

NOTICE

This technical data was produced for the U. S. Government under Contract No. W15P7T-13-C-A802, and is subject to the Rights in Technical Data-Noncommercial Items clause at DFARS 252.227-7013 (FEB 2012)

© 2014 The MITRE Corporation. All Rights Reserved.



# Developmental Evaluation Framework & STAT-based Integrated Test Design

Dr. Suzanne Beers

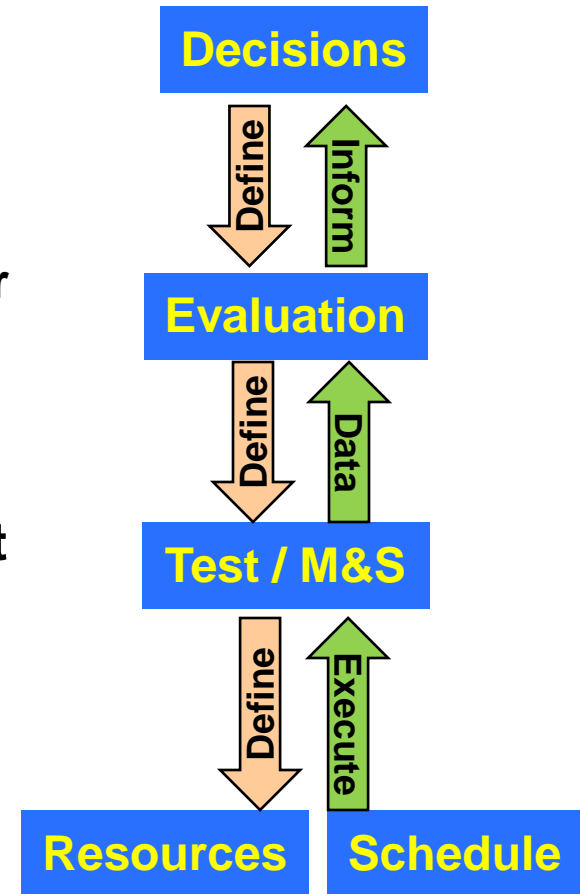
NDIA T&E Conference

22-23 July 2014

# Briefing Purpose & Overview



- “E”: Developmental Evaluation Framework (DEF) part of TEMP’s SE-V story:
  - How acquisition, technical and programmatic *decisions* will be informed by evaluation
  - How system will be *evaluated*
  - How *test and M&S events* will provide data for evaluation
  - What *resources* are required to execute test, conduct evaluation, and inform decisions
- “T”: DEF and OT Evaluation Summary Chart
  - Define data needs
  - Basis for integrated test planning
  - Statistical Test and Analysis Techniques (STAT) build optimal design

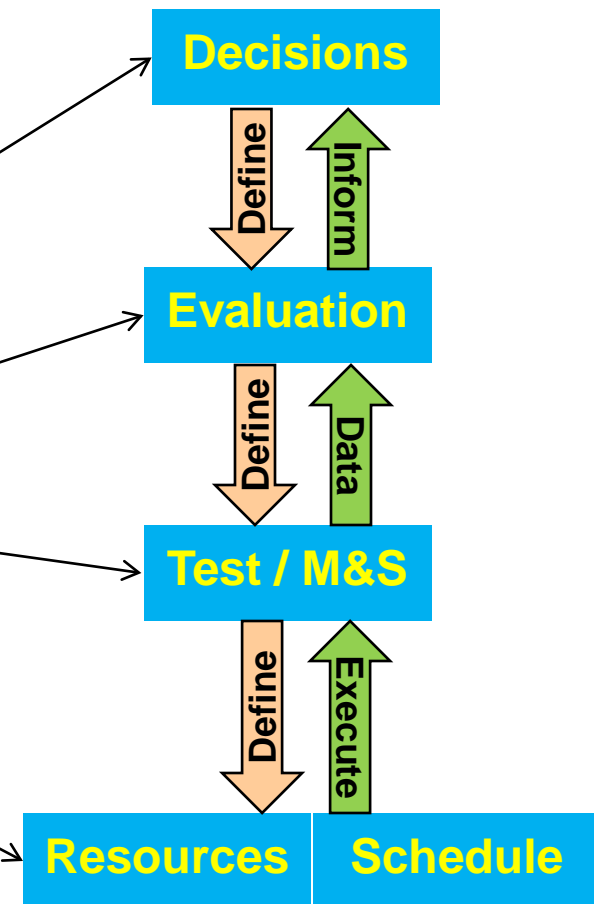


# DT&E Strategy Overview



Articulate a logical *evaluation* strategy that informs decisions

- How acquisition, programmatic, technical and operational decisions will be *informed* by evaluation
- How system will be *evaluated*
- How test and M&S events will provide *data* for evaluation
- What *resources* are required to execute test, conduct evaluation, and inform decisions



DT&E story thread: decision – evaluation – test & resources

# Developmental Evaluation Framework

(Enclosure 4, DoD Interim Instruction 5000.02)



## Test and Evaluation Master Plan (TEMP) includes a Developmental Evaluation Framework (“T&E Roadmap”)

- Knowledge gained from testing provides information for technical, programmatic, and acquisition decisions.

### DoDI 5000.02 (Interim)

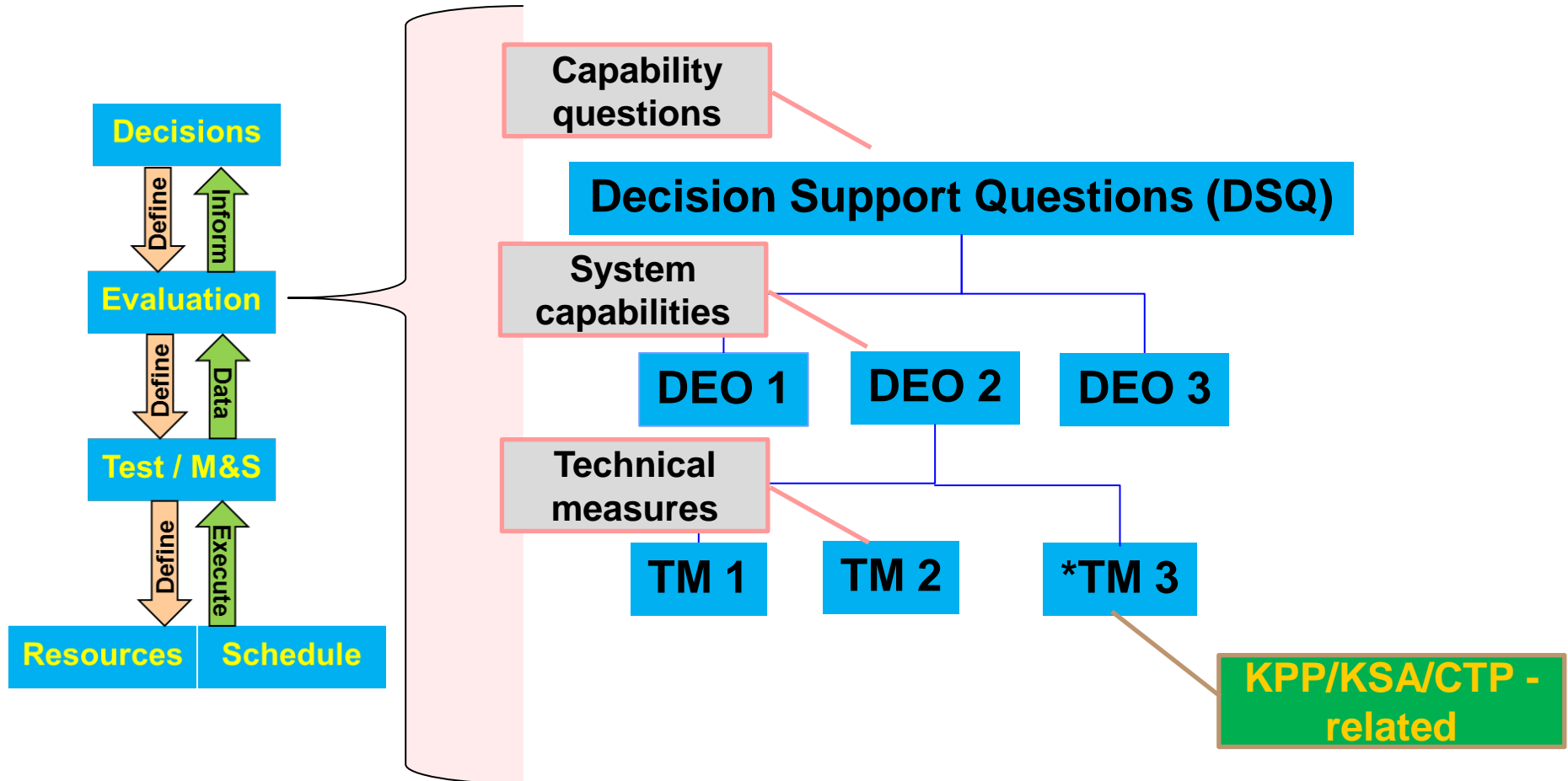
### Developmental Evaluation Framework:

- Identifies key data that contributes to assessing progress on:
  - Key Performance Parameters
  - Critical Technical Parameters
  - Key System Attributes
  - Interoperability requirements
  - Cybersecurity requirements
  - Reliability growth
  - Maintainability attributes
  - Developmental test objectives
  - Others as needed
- Show the correlation/mapping between:
  - Test events
  - Key resources
  - Decision supported

Developmental Evaluation Objectives		Decisions Supported									
System Requirements and T&E Measures											
Functional evaluation areas	Technical Reqrmts	Identify major decision points for which testing and evaluation phases, activity and events will provide decision supporting information. Cells contain description of data source to be used for evaluation information, for example: 1) Test event or phase (e.g. COT1...) 2) M&S event or scenario 3) Description of data needed to support decision 4) Other logical data source description									
System capability categories	Document Reference	Description									
Performance											
Interoperability											
Cybersecurity											
Reliability											

**Developmental Evaluation Framework  
(Defense Acquisition Guidebook)**

# Developmental Evaluation Framework (DEF)



**System Engineering decomposition:  
Evaluate system capability - Inform decisions**

# The TEMP's DT&E Strategy Story

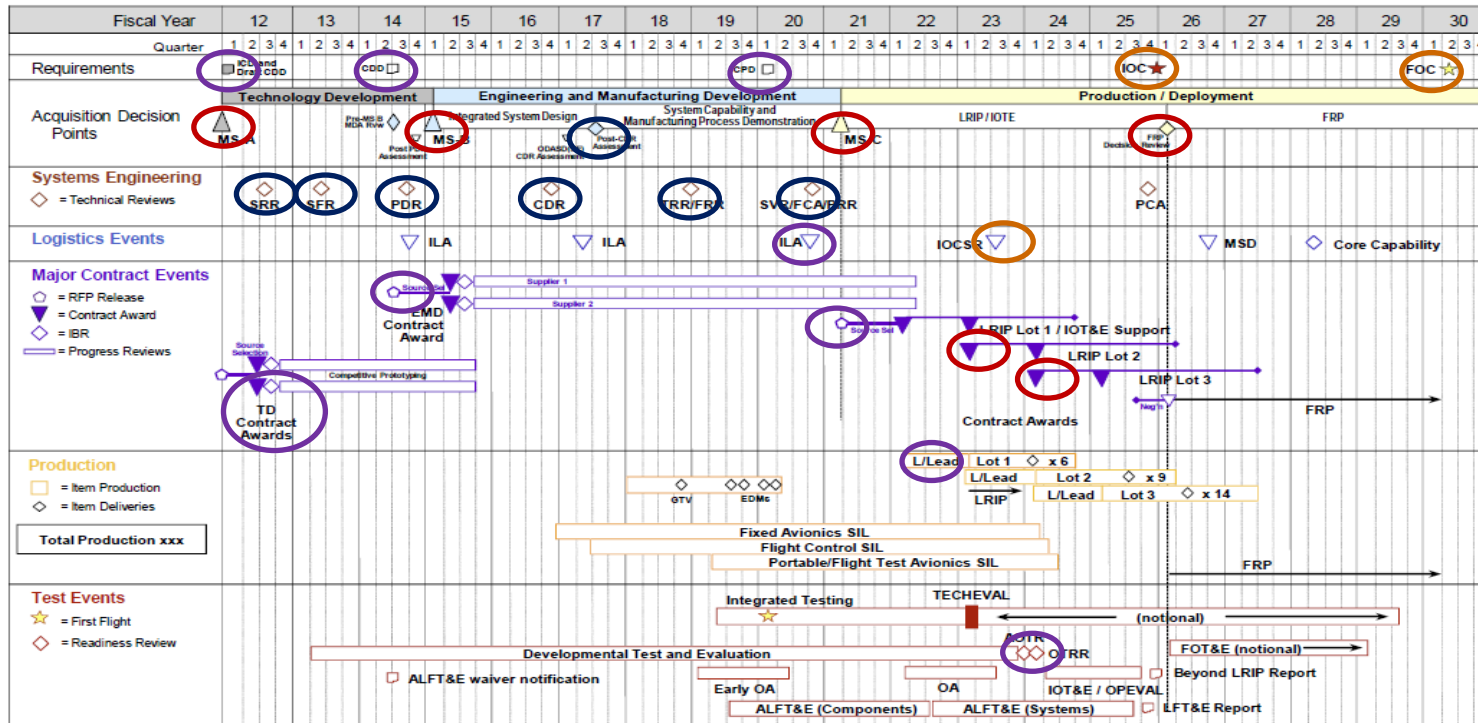


- TEMP tells the decision – evaluation – test/M&S story
  - Section 3.1 – T&E Strategy. Describe how T&E informs Acquisition Strategy decisions
    - Figure to accompany verbiage: Decision Support Key (DSK)
      - Describes decisions and T&E information needed
  - Section 3.3 – Developmental Evaluation Approach. Describe how system will be evaluated to inform decisions
    - Figure to accompany verbiage: Developmental Evaluation Framework (DEF)
      - Links decisions – evaluation – test/M&S events
  - Section 3.6 – Operational Evaluation Approach.
- Decision Support Key (DSK) & Developmental Evaluation Framework (DEF) built by Chief Dev Tester
  - DEF Core Team is subgroup of T&E WIPT including Chief Developmental Tester and select SME's



# DT&E Informed Decisions

Figure 3. Notional depiction of the Integrated Schedule for Program



- Acquisition
- Programmatic
- Technical
- Operational

AOTR: Assessment of Operational Test Readiness  
 ALFT&E: Alternative Live Fire Test & Evaluation  
 CDR: Critical Design Review  
 EDM: Engineering Development Model  
 EMD: Engineering & Manufacturing Development  
 FCA: Functional Configuration Audit  
 FOT&E: Follow-On Operational Test & Evaluation  
 FRP: Full Rate Production  
 FRR: Flight Readiness Review  
 GTV: Ground Test Vehicle  
 ILA: Integrated Logistics Analysis

IOCSR: Initial Operational Capability Supportability Review  
 IOT&E: Initial Operation Test & Evaluation  
 LFT&E: Live Fire Test & Evaluation  
 LRIP: Low-Rate Initial Production  
 MDA: Milestone Decision Authority  
 MSD: Material Support Date  
 OA: Operational Assessment  
 OASD(SE): Office of the Assistant Secretary of Defense (Systems Engineering)  
 OPEVAL: Operational Evaluation

OTRR: Operational Test Readiness Review  
 PCA: Physical Configuration Audit  
 PDR: Preliminary Design Review  
 PRR: Production Readiness Review  
 SFR: System Functional Review  
 SIL: Systems Integration Lab  
 SRR: System Requirements Review  
 SVR: System Verification Review  
 TD: Technology Development  
 TECHEVAL: Technical Evaluation  
 TRR: Test Readiness Review

**Informing decisions throughout lifecycle:  
 Same DEF concept/form; *Different* decisions and system info**



# Decision Support Key

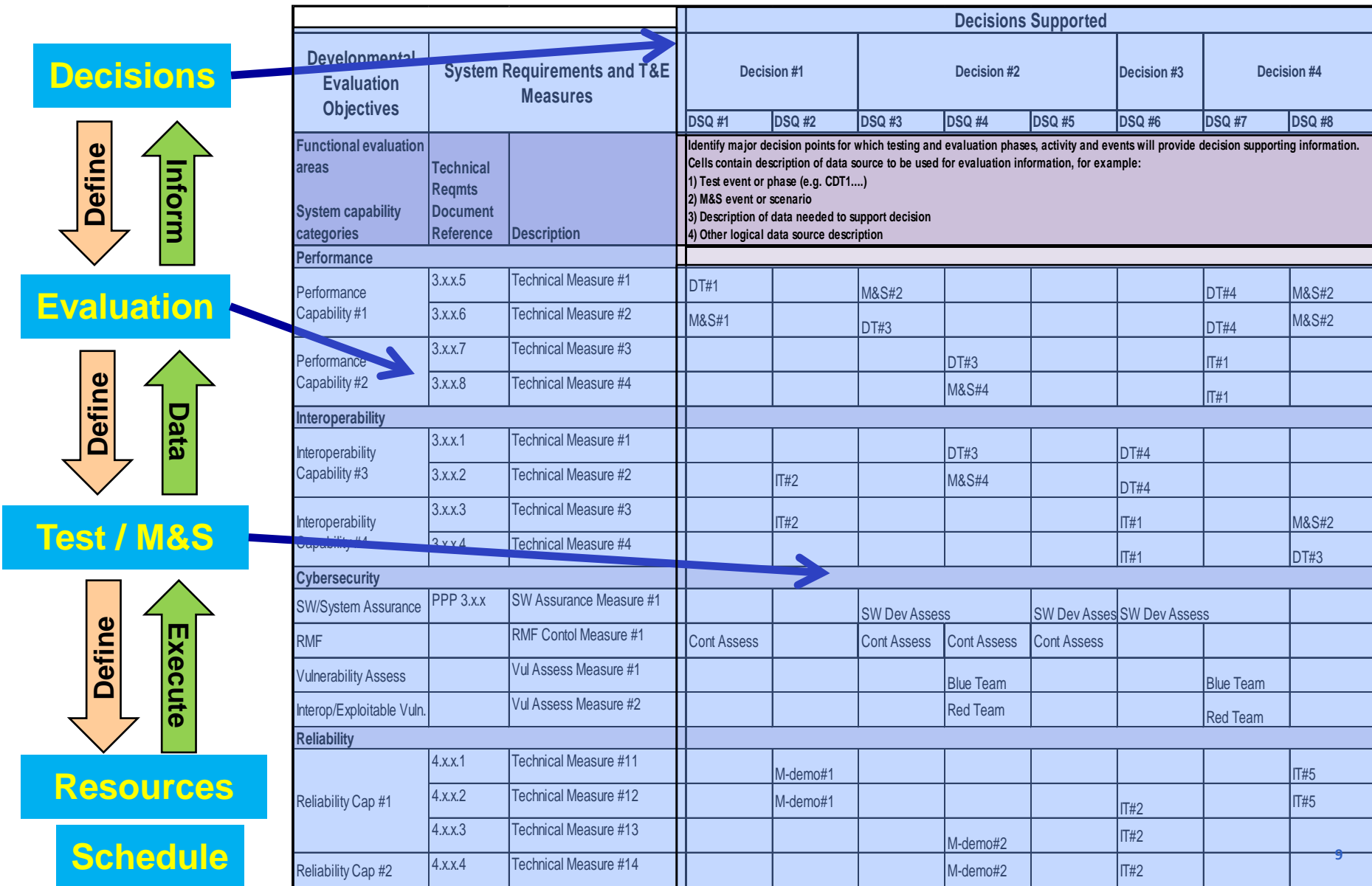
**Show how decisions will be informed by answering T&E focus questions: DT&E Decision Support Questions (DSQs) and OT&E Critical Operational Issues (COIs)**

- TEMP Section 3.1 – T&E Strategy – Describes how program’s Acquisition Strategy is informed by T&E Strategy**

<b>Decision</b>	<b>Decision Description</b>	<b>T&amp;E Info Source</b>
Decision#1 (Component maturity)	Major component technical maturity	DSQ#1, DSQ#4, DSQ#5
Decision#2 (Platform maturity)	Adequacy of host platform to accept major component integration	DSQ#2
Decision#3 (Component integration readiness)	Major component integration readiness	DSQ#1, DSQ#2, DSQ#5
Decision#4 (Initial sea trials)	Integrated system performance in ops environment	DSQ#1-5; CO#1
Decision#5 (IOC)	Initial operational capability	CO#1-4
Decision#6 (Sustainment mod)	Adequacy of sustainment modification	DSQ#4, DSQ#5, CO#1-4
Decision#7 (FOC)	Full operational capability	CO#1-4



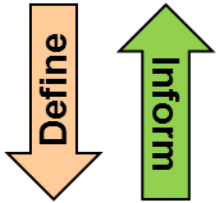
# Developmental Evaluation Framework



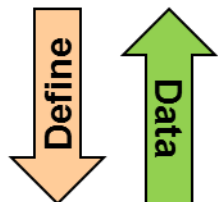


# Link Resources & Schedule

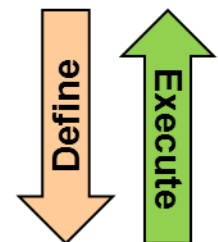
Decisions



Evaluation



Test / M&S



Resources

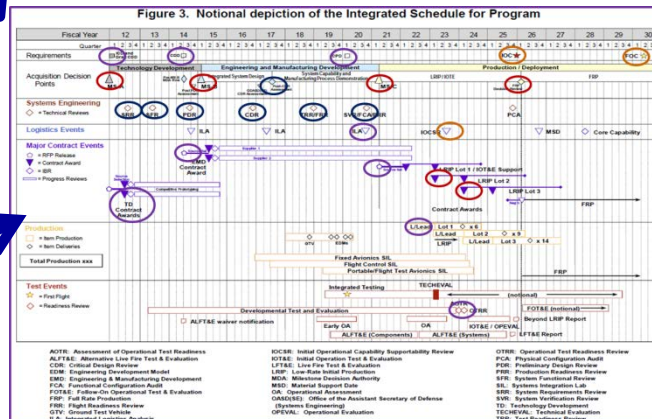
Schedule

Fiscal Year		14	15	16	17	18	19	18-20	16-20
TEST EVENT									
		DT1	DT2	DT3	IST1	IST2	IST3	Demo1-2	Exercise 1-5
RESOURCE	Units								
Resource#1: TVAC	Hours	50	80				40		
Resource#2: Acoustic Chamber	Hours		50	80			40		
Resource#3: RF Chamber	Hours		40	80			40		
Resource#4: SIL	Hours	25	25	25	80	80	80	40	200
M&S Model#1	Runs		50	132	60	100	140	30	30
M&S Model#2	Runs		50	132	60	100	140	30	30
Resource#5: Arnold AFS 6' Chamber	Hours			40			40		120

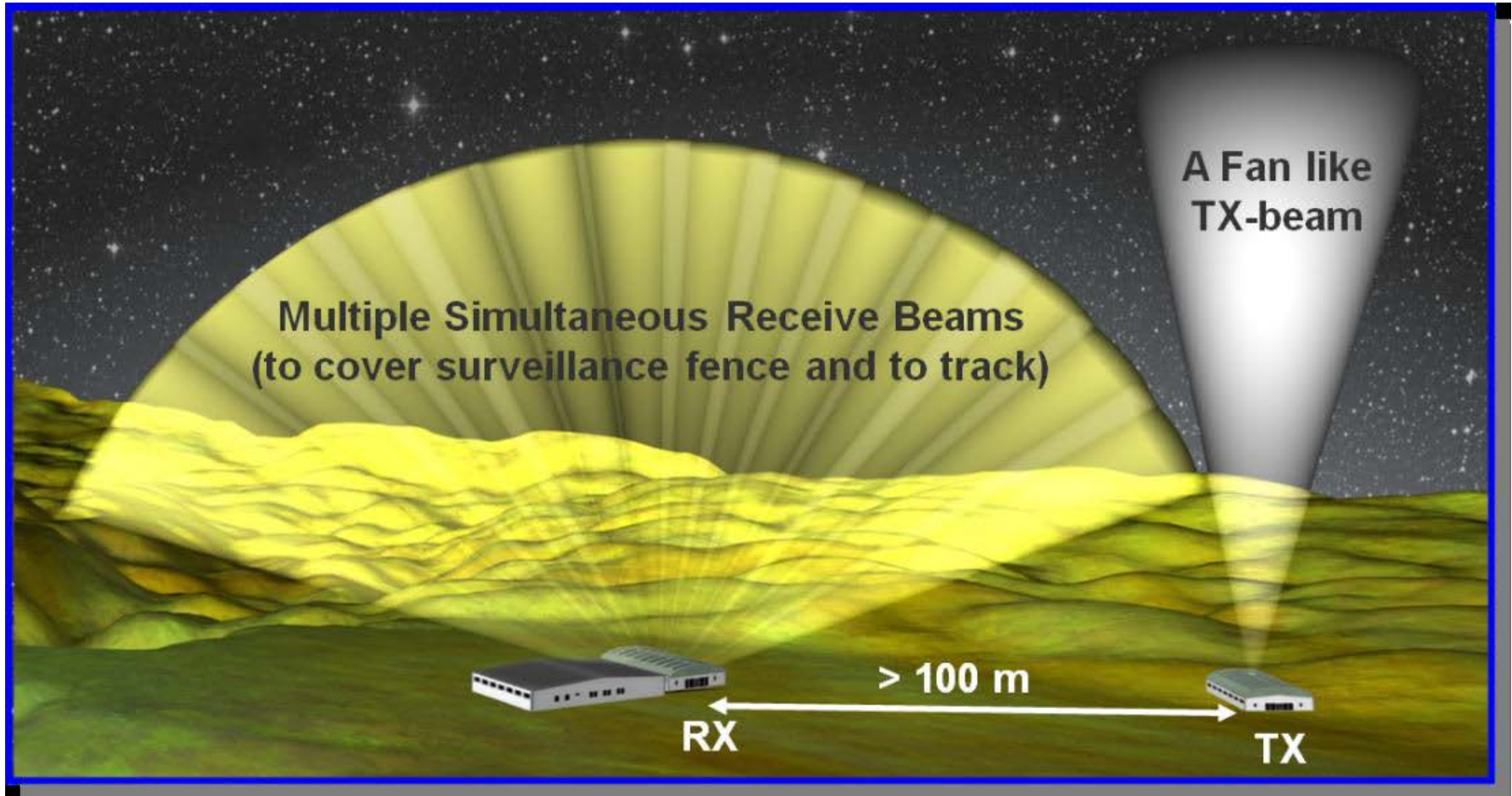
Link **key resources** and **schedule** to DEF

- Describe logical linkage of test/M&S events to necessary resources in Section IV
- Describe linkage of decisions, evaluation, test, and M&S events to schedule in programmatic schedule in Section II

Figure 3. Notional depiction of the Integrated Schedule for Program



# Example 1 - Space Fence

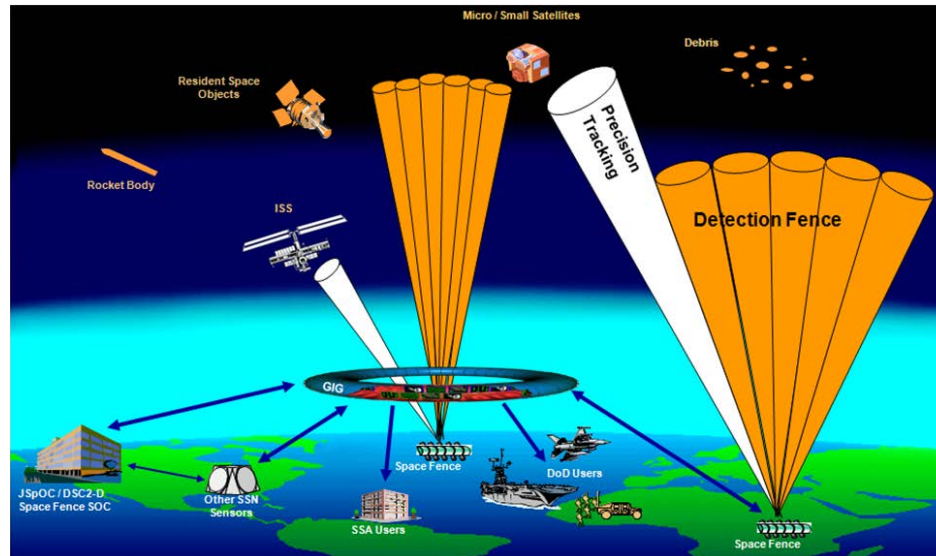


**Ground-based S-band radar to detect, track, and report on space objects to provide space situational awareness**

# Inform Capability Decisions



**Technical Mission Statement: Design and build a ground based radar system to provide LEO and MEO coverage to meet space situational awareness mission requirements**



Does the **radar** provide coverage, sensitivity, and accuracy sufficient to detect and track LEO and MEO objects?

Is the radar **data processing**, handling, and storage sufficient to characterize, correlate, track, and report space objects?

Are **command and control** and interfaces sufficient to provide tasking to the radar and surveillance information to the SSA customer

Are **environmental effects** sufficiently planned for and executed?

Are planned and executed system and **information protections** sufficient to ensure information assurance and physical security?

Are **Life Cycle Cost** factors considered and balanced with other design factors sufficient to provide a reliability, maintainable, available, and economical system?

# Space Fence DEF



## Mission & DSQs

## Dev Eval Objectives

Technical Mission Statement: Design and build a ground based radar system to provide LEO and MEO coverage to meet space situational awareness mission requirements

Critical Developmental Issues  Developmental Test Objectives	CDI #1: Does the radar provide coverage, sensitivity, and accuracy sufficient to detect and track LEO and MEO objects?	CDI #2: Is the radar data processing, handling, and storage sufficient to characterize, correlate, track, and report space objects?	CDI #3: Are command and control and interfaces sufficient to provide tasking to the radar and surveillance information to the SSA customer	CDI #4: Are environmental effects sufficiently planned for and executed?	CDI #5: Are planned and executed system and information protections sufficient to ensure information assurance and physical security?	CDI #6: Are Life Cycle Cost factors considered and balanced with other design factors sufficient to provide a reliability, maintainable, available, and economical system?
Radar coverage	*LEO uncued search coverage *LEO cued search coverage *Coverage flexibility					
Radar sensitivity	*LEO sensitivity *MEO sensitivity LEO/MEO/HEO simultaneous operations Closely spaced operations resolution					
Observation accuracy	*Angle (az/el) accuracy *Range accuracy *Time accuracy *RCS accuracy *Obs tagging integrity (includes correlate & tag)					
System calibration	Atmospheric calibration  Systematic error calibration RCS calibration Radar calibration					
Surveillance and Characterization process		Metric obs formation and dissemination RCS determination and dissemination Space object identification				

## Measures

# THEN Plan the Test -- Integrated DT/OT

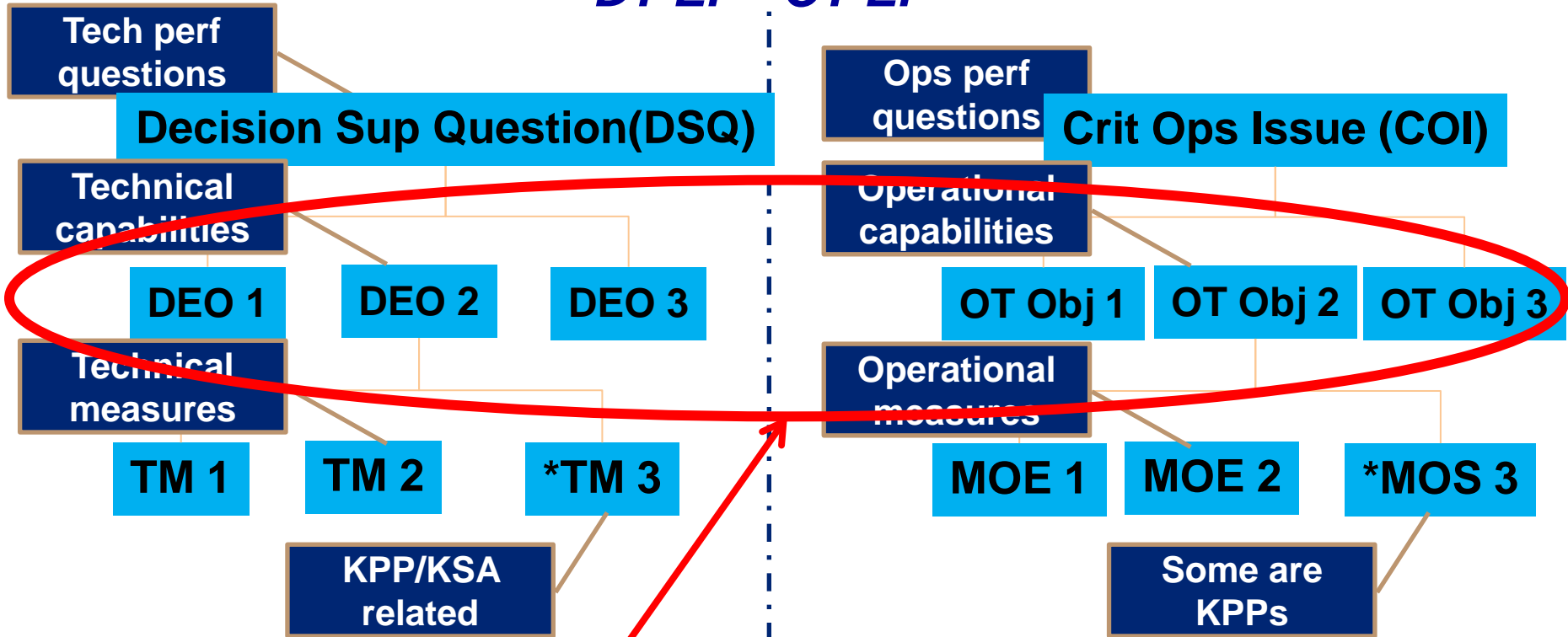


- **Integrated Test (IT) is intended to...**
  - Combine test resources (events, assets, ranges)
  - Generate data to evaluate using DT or OT evaluation framework – *independent evaluation*
  - Inform DT or OT decision-makers – *different decisions*
  
- **Integrated Test is NOT intended to be...**
  - DT&E graduation exercise
  - OT&E pre-exam
  
- **How should I design an analytically-rigorous IT?**
  - At objective level, define common input factors/conditions, output measures of interest
  - Develop input, process, output (IPO) diagram to illustrate IT design
  - Apply STAT to generate common test cases

# EFs Defines IT Data Needs



*DT EF* | *OT EF*



Potential common data for IT

# IT Design – Objective Comparison



Objective (Capability)	Measure Description	Measure Quantitative Value	Factors	Note
<b>OT&amp;E</b>				
Coverage	Number of tracks per object (KPP)	# vary by altitude (CDD Table 6-1, pg25)	Altitude, Inclination, Cued, Uncued, Time (27 hour period), Orbit shape	
	Number of objects simultaneously tracked	≥200		Similar DT measure in Capacity objective
	Detectable target size (KPP)		Size, altitude	
	Object discrimination	"best available"		Similar DT measure in Sensitivity objective
<b>DT&amp;E</b>				
Radar coverage	Range	Min = 100Km, Max ≥ 40K Km		TRD Para 3.1.4.5.1 pg 18
	Track angle	±70 degrees		
	Configurable	Operator configurable		
	LEO Uncued surveillance coverage		Altitude, Inclination, Time (27 hour period), Orbit shape	
	LEO Cued surveillance coverage		Altitude, Inclination, Time (27 hour period), Orbit shape	
	Coverage flexibility	Enhanced sensitivity over a settable region in space to meet LEO performance requirements	Target size, polarization, altitude	TRD Table 3, pg 20 Measure may be a better fit in Sensitivity DT Obj

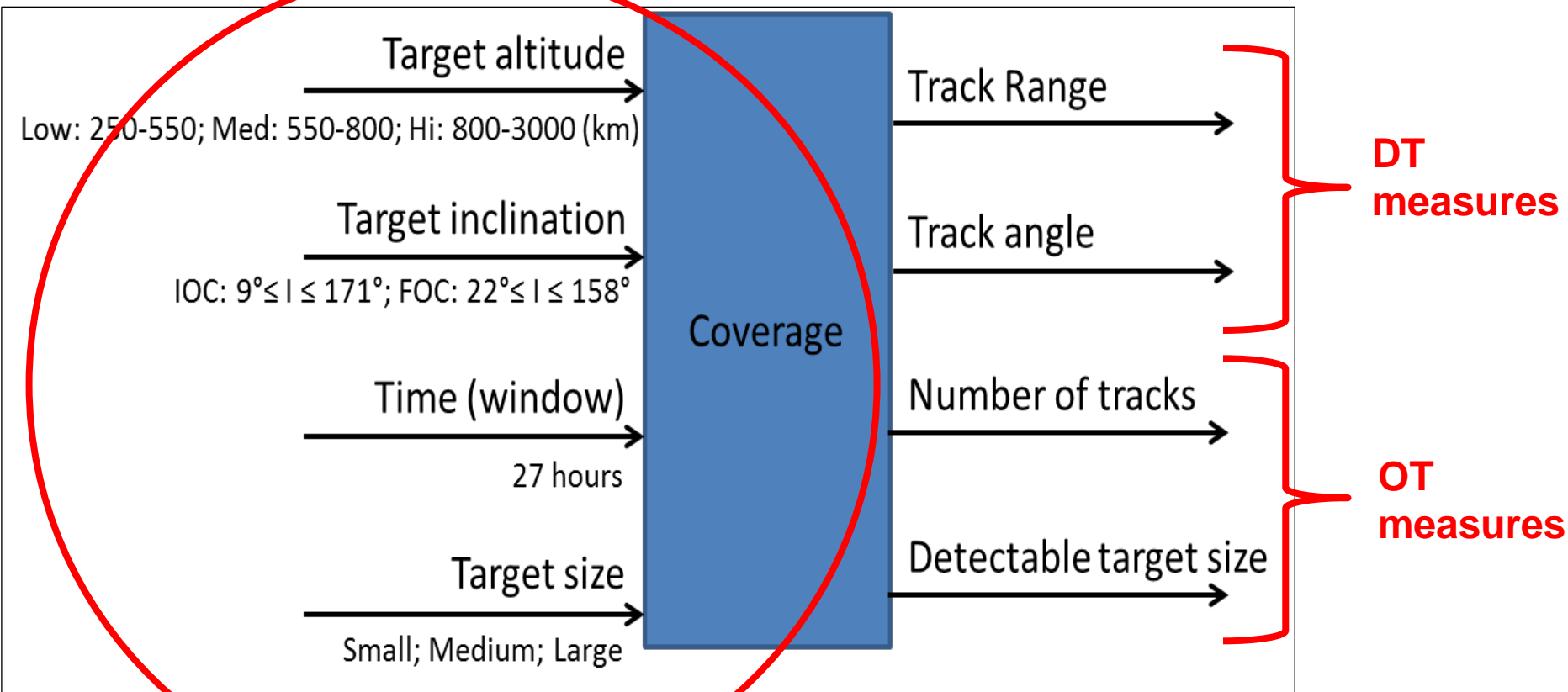
**Common DT and OT objectives (process)**

**Associated measures (output)**

**Common factors & levels (input)**

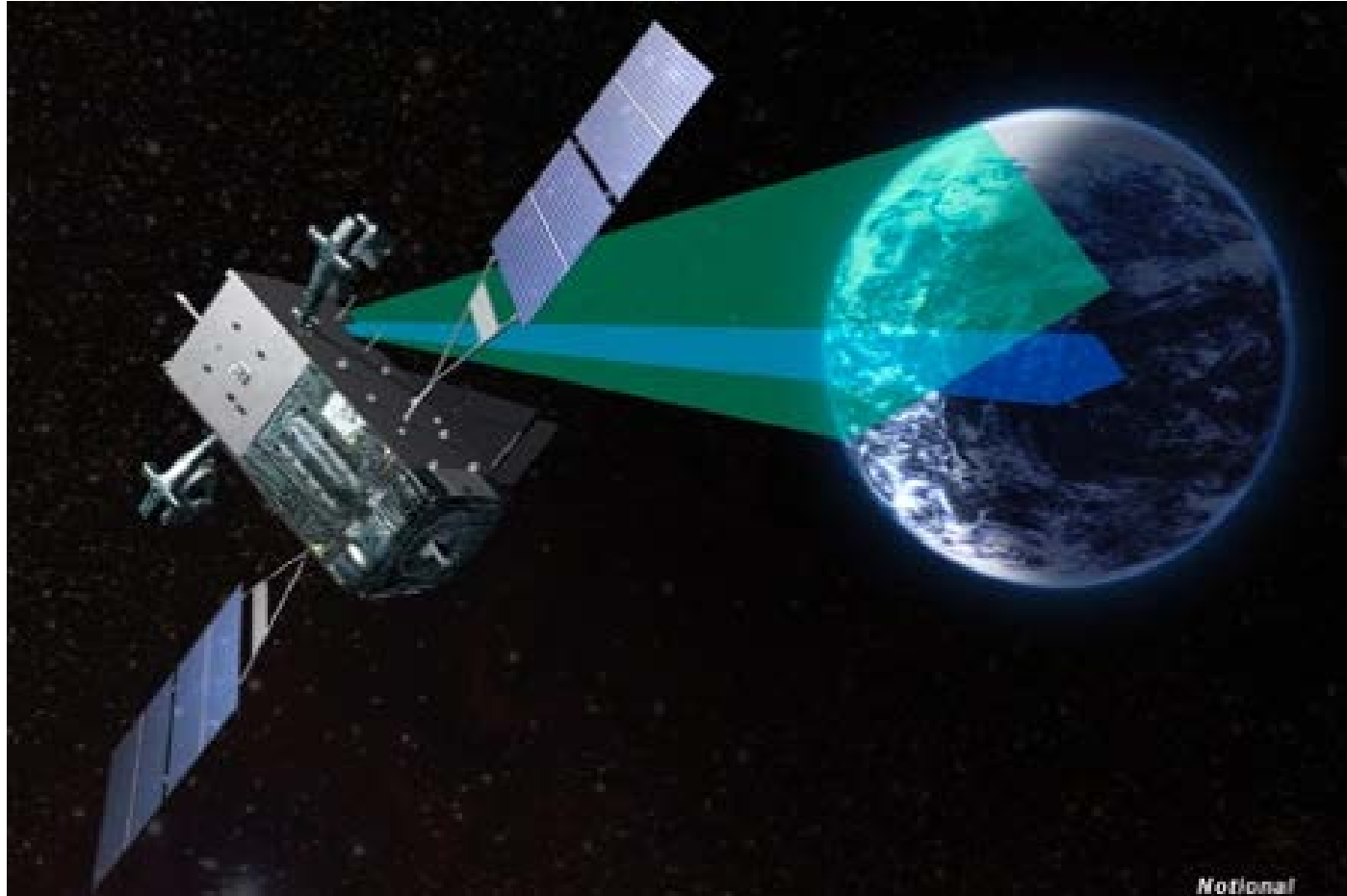


# IT Design Example – IPO Diagram



**Common DT & OT objective, factors, levels  
create test design**

# Example 2 - SBIRS

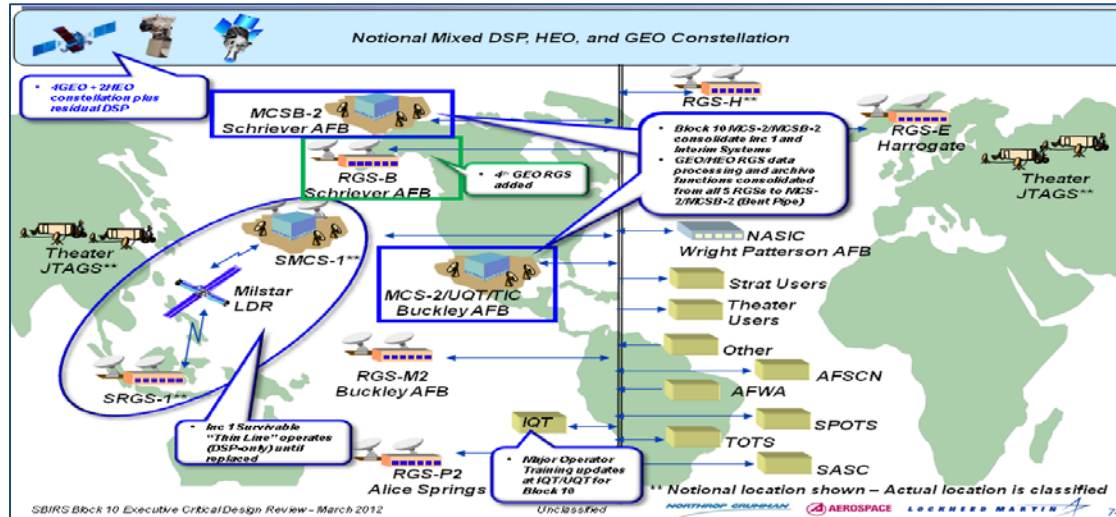


**Space-based infrared sensors and ground-based control and processing to provide missile warning, missile defense, technical intelligence, and battlespace awareness**

# Upcoming Acquisition Phases Crossed with Mission Areas



**Technical Mission Statement: Design and build satellites, infrared sensors, ground command and control, mission data processing to inform MW, MD, TI, and BA mission with IR information**



Ground Block 10.3

Ground Block 20 (Inc 2)

Space Vehicle Readiness

Mobile Ground System (S2E2)

Missile Warning

Missile Defense

Technical Intelligence

Battlespace Awareness

# SBIRS DEF



DT Objectives	Parameters	Specification Document Para #		Decision Milestones Informed				Mission Areas Supported			
		HCS	S2E2 SRD	SV Readiness	Block 10.3	Block 20	S2E2	MW	MD	TI	BA
Mission Warning / Collection Capability	Probability of Warning	3.2.1.4.1	3.2.1.3.4		X	X	X	Y	Y		
	Probability of Collection	3.2.1.5.5			X	X				Y	
Mission Coverage	Coverage	3.2.1.3	3.2.1.2		X	X	X	Y	Y	Y	
	Focused Areas	3.2.1.1, 3.2.1.3			X	X				Y	
Minimum Detection Capability	Minimum Threat	3.2.1.4	3.2.1.3.4		X	X	X	Y	Y	Y	
Mission Report Time	Report Time	3.2.1.4.2	3.2.1.1, 3.2.1.4		X	X	X				
Mission Loading Capability	Mission Simultaneity	3.2.1.1			X	X					
	Attack Magnitude	3.2.1.9	3.2.1.3.4		X	X	X				
Mission Data Availability	Data Availability	3.2.1.6, 3.3.2.3.5			X	X					
False Reporting	False Track Rate	3.2.1.8	3.2.1.5.4		X	X	X	Y	Y		
	False Report Rate	3.2.1.7	3.2.1.5.4		X	X	X	Y	Y		
Mission Data Collection	Line of Sight	3.3.2.3.1.4		X	X	X	X	Y	Y	Y	Y
	Radiometric Accuracy	3.3.2.3.1.2		X	X	X	X	Y	Y		
	Closely Space Object **Min Spatial Res**	3.3.2.3.1.3		X	X	X	X	Y	Y		
Reporting	Report Data and Contents	3.2.1.5	3.2.1.5, 3.1.4, 3.2.1.3.2		X	X	X	Y	Y		
	See to Ground Focus Area	3.2.1.5.4			X	X				Y	
Suitability	Data Fidelity	3.2.1.5.3.1			X	X				Y	
	Dependability	3.4.3	3.3.5		X	X	X	Y	Y		
	Reliability	3.4.4	3.3.4		X	X	X	Y	Y		
	Maintainability	3.4.6	3.3.6		X	X	X	Y	Y	Y	Y
	EMC	3.4.2	3.5.5		X	X	X	Y	Y	Y	Y
	Human Factors	3.7	3.5.14		X	X	X	Y	Y	Y	Y
	Supportability	3.4.8			X	X		Y	Y	Y	Y
	Ground Segment Loading	3.2.2.2	3.2.1.20		X	X	X	Y	Y	Y	
	Availability		3.3.5				X	Y	Y		
	Data Archiving	3.2.3.2.4	3.2.1.6		X	X	X	Y	Y	Y	
Data Fusion	3.2.3.2.5			X	X	X	Y	Y	Y		
Functionality	Track Telemetry and Control	3.2.3.2.6, 3.3.2.2	3.2.1.1	X	X	X	X	Y	Y	Y	
	Commanding	3.2.3.2.7, 3.3.2.2	3.2.1.9	X	X	X	X	Y	Y	Y	Y
	Autonomous Ephemeris		3.2.1.8				X	Y	Y		
	Failover/Activation Timelines		3.2.1.11				X	Y	Y		
	Anomaly Resolution	3.3.2.1.3.3		X	X	X		Y	Y	Y	Y
	Set up		3.2.1.12, 3.2.1.13, 3.2.1.14				X	Y	Y		
	Text Message Handling		3.2.1.16				X	Y	Y		
	MGS Backwards Compatibility		3.2.1.15				X	Y	Y		
	Environmental Characteristics		3.4				X	Y	Y		
	Design and Construction		3.5				X	Y	Y		
Information Assurance	System Security	3.4.9		X	X	X	X	Y	Y	Y	Y
System Survivability/Endurability	Space Segment	3.6		X				Y	Y		
	Ground Segment	3.2.2.2.2, 3.6					X	Y	Y		
Mission Interface Compliance	USNDS	3.2.3.1	3.2.2.1.2.8		X	X	X	Y	Y		
	External Comms	3.2.3.2.12	3.2.2		X	X	X	Y	Y	Y	Y
Mission Management	Surveillance	3.3.2.1.1		X	X	X		Y	Y	Y	Y
	Mission Data Processing Priorities	3.3.2.1.2			X	X		Y	Y	Y	Y
	Managing System Resources	3.3.2			X	X		Y	Y	Y	Y
	Collaborative OPIR Functionality	3.3.2			X	X		Y	Y	Y	Y
Logistics	ILS Support	3.5			X	X	X	Y	Y	Y	Y
	Personnel & Training	3.5.5, 3.5.6, 3.5.7			X	X	X	Y	Y	Y	Y

**Decision Support Questions**

**Mission Areas Supported**

**Developmental Evaluation Objectives**

**Measure for DSQ/DEO eval**

**Technical Measures**



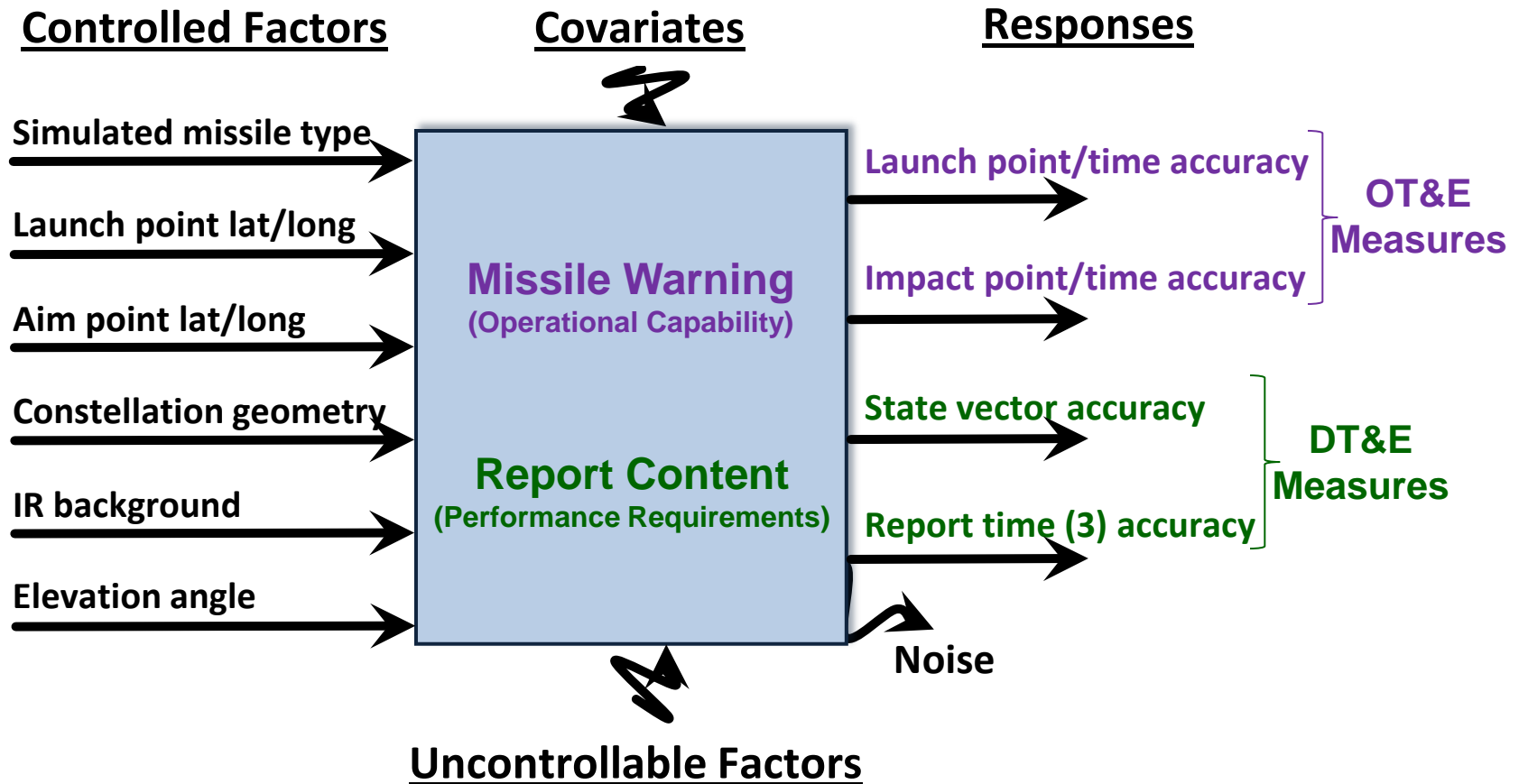
# SBIRS COI 1 Factor Space

COI 1 Factors (Version 2.1 - 30 April 2014)

Factor Name	OLGASim Factor	NG Rank	LM Rank	Factor Type	Factor Subtype	Levels	Factor Management	Level Descriptors	Factor Type
Missile Type <sup>1</sup>	x	1	1	Categorical	Nominal	5	Vary	ICBM, SLBM, IRBM, MRBM, SRBM Small, Medium, Large None, A, B, C (Demo A, B, C) Min, Max Short, Intermittent, Long Min, Max	Threat
Attack Magnitude		1		Categorical	Nominal	3	Vary		
Threat				Categorical	Nominal	4	Vary		
Source Missile Intensity <sup>2</sup>	x			Numeric	Continuous	2	?		
Burn Duration <sup>2</sup>	x			Categorical	Nominal	3	?		
Missile Acceleration <sup>2</sup>	x			Numeric	Continuous	2	?		
Launch Origin Latitude	x	4	2	Numeric	Continuous	129	Vary	Min, Max Min, Max Min, Max Min, Max -180, +180	Trajectory
Launch Origin Longitude	x			Numeric	Continuous	129	Vary		
Aim Point Latitude	x	4		Numeric	Continuous	1108	Vary		
Aim Point Longitude	x			Numeric	Continuous	1108	Vary		
True Launch Azimuth	x		2	Numeric	Continuous	n/a	Log		
Local Zenith Angle	x	4		Numeric	Continuous	n/a	Log	0, +180 Eclipsed, Non-Eclipsed Day, Night Cirrus, Cirrostratus, Cirrocumulus, None	Environmental
Solar Season		3		Categorical	Nominal	2	Vary		
Time of Day		1	3	Categorical	Nominal	2	Vary		
Cloud Cover			3	Categorical	Nominal	n/a	Log		
Atmospheric Transmission			3						
HEO Coverage				Categorical	Ordinal	n/a	Log	0, 1, 2	Constellation
Sensor Type		2	1	Categorical	Nominal	n/a	Log	GEO Scanner, GEO Starer, HEO, DSP, Combination, Other Data	
Sensor with Sufficient Angle				Categorical	Nominal	n/a	Log	0 thru N (N = Classified)	
Launch Notice		4		Categorical	Nominal	3	Vary	None, Short, Advanced	Operational
Operator Experience				Categorical	Nominal	n/a	Log	Begginer, Intermediate, Advanced	
Number of Strategic Events				Numeric	Continuous	n/a	Log	Classified	
Concurrent Strategic Events				Numeric	Continuous	n/a	Log	Classified	
Release Mode			4	Categorical	Nominal	n/a	Log	Operator, Auto Release	
Communication Link				Categorical	Nominal	n/a	Log	Given	



# Test Design Illustration Example

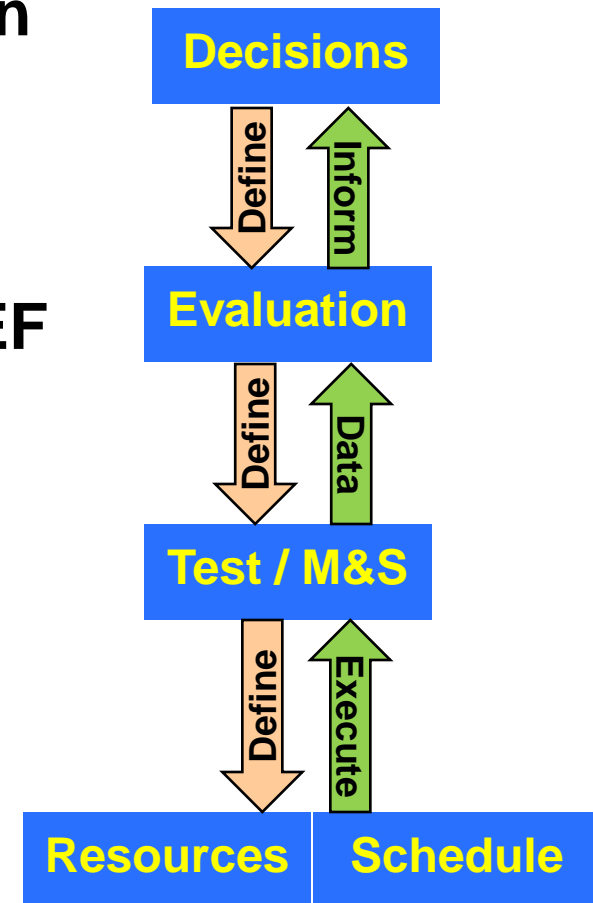


Ref: Beers, S. M., Brown, C. D., Cortes, L. A. (2014). The "E" before the efficient & rigorous "T": From Developmental Evaluation Framework to Scientific Test and Analysis Techniques implementation. *ITEA Journal* 2014; 35: 45-50.



# Summary & Way Ahead

- **“E”**: DEF focuses system evaluation (in mission context) to inform decisions
  - DSQ (decision) → DEO (capability) → TM (measure)
- **“T”**: Test plans generate data to feed EF
  - Use STAT / DOE to design rigorous and complete test campaigns





# ***SBIRS' STAT-based Test Design***

- Identify and rank candidate Developmental Evaluation Framework (DEF) and Operational Test Evaluation Framework (OT EF) critical performance parameters (responses) for test and analysis via design of experiments
- Complete defining their respective factor spaces
  - Threat factors
  - Trajectory factors
  - Environmental factors
  - Constellation factors
  - Operational factors
- Screen DEF and OTEF responses for common influential factors
- Identify test constraints and limitations
- Review E5 data analysis
- Plan for building a representative number of experimental designs taking advantage of historical data analysis and Block 20 test plans
- Start planning the strategy for Integrated Testing





# Background - STAT Strategy

