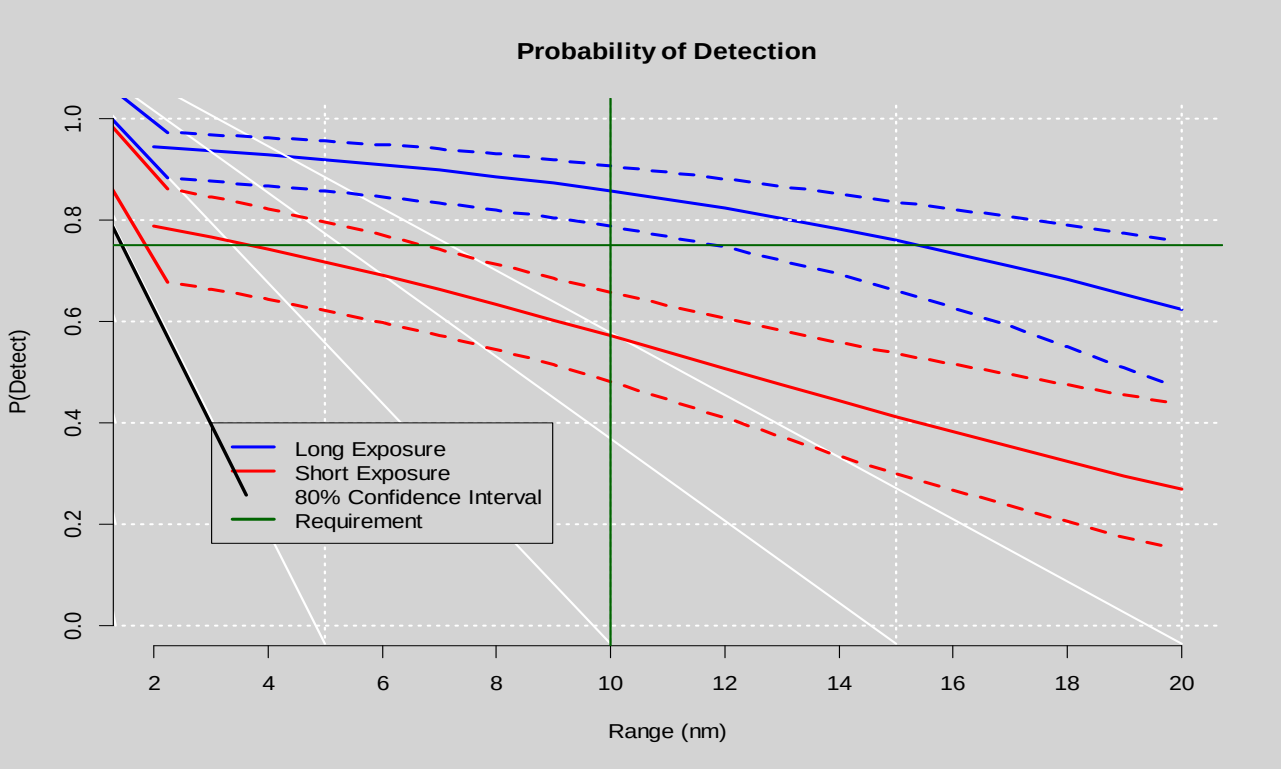
# DOE provides the analytical basis for test planning tradeoffs

Four Challenges Every Test Faces:

1. *How many? Depth of Test*
2. *Which Points? Breadth of Testing* – spanning the operational envelope
3. *How Execute? Order of Testing*
4. *What Conclusions? Test Analysis* – drawing objective, robust conclusions while controlling noise

# Operational testing should focus on characterizing capability/performance across a variety of operational conditions.

- Must be able to use test data to determine whether and to what degree system performance depends on each factor
- Determine if a system meets requirements across operational conditions



# The objective (e.g., screen/characterize) of the testing drives the complexity required in the analysis

Common Terminology:

- **Main Effect**: the change in the response produced by changing the level of a factor
- **Interaction effect**: occurs when the change in the response between the levels of one factor is not the same at all levels of the other factors (e.g., factors work in a synergistic fashion)
- **First order model**: a model form that allows for the estimation of main effects only
- **Second order model**: a model form that allows for the estimation of main effects, two-way interaction effects, and quadratic effects

First Order Model

$$y = \beta o + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon$$

Second Order Model

$$y = \beta o + \beta_1 x_1 + \beta_{11} x^2{}_1 + \beta_2 x_2 + \beta_{22} x^2{}_2 + \beta_{ij} x_i x_j + \varepsilon$$

Main Effect

Quadratic Effect

Two-way interaction

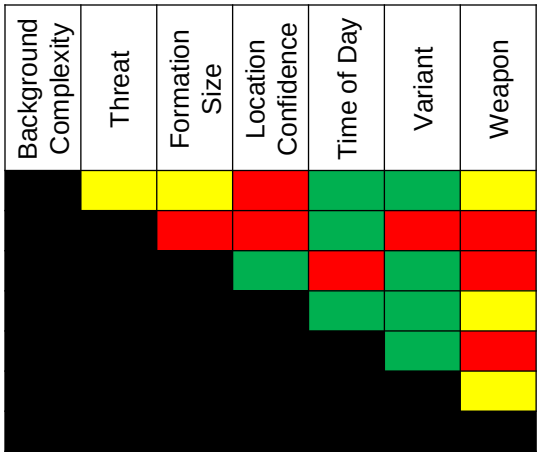
# F-35 Joint Strike Fighter
# Air-to-Ground Mission Testing

# Operational Envelope Defined – 128 possible cases

| | | | Variant - B | | | | | | | | Variant - A | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Category-B Threat | | | | Category-C Threat | | | | Category-B Threat | | | | Category-C Threat | | | |
| | | | Low TLC | | High TLC | | Low TLC | | High TLC | | Low TLC | | High TLC | | Low TLC | | High TLC | |
| | | | L | H | L | H | L | H | L | H | L | H | L | H | L | H | L | H |
| 2-Ship | Day | JDAM | | | | | | | | | | | | | | | | |
| | | LGB | | | | | | | | | | | | | | | | |
| | Night | JDAM | | | | | | | | | | | | | | | | |
| | | LGB | | | | | | | | | | | | | | | | |
| 4-Ship | Day | JDAM | | | | | | | | | | | | | | | | |
| | | LGB | | | | | | | | | | | | | | | | |
| | Night | JDAM | | | | | | | | | | | | | | | | |
| | | LGB | | | | | | | | | | | | | | | | |

# Test Design Process was a lot of work!

Test team used combination of subject matter expertise, and test planning knowledge to efficiently cover the most important aspects of the operational envelope




| | |
|---|---|
| | No significant interaction expected |
| | Significant interaction in one response |
| | Significant interaction in multiple responses |

Identified factors and their interactions and refined them to identify the most important aspects of the test design

# Determined that 21 trials was the minimum test size to adequately cover the operational space

Provided the data are used together in a statistical model approach, plan is adequate to evaluate JSF performance across the full operational envelope.

Note the significant reduction to the 128 possible conditions identified.

| | | | Variant - A | | | | | | | | Variant - B | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Category-B Threat | | | | Category-C Threat | | | | Category-B Threat | | | | Category-C Threat | | | |
| | | | Low TLC | | High TLC | | Low TLC | | High TLC | | Low TLC | | High TLC | | Low TLC | | High TLC | |
| | | | L | H | L | H | L | H | L | H | L | H | L | H | L | H | L | H |
| 2-Ship | Day | JDAM | | | 1 | | | | | | | 1 | | | | | | |
| | | LGB | | | | | | | | 1 | 1 | | | 1 | | | | |
| | Night | JDAM | 1 | | | | | | | 1 | | | | | 1 | | | |
| | | LGB | | 1 | | | | | | | | | 1 | | | 1 | | |
| 4-Ship | Day | JDAM | | | | | 1 | | | | | | | 1 | | | | |
| | | LGB | | | 1 | | | 1 | | | | | | | | | 1 | |
| | Night | JDAM | | 1 | | | | | | | | | 1 | | | | | 1 |
| | | LGB | | 1 | | | 1 | | | | | | | | | | | |

# Defensible Design enabled a departure from the TEMP

TEMP test design required 16 trials

- Would have been insufficient to examine performance in some conditions

Updated test design requires 21 trials but provides full characterization of JSF Pre-planned Air-to-Ground capabilities.

New test design answers additional questions with the addition of only 5 trials:

- Is there a performance difference between the JSF variants?
  - Do those differences only manifest themselves only under certain conditions?
- Can JSF employ both primary weapons with comparable performance?

"There's no requirement for that."

"We'll accept the risk."

# Littoral Combat Ship Radar Tracking Characterization against Small Boats

# Navy initially planned a case-based test program

Run conditions were selected to collect track quality data for specific engineering cases of interest

| Event | Number Targets | Spacing | Radar Mode | Weave Type | Pattern | Aviation |
|---|---|---|---|---|---|---|
| 1A | 8 | 750 | A | None | abreast | Any |
| 1B | 8 | 750 | A | None | abreast | Any |
| 2A | 8 | 600 | A | None | abreast | Any |
| 2B | 8 | 600 | A | None | abreast | Any |
| 3A | 8 | 450 | A | None | abreast | Any |
| 3B | 8 | 450 | A | None | abreast | Any |
| 4A | 8 | 300 | A | None | abreast | Any |
| 4B | 8 | 300 | A | None | abreast | Any |
| 5A | 8 | 150 | A | None | abreast | Any |
| 5B | 8 | 150 | A | None | abreast | Any |
| 6A | 8 | 50 | A | None | abreast | Any |
| 6B | 8 | 50 | A | None | abreast | Any |
| 7A | 1 | na | A | A | none | 60R |
| 7B | 1 | na | A | B | none | 60R |
| 7C | 1 | na | A | C | none | 60R |
| 7D | 1 | na | A | D | none | 60R |
| 7E | 1 | na | A | E | none | 60R |
| 7F | 1 | na | A | F | none | 60R |
| 7G | 1 | na | A | G | none | 60R |
| 8 | 6 | Multiple | A | None | Line | 60R |
| 9A | 1 | na | A | B | none | none |
| 9B | 1 | na | B | B | none | none |
| 9C | 1 | na | C | B | none | none |
| 10A | 4 | 200 | A | B | abreast | none |
| 10B | 4 | 200 | B | B | abreast | none |
| 10C | 4 | 200 | C | B | abreast | none |
| 11 | 6 | 200 | A | None | Line | 60R |
| 12 | 9 | Multiple | A | B | Blob | None |
| 13A | 9 | 200 | A | B | Diamond | None |
| 13B | 9 | 200 | B | B | Diamond | None |
| 13C | 9 | 200 | C | B | Diamond | None |
| 14 | 7 | 200 | A | B | Delta | None |

# A closer examination reveals some problems

32 Runs

Primarily One-Factor-At a-Time (OFAT)

- E.g., fix radar mode, fix pattern, fix weave, vary spacing
- Then: vary radar mode, fix pattern, fix weave, fix spacing
- Etc.

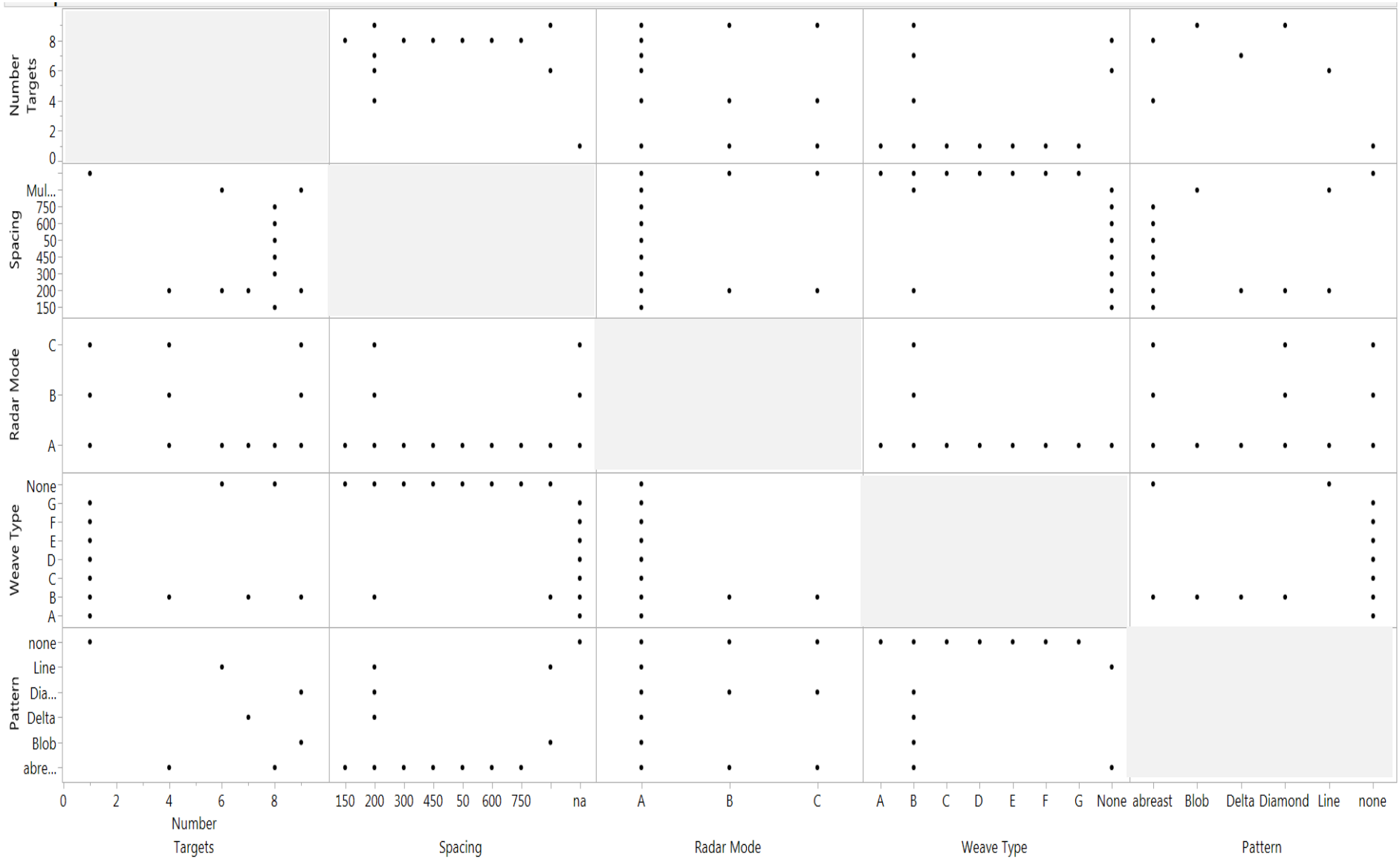
Lose ability to see interactions between factors

- E.g., Radar mode may have differing effects for different spacing and/or weave types and/or pattern types

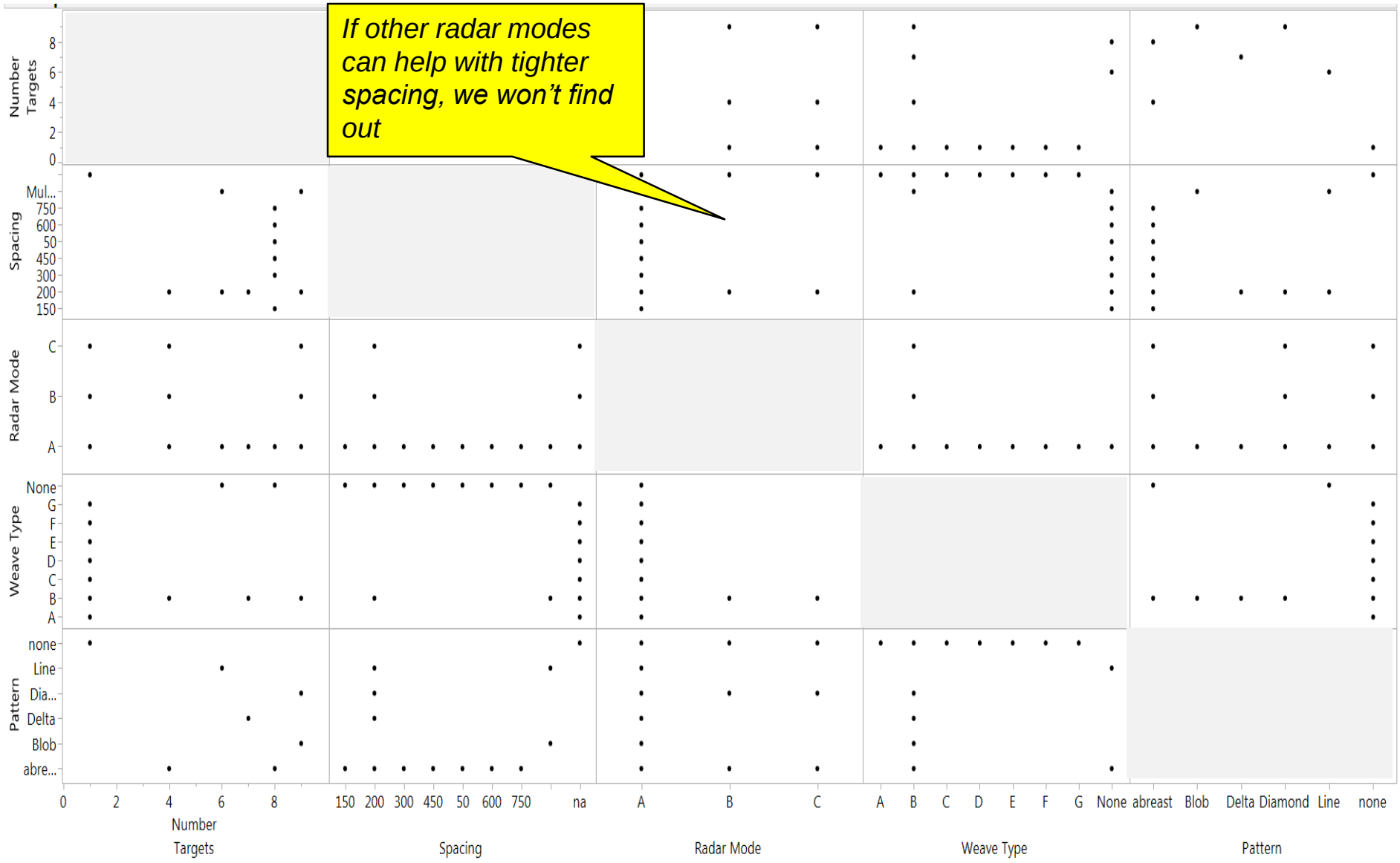
Some factors are confounded

- E.g, Radar mode is changed simultaneously with weave type (all none-weaving runs in Radar mode A)

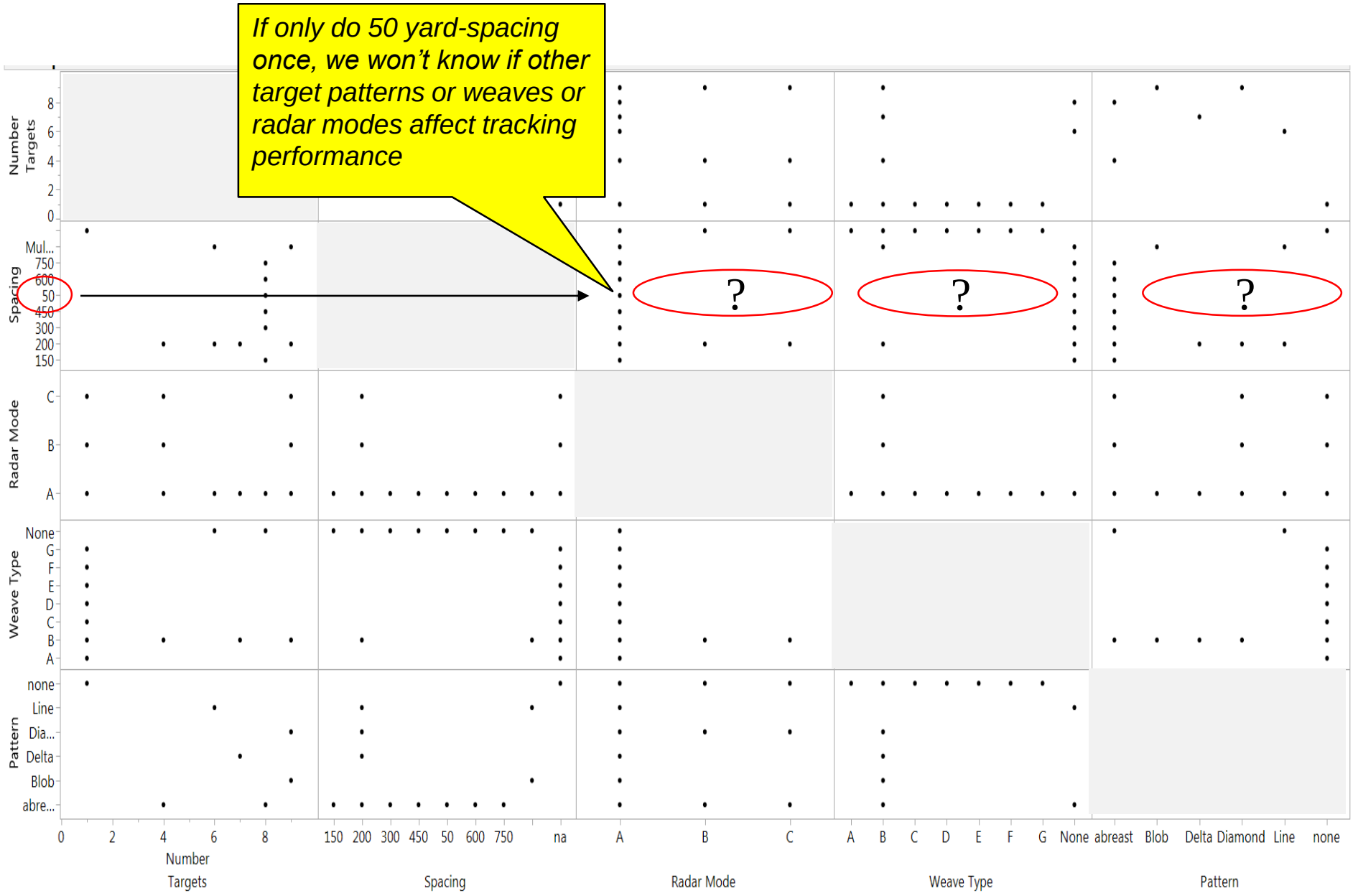
# Visual Presentation of Test Points



# Visual Presentation of Test Points



# Visual Presentation of Test Points



# Statistical Power (ability to discover performance changes across conditions) is low or inestimable

Power to Observe Significant* Performance Differences

- Near 1.0 is desired

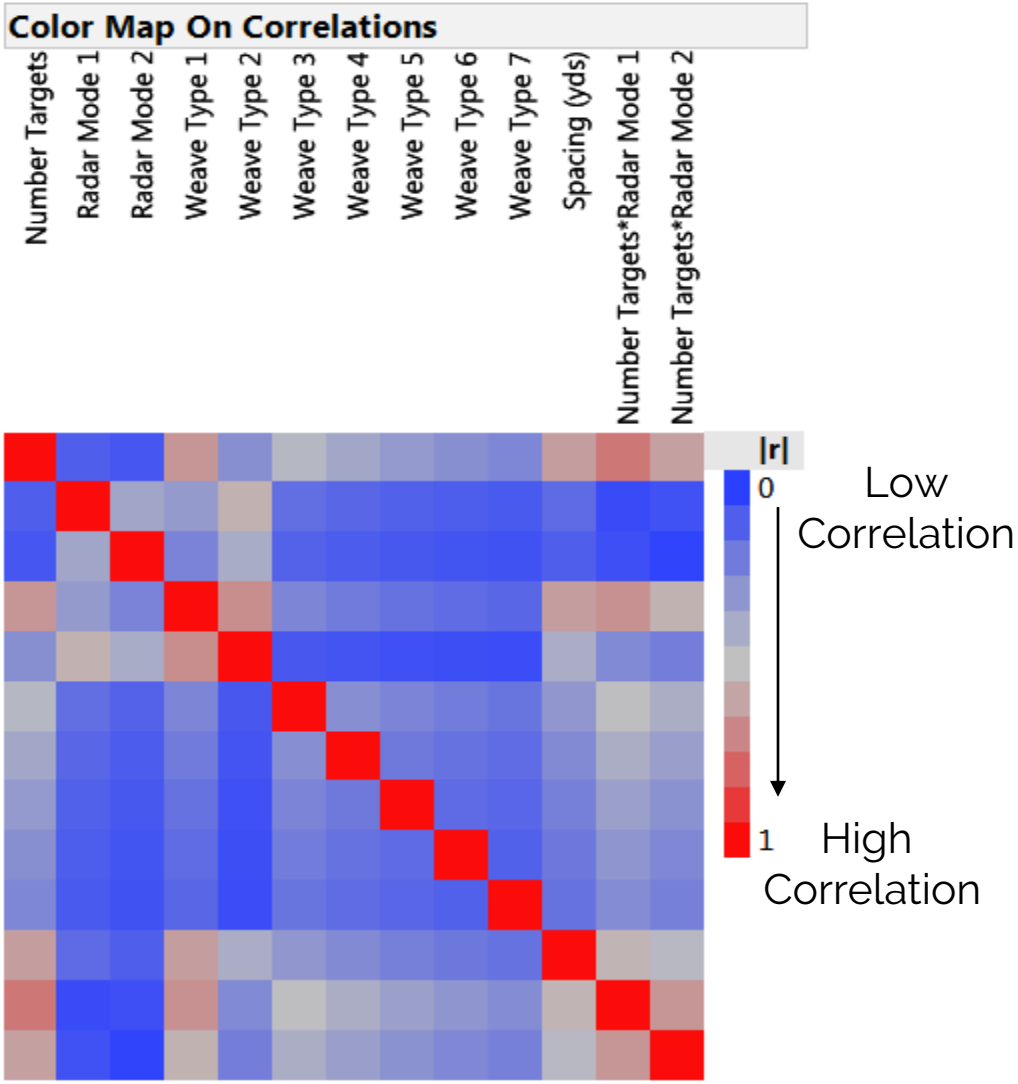
| Main Effects | |
|---|---|
| Number Targets | Not estimable |
| Radar Mode | 0.33 |
| Weave | 0.41 |
| Spacing | 0.27 |
| Pattern | 0.41 |

| Interactions | |
|---|---|
| Targets x Radar | Not estimable |
| Targets x Weave | Not estimable |
| Targets x Spacing | Not estimable |
| Targets x Pattern | Not estimable |
| Radar x Weave | Not estimable |
| Radar x Spacing | 0.27 |
| Radar x Pattern | Not estimable |
| Weave x Spacing | Not estimable |
| Weave x Pattern | Not estimable |
| Spacing x Pattern | None estimable |

*Defined as 2σ difference in response, at 80% confidence

# Correlations between factors is high..

High correlations between terms means we will not be able to ascribe performance differences to specific factors



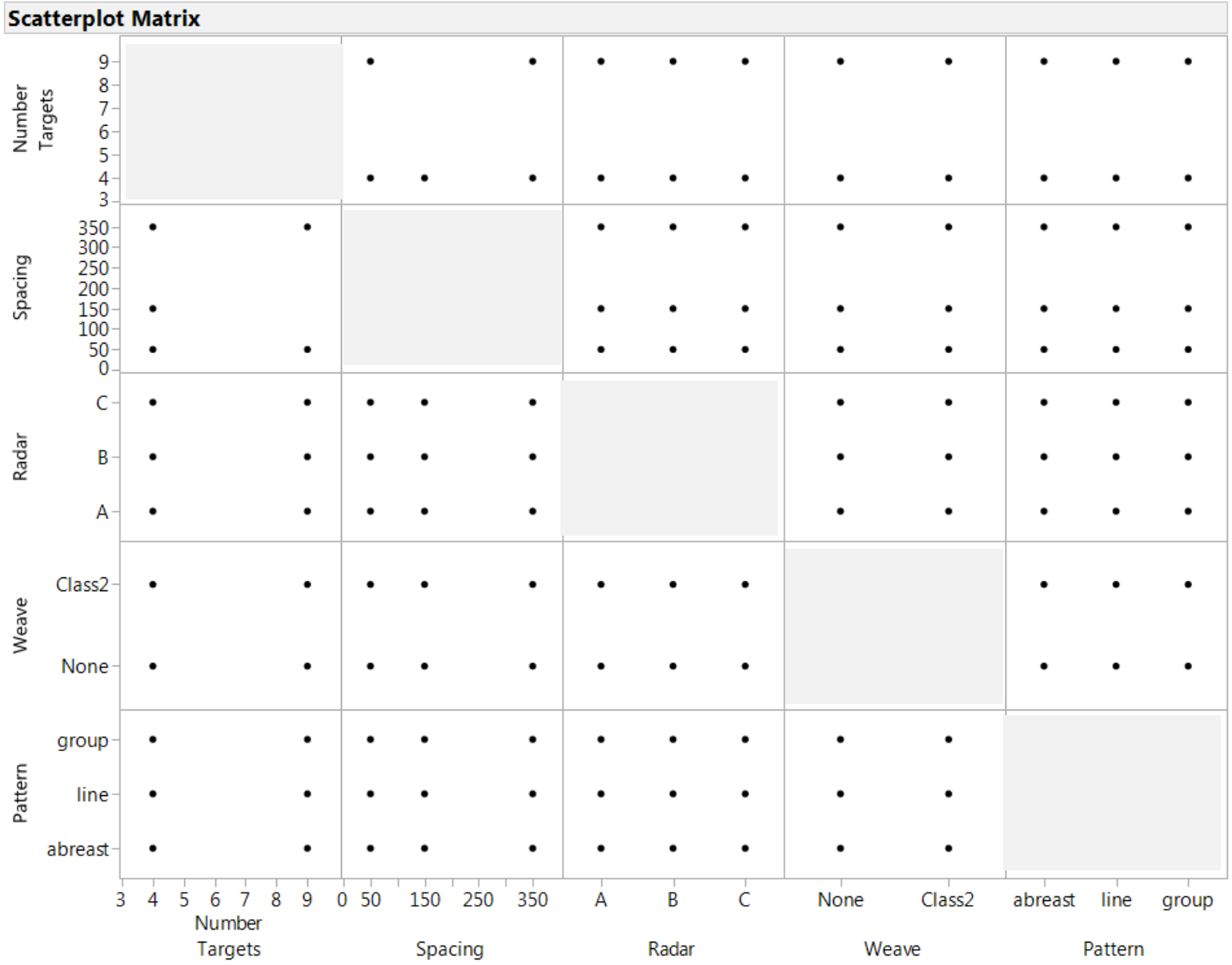
# Navy agreed to employ a DOE approach

Hard work ensued

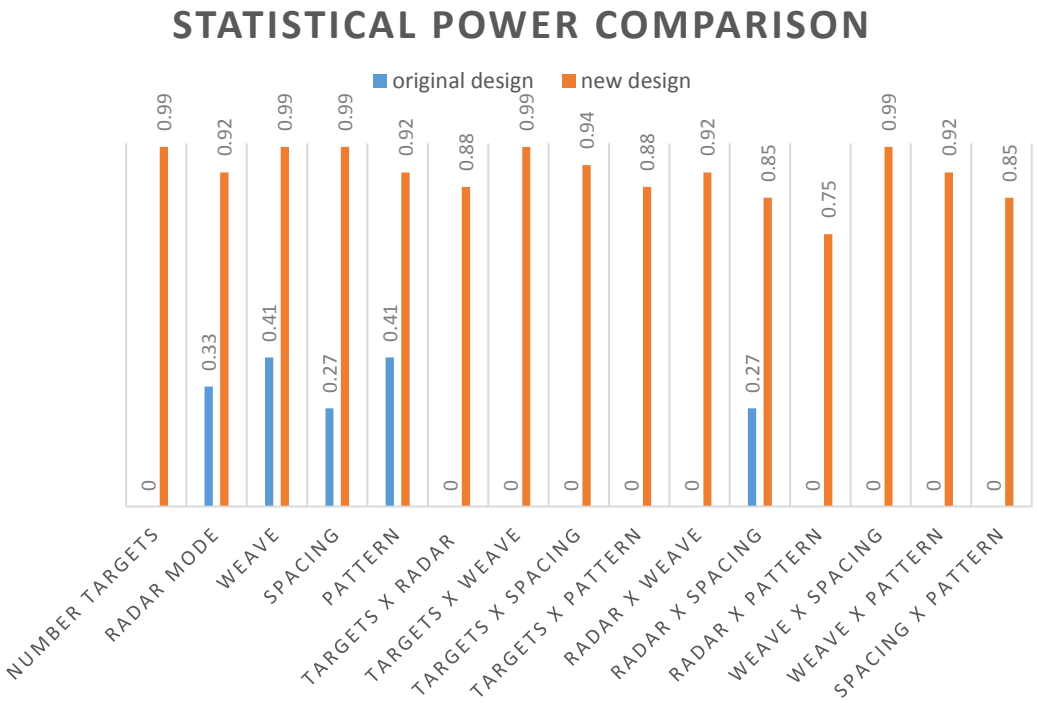
Multiple meetings between testers, engineers, SME, program manager, etc.

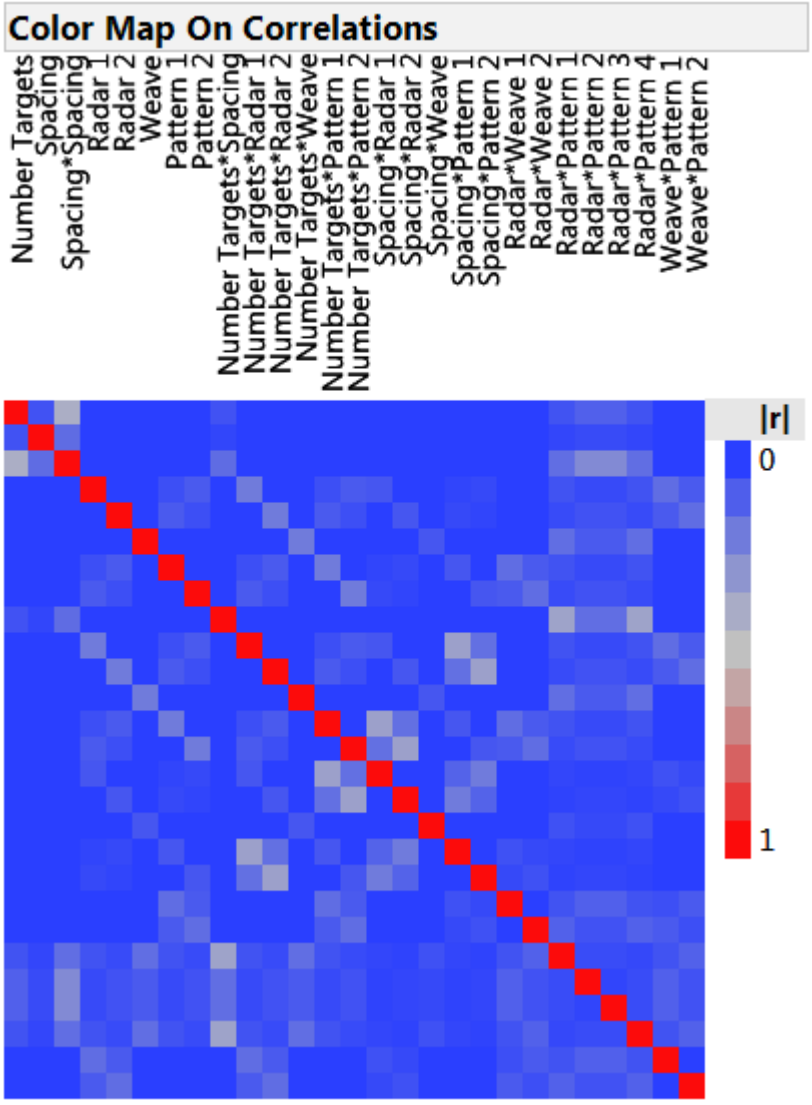
Considered multiple different designs and redesigned several times to account for execution complexities (range time)

# Final Design provides better coverage of the conditions of interest, but with the same or fewer runs



# Better Power to See Performance Differences and Low correlation among Factors




Able to estimate interactions!

Able to uncorrelate all factors!

# Applying DOE enables a better characterization of LCS's radar, for approximately the same number of runs

DOE-based test enables better development of tactics

- Can now determine which radar mode is best for different tactical situations

Enabled more informed development of the system

- Missile employment on radar tracks might require different initialization depending on radar's accuracy

Provides a better understanding of performance shortfalls and strengths across operational envelope

# Act 3 – Realism is Key

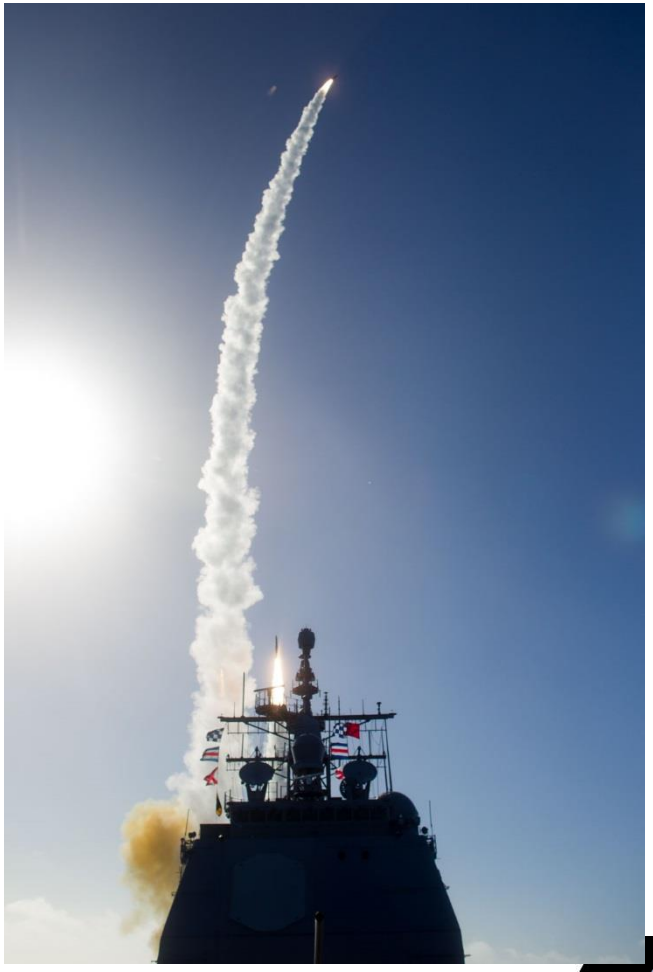
# Key Principles for OT, in 3 Acts

Testing merely to determine if KPPs or requirements have been satisfied often does not capture improvements in mission accomplishment, or achievement of intended capabilities in an operational environment

Statistical methods (e.g., Design of Experiments) are essential for designing rigorous and defensible operational tests, determining quantitatively why a test design is good, and allocating resources in the most efficient and powerful way

Operational testing has to be performance against expected realistic operational threats and in realistic conditions, and is often the ONLY means of identifying critical performance problems

# Ship Self Defense against Anti-Ship Cruise Missiles



# Operational Testing must focus on mission success for the System of Systems (i.e., the ship and its crew)

Service operational test agencies and program managers are often focused on testing their piece of the SoS.

- "Stovepipe" testing

However, operational testing usually requires testing the entire SoS, because it is often the only means to assess mission performance

- System A works
- System B works
- System A+B does not work

Individual system requirements are often inconsistent with SoS and overall mission requirements

- System A+B has to defeat threat X
- System A and B have no requirement to defeat threat X

# Probability of Raid Annihilation (PRA)

Background:

In the wake of the USS *Stark* attack (17 May 1987), the Navy took an initiative to improve ship self-defense against ASCMs.

In 1996, the Chief of Naval Operations defined the minimum self-defense requirements for all current and planned ship classes.

The requirement is known as the Probability of Raid Annihilation (PRA) requirement.

USS San Antonio (LPD 17) was the first ship class required to demonstrate the CNO's requirement.

Struck by two Iraqi Exocet Anti Ship Cruise Missiles

Other ship classes that must demonstrate PRA include the following:

- USS America (LHA 6) amphibious assault ship
- USS Zumwalt (DDG 1000) destroyer
- USS Freedom (LCS 1) and Independence (LCS 2) littoral combat ships
- USS Gerald R. Ford (CVN 78) aircraft carrier
- USS Arleigh Burke (DDG 51) Flight III guided missile destroyer

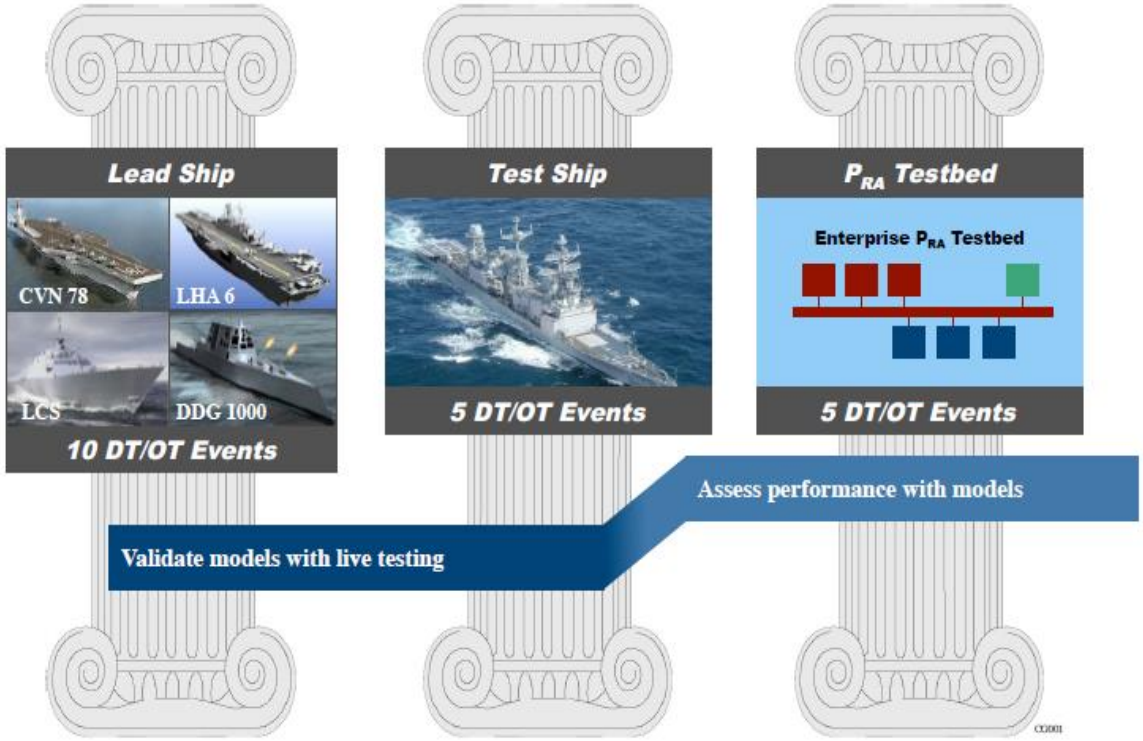
# The Navy developed a hybrid strategy of live testing and M&S, known as the Ship Self-Defense Enterprise

Range safety restrictions would not allow testers to fly ASCM surrogates close enough to manned ships to allow for self-defense engagements and they do not permit radial inbound profiles

Of five threat classes, we can test only one on manned ship and layered defense cannot be tested for any

The probabilistic nature of the $P_{RA}$ requirement, and its numeric value, would be too expensive to demonstrate via a traditional live-fire only test

# The Navy developed a hybrid strategy of live testing and M&S, known as the Ship Self-Defense Enterprise




$P_{RA}$ Assessment

# Ship Self-Defense Testing has identified Major Deficiencies that the Navy is working to fix

Many deficiencies could not have been found except with an operationally realistic presentation of threats trajectories against the self-defense test ship

- Evolved Sea Sparrow Missile (ESSM) deficiencies for specific types of ASCM threats and raids
- Rolling Airframe Missile (RAM) deficiencies against specific types of threats and raids
- Ship Self Defense System (SSDS) combat system deficiencies with respect to sensor integration and engagement scheduling
- Cooperative Engagement Capability (CEC) tracking problems against specific threats
- Radar system (e.g., SPS-48E and SPQ-9B ) detection gaps for specific threats and raid types

# Ship Self-Defense Testing has identified Major Deficiencies that the Navy is working to fix

Many deficiencies could not have been found except with an operationally realistic presentation of threats trajectories against the self-defense test ship

The $P_{RA}$ Test Bed was used by the Navy to measure LPD 17's $P_{RA}$ requirement

- Building on the success of the Test Bed, the Navy is using the Test Bed as system engineering tool to evaluate potential combat system upgrades to the LPD 17 class
- Analysis of $P_{RA}$ Test Bed results can support statistical characterizations of the ships' capabilities against ASCMs
  - Threat 1 – $P_{RA}$ = aa
  - Threat 2 – $P_{RA}$ = bb
  - Threat 5 – $P_{RA}$ = cc
  - Threat 7 – $P_{RA}$ = dd

# Parting Thoughts

Tremendous pressures to eliminate or curb operational testing, especially late in a program

- Cost argument
- Schedule argument
- Report card argument (pass/fail)
- Requirement argument (don't test "beyond" threshold)

Must resist these pressures – the goal of OT is to find and discover performance shortfalls BEFORE we go to war, so they can fixed

It is ALWAYS worth the effort and $$ to get good information

"We are not engaged in bureaucratic game play here; testing is not a game to be won. What we do is very serious. And yes, we need to highlight the performance problems that need to be fixed so that they can be fixed."

-- Dr. J. Michael Gilmore, 2016 DOT&E Annual Report