



Approved for Public Release; Distribution Unlimited. Case Number 15-3362  
©2015 The MITRE Corporation. ALL RIGHTS RESERVED.



# **The Developmental Evaluation Framework (DEF) and Design of Experiments (DOE) - Helping Decision Makers make Better Informed Decisions**

**3 March 2016**

Luis A. Cortes, Ph.D., P.E.  
The MITRE Corporation  
lcortes@mitre.org

Suzanne M. Beers, Ph.D.  
DASD(DT&E) DEF Technical Lead  
The MITRE Corporation  
sbeers@mitre.org



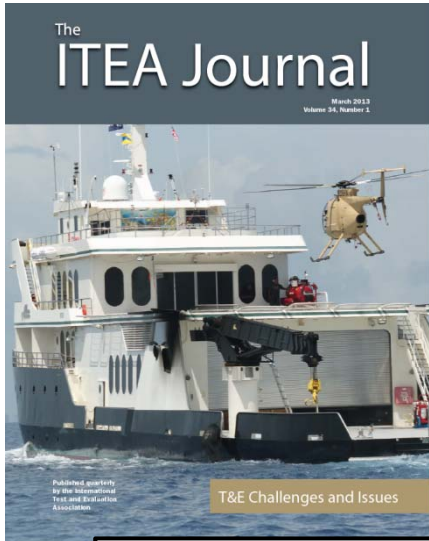
# Discussion Topics



- Context - DT&E's Purpose is to Inform Decisions
- First the **"E"**: Developmental Evaluation Framework (DEF)
- Then the **"T"**: Design and Analysis of Experiments (DOE/STAT)



# Improving Acquisition Decision-Making



## ■ Honorable Frank Kendall (USD/AT&L) on the importance & benefits of DT&E-informed acquisition decision-making

- *As Defense Acquisition Executive, I rely heavily on the implications of developmental test results for investment decisions, particularly for entry into low rate production. Developmental testing is a core activity in our acquisition programs.*
- *The purpose of developmental testing is simple: to provide data to program leadership so that good decisions can be made as early as possible.*
- *Formal design of experiments techniques are being used widely now to ensure that tests are structured to extract meaningful information as efficiently as possible and I applaud this development.*
- *The bottom line: Developmental testers are critical professionals who make a major contribution to DoD's programs. Working with program and engineering leadership as key members of the management team, developmental testers provide the information that makes program success possible and much more probable.*

Guest Editorial  
ITEA Journal 2012, Vol. 34, No. 1

### Perspectives on Developmental Test and Evaluation

Frank Kendall  
Under Secretary of Defense for Acquisition, Technology, and Logistics  
The Pentagon, Washington, D.C.


*During my first tour in the Pentagon in Acquisition, Technology, and Logistics (AT&L) from 1986 to 1994, I was responsible initially for strategic defense systems and then for tactical warfare programs. During this time, I had the opportunity to work with a Developmental Test and Evaluation (DT&E) organization that was very professional and led by an outstanding civil servant. His deep knowledge along the way, in processes and personalities changed in the Office of the Secretary of Defense (OSD), the DT&E organization expanded and all but disappeared. For the last few years, under the auspices of the Program System Acquisition Reform Act, we have been strengthening the DT&E organization within OSD. Ed Chen, who retired from public service recently, has rebuilt the DT&E organization in the past that is now performing a role much closer to the one I remember from the 80s and 90s. As Defense Acquisition Executive, I rely heavily on the DT&E office and staff for sound advice on the adequacy of the test programs being proposed for major programs and on the implications of developmental test results for investment decisions, particularly for entry into low rate production. Developmental testing is a core activity in our acquisition programs, however, not just an OSD oversight function. In this article I discuss the role DT&E plays in our programs, some important principles I believe should be applied to developmental testing, and some common problems I have encountered that relate to the effectiveness of DT&E.*

#### Role of developmental testing

The purpose of developmental testing is simple: to provide data to program leadership so that good decisions can be made as early as possible. I have a sign outside my office displaying a quote from W. Edwards Deming: "To God we trust, all others must bring data." It is our developmental testers who "bring the data" needed to make sound decisions during product development. Programs are organized in various ways, but whenever the specific organizational model, writing is the source of the critical information that provides feedback to program managers, chief engineers, lead system engineers, integrated product teams, and military users on whether their design meets requirements or not. The spectrum of testing types and sensors that is captured in compliance

maintains for system specifications runs the gamut of laboratory testing and field testing. All of these sources of information can be valuable, but integrating them into a test program and an overall program plan and schedule that meet the needs of developmental testers' customers require a high degree of professionalism and a deep understanding of how test results can influence design and program decisions. In my experience, a well-structured test plan makes all the difference in determining whether a program is efficiently executed or not. There are two layers of DT&E: organizational role and relationships both are important in determining DT&E's contributions to program success.

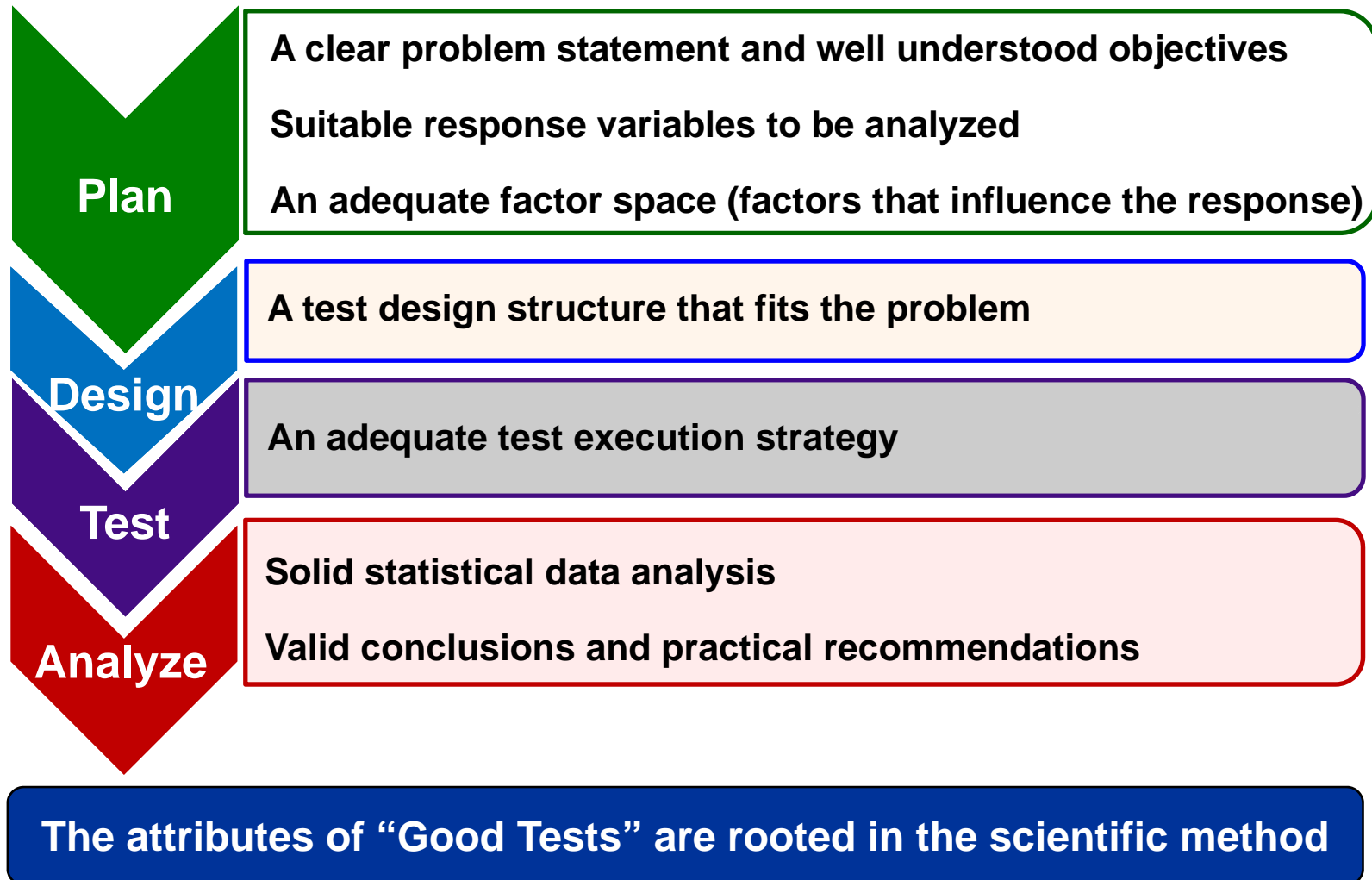
The first layer of DT&E organization exists within the program office. I have seen several organizational



6 ITEA Journal

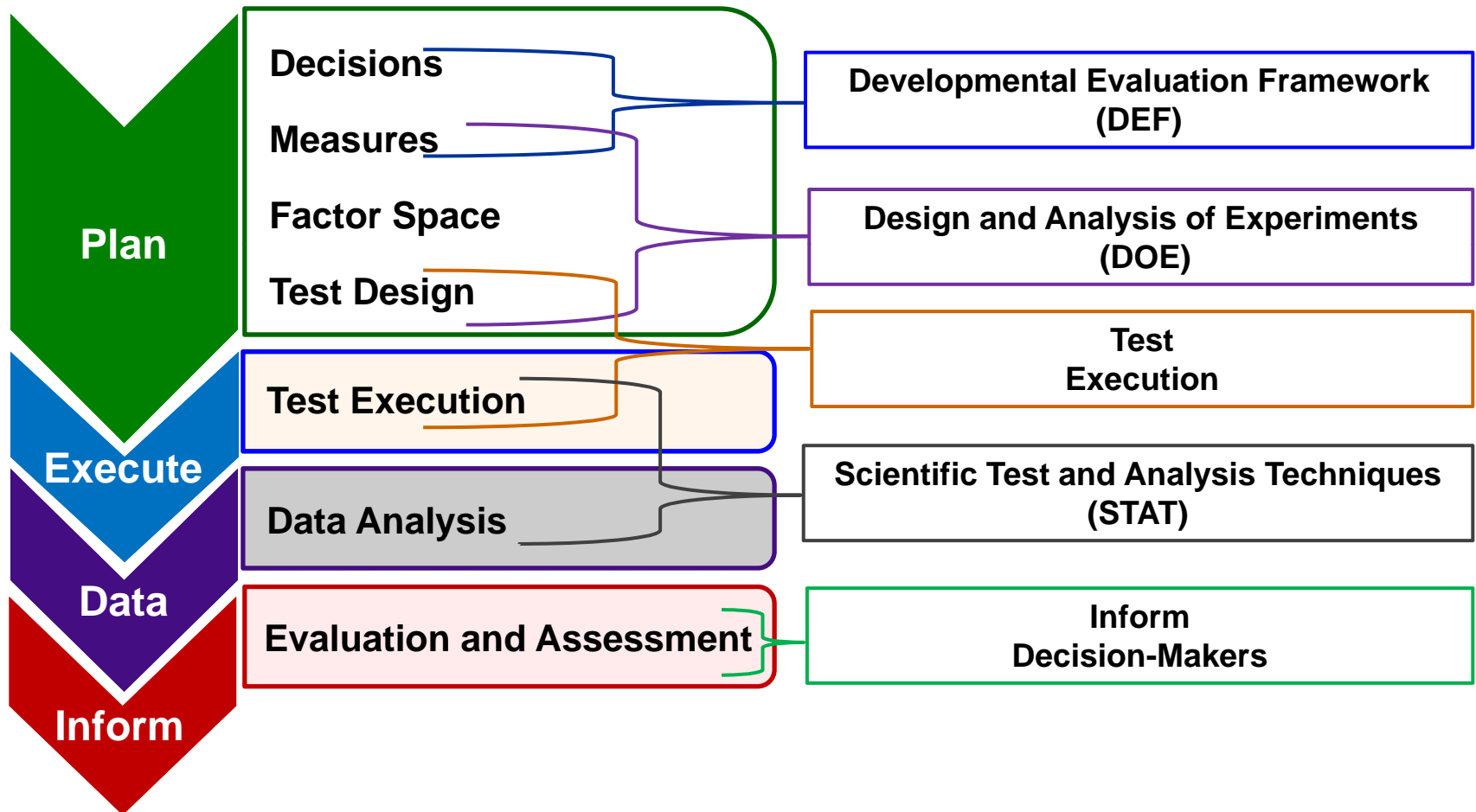


# Attributes of a “Good Test”





# DEF & DOE Contributions to a “Good Test”

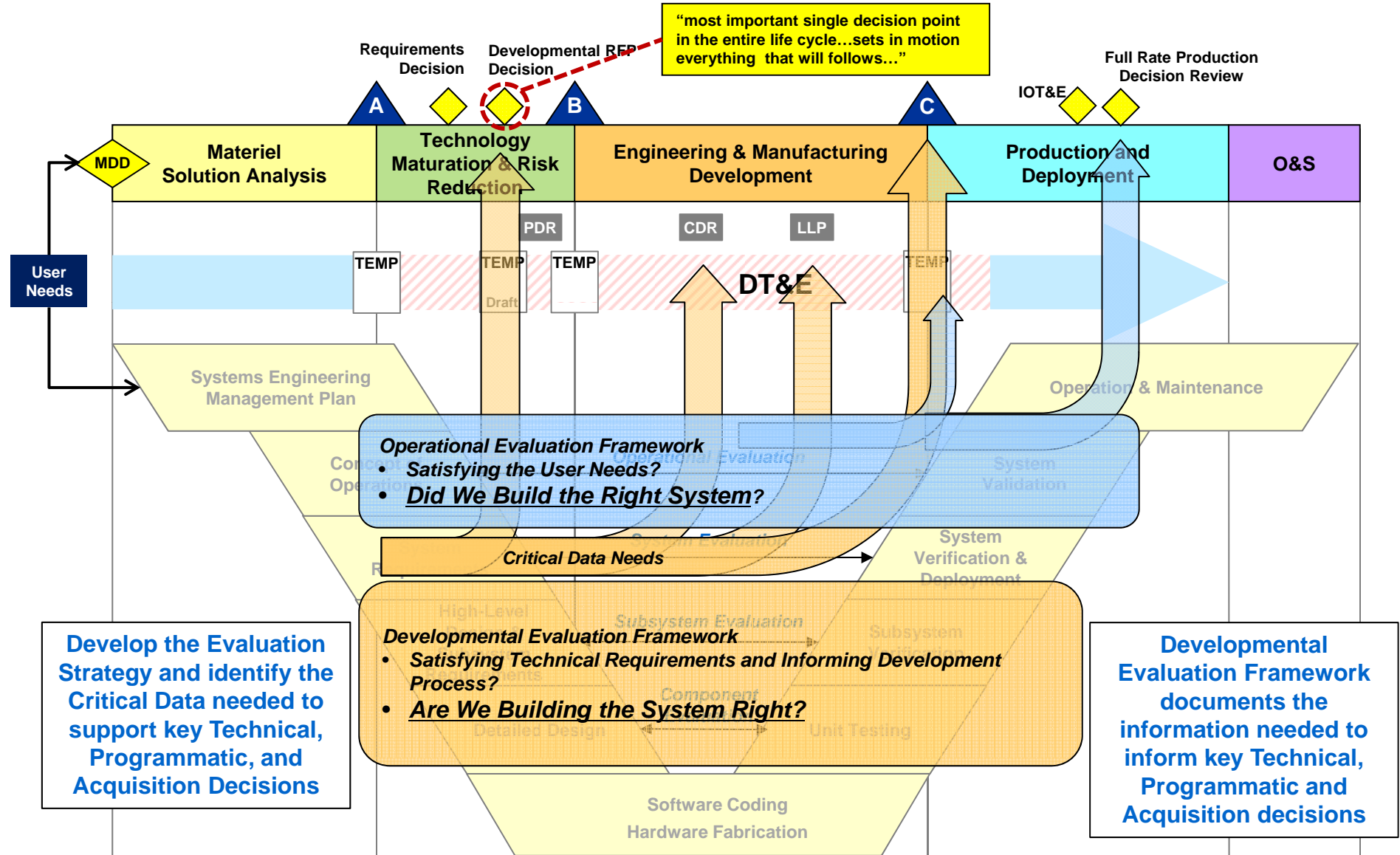


*“The purpose of developmental testing is simple: to provide data to program leadership so that good decisions can be made as early as possible.”*



# SE, DT&E, and DoDI 5000.02

## Plan the Evaluation & Inform the Decisions





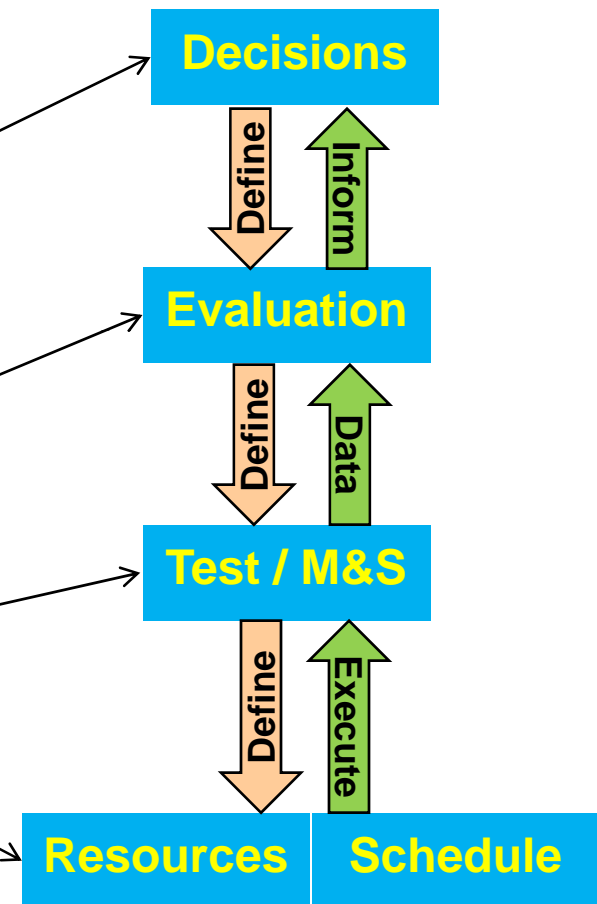


# DEF: Articulating the DT&E Strategy



Articulate a logical *evaluation* strategy that informs decisions

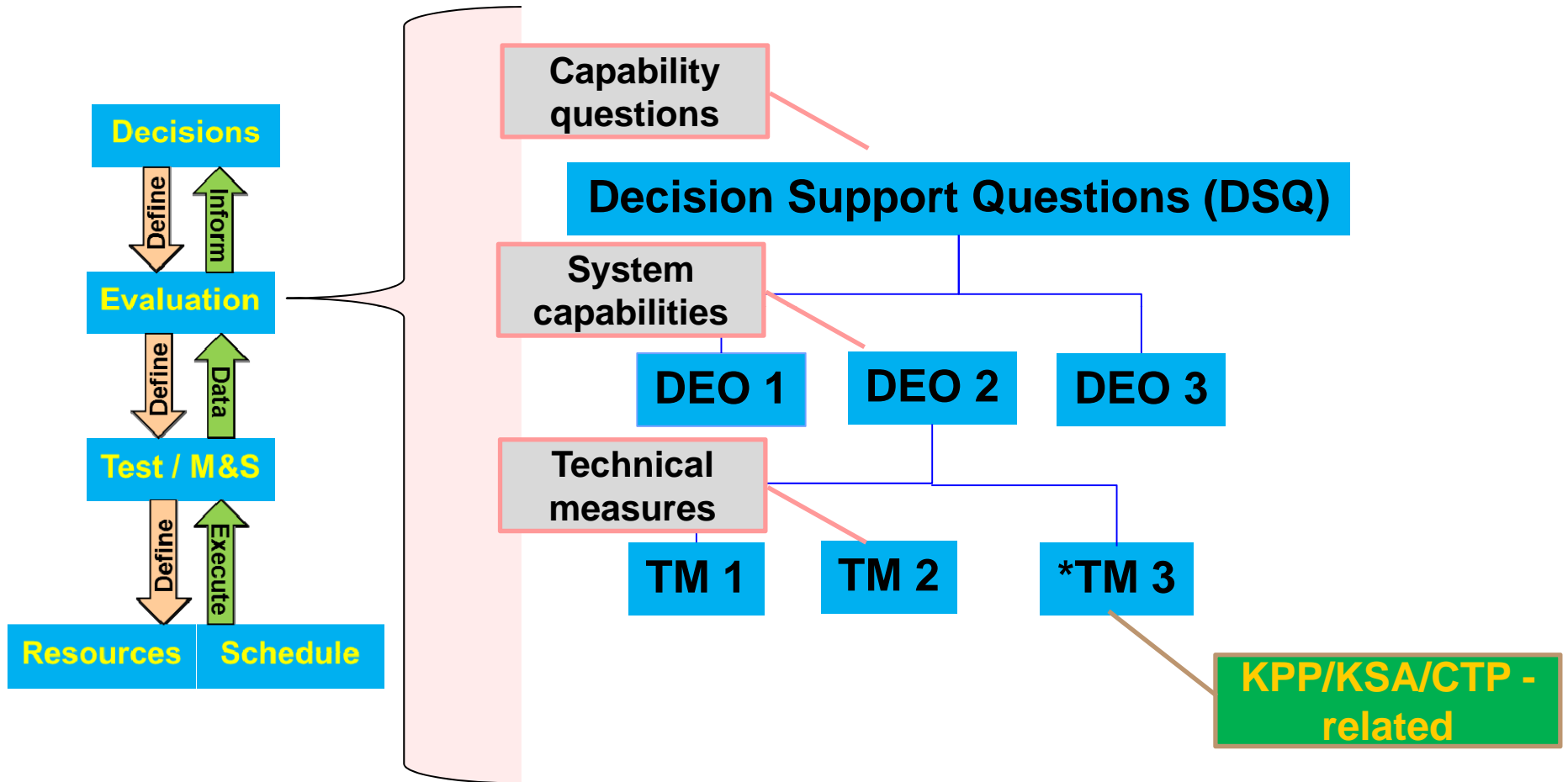
- How acquisition, programmatic, technical and operational decisions will be *informed* by evaluation
- How system will be *evaluated* to generate the knowledge needed to inform decisions
- How test and M&S events will provide *data* for evaluation
- What *resources* are required to execute test, conduct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



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



# Decision Support with an Evaluation Focus

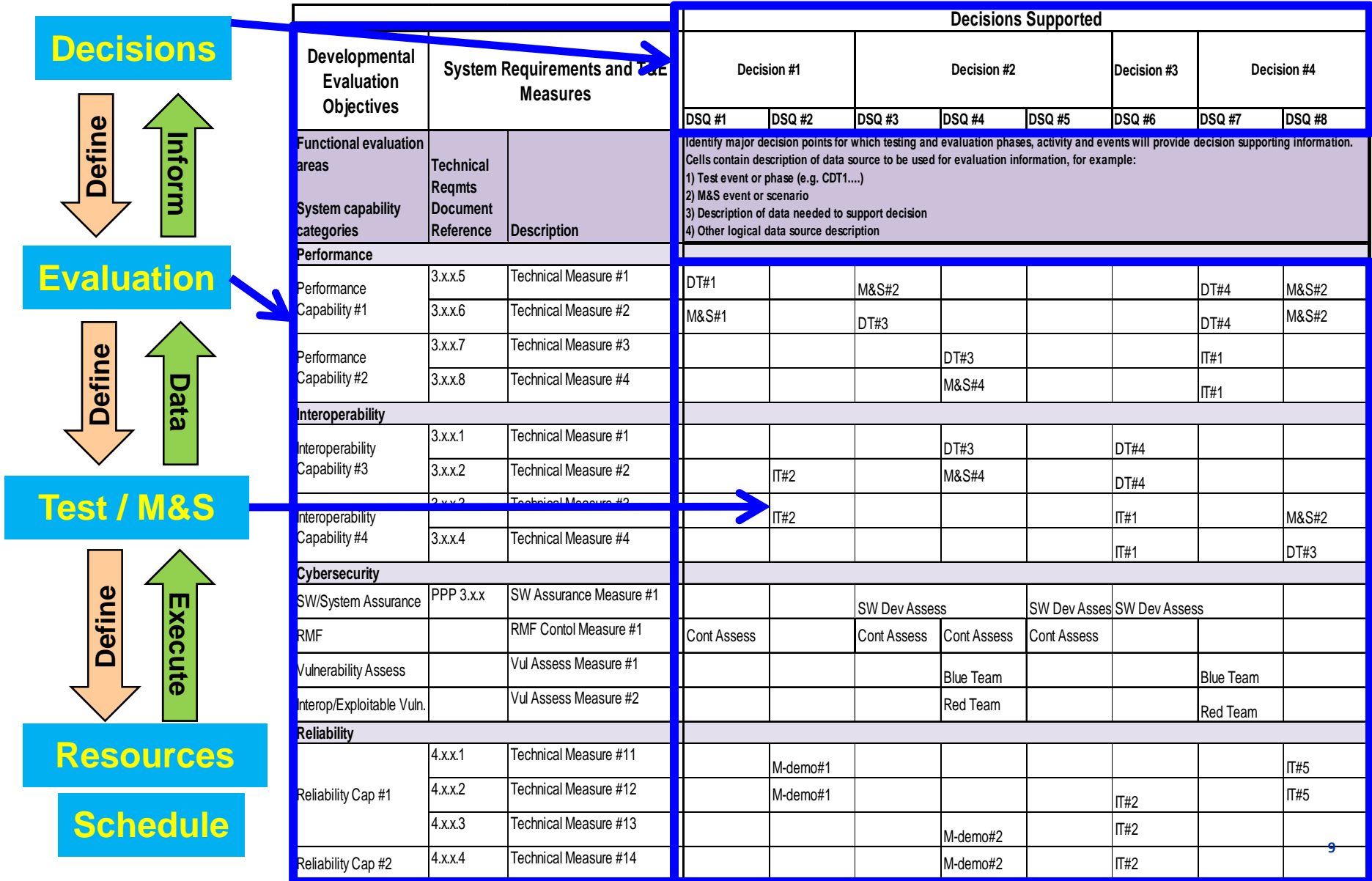


**System Engineering decomposition:  
Evaluate system capability - Inform decisions**





# Developmental Evaluation Framework





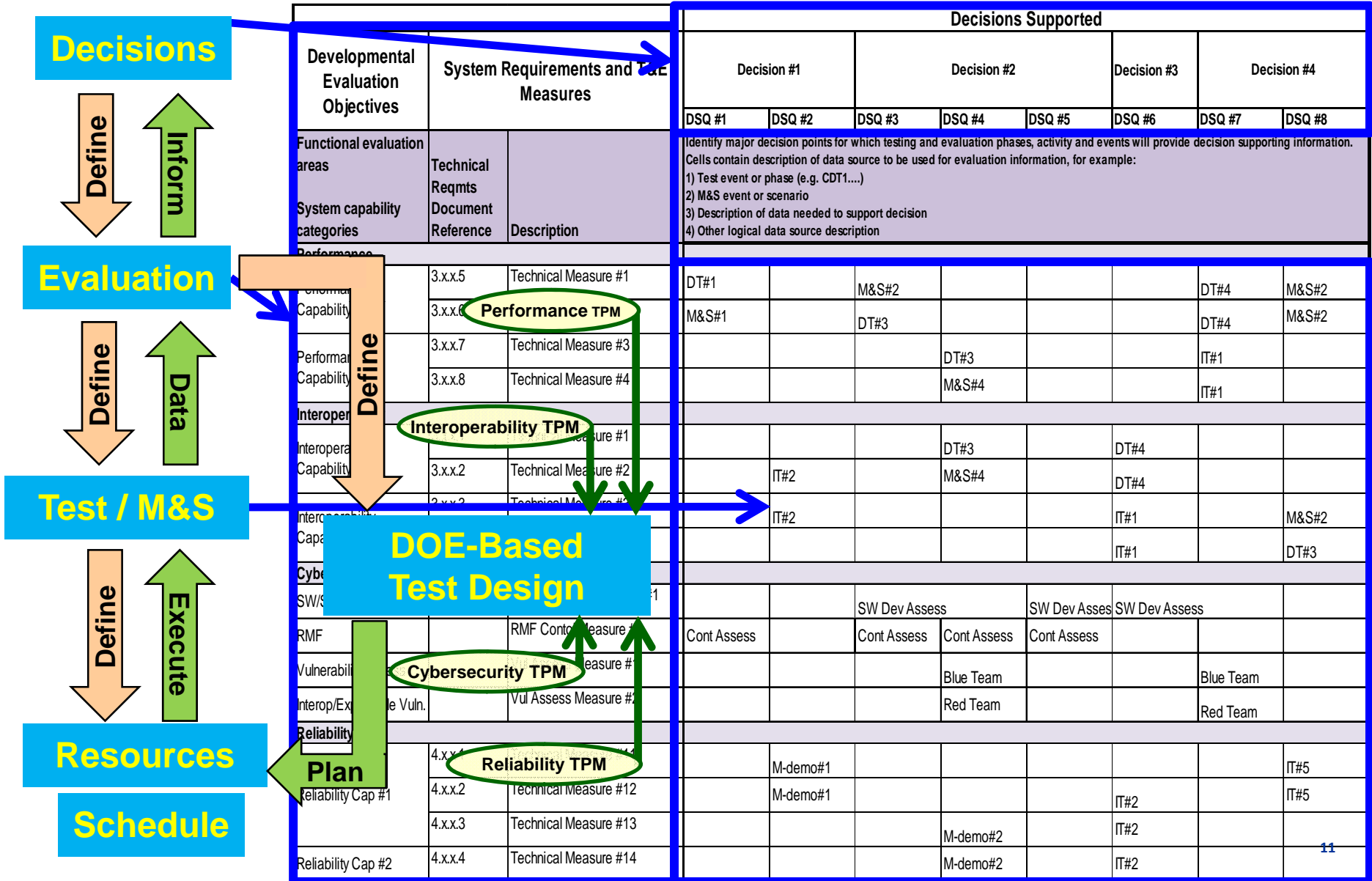
# *Then Plan the Test or Bringing it Back Together as IT*



- **With Evaluation Frameworks developed, informed integrated test planning can proceed**
  - 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*
- **How to 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 DOE to generate common test cases

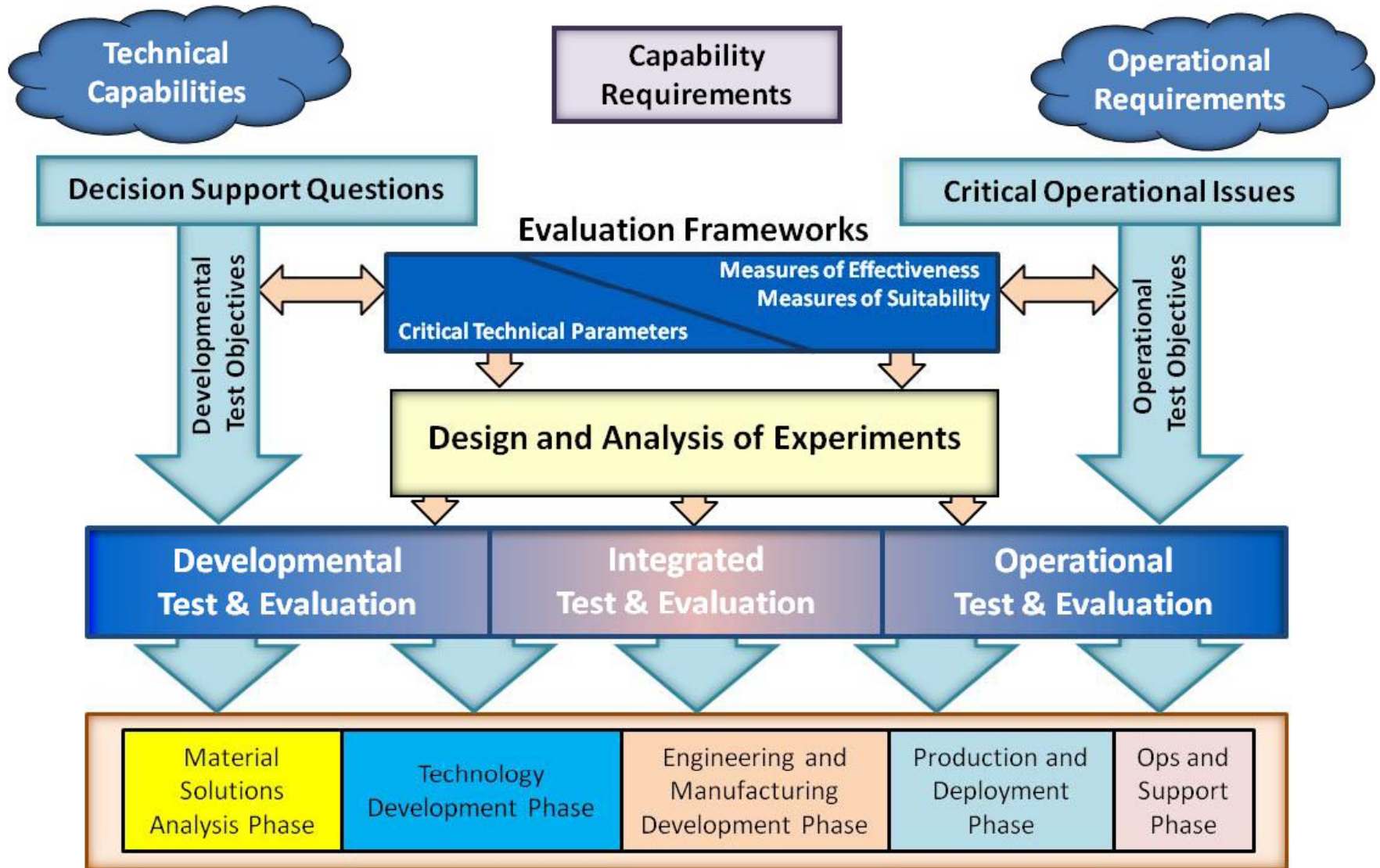


# Developmental Evaluation Framework



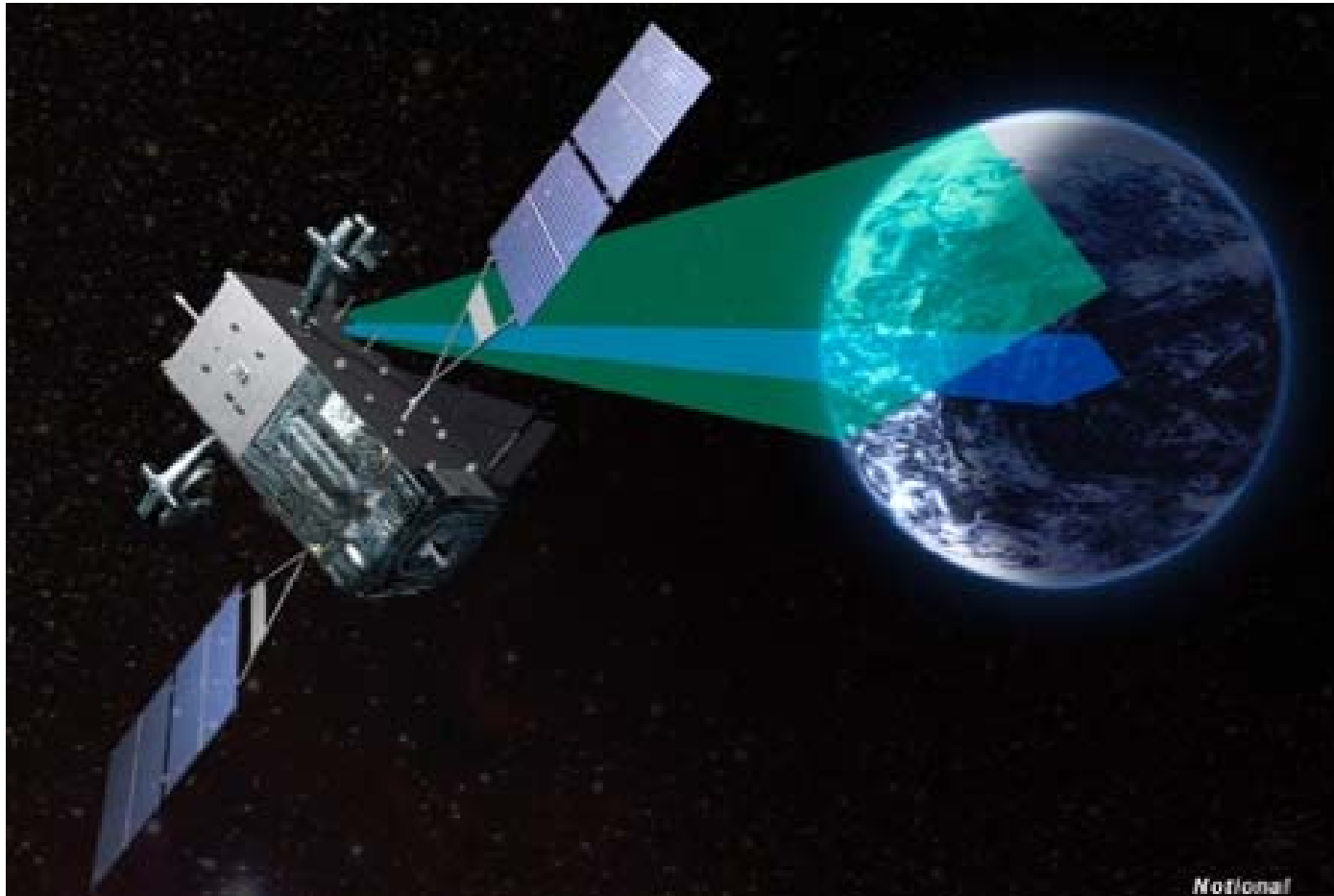


# The Place for Design and Analysis of Experiments in T&E





# Example - Space Based Infrared System (SBIRS)



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



# SBIRS DEF



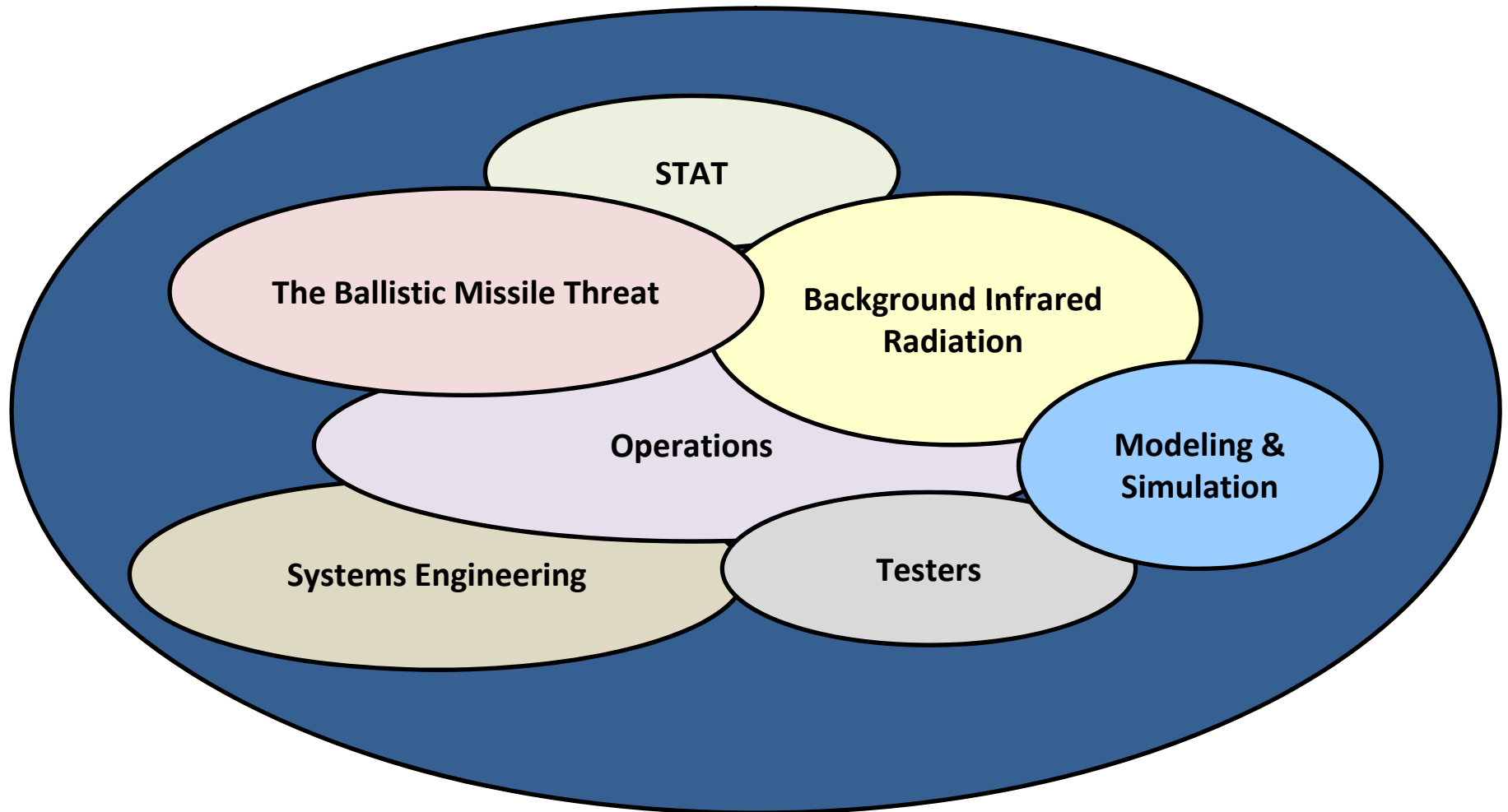
Developmental Evaluation Objectives	System Requirements / Measures		Decisions Supported				
			Space Segment		Ground Segment	S2E2	
			SV Flight Readiness	SV ops acceptance	Ground (Blk 20) readiness (PEO Certification)	Govt DT TRR	Inc 1 (DSP) PEO Certification
<b>DEO/Capabilities</b>			SV performance and stability for launch	<b>Decisions &amp; DSQs</b>			
Performance				Space Segment			
System Mission Performance				SV Flight Readiness			
Mission Data Collection				SV performance and stability sufficient to launch			
Reporting				SV Ops Acceptance			
Functionality				Tuning & on-orbit performance sufficient to accept into constellation?			
System Survivability/Endurability				Ground Segment			
Missino Management				Ground (Blk 20) readiness (PEO Certification)			
Interoperability				Blk 20 performance satisfy HCS requirements?			
Mission interface compliance				S2E2			
Net ready				Govt DT TRR			
Cybersecurity				Does S2E2 operate thru environmental reqmts?			
System/SW Assurance				Inc 1 (DSP) PEO Certification			
Compliance (RMF)				Mono/DSP performance satisfy S2E2 SRD reqmt?			
Vulnerability Assessment				Inc 2 (GEO) PEO Certification			
Reliability				GEO performance satisfy S2E2 SRD reqmt?			
Suitability							
Logistics							
<b>Requirements/Technical Measures</b>							
Suitability	3.4.3	3.3.6	Dependability	High Component Spec (HCS), S2E2 SRD			
	3.4.4	3.3.4	Reliability				
	3.4.6	3.3.6	Maintainability				
	3.4.2	3.6.6	EMC				
	3.7	3.6.14	Human Factors				
	3.4.8		Supportability				
	3.2.2.2	3.2.1.20	Ground Segment Loading				
		3.3.5	Availability				
Logistics	3.5	3.6	ILS support				
	3.5.5; 3.5.6	3.7	Personnel & training				





# SBIRS DOE Working Group

## Enterprise-wide SME Involvement



"An approximate answer to the right problem is worth a good deal more than an exact answer to an approximate problem." -- John Tukey





# SBIRS COI 1 Factor Space



COI 1 Factors (Version 2.1 - 30 April 2014)

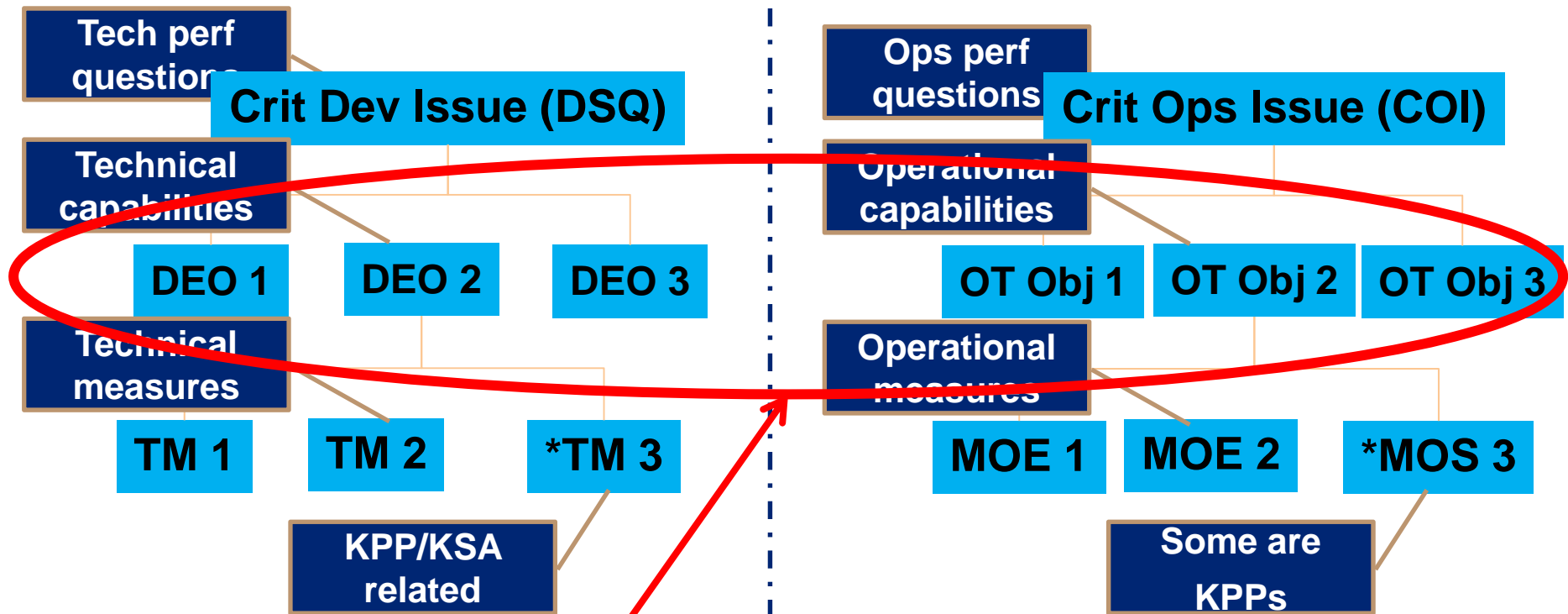
Factor Name	OLGASim Factor	NG Rank	LM Rank	Factor Type	Factor Subtype	Levels	Factor Management	Level Descriptors	Factor Type
Missile Type <sup>1</sup>	x	1	1	Categorical	Nominal	5	Vary	ICBM, SLBM, IRBM, MRBM, SRBM Small, Medium, Large None, A, B, C (Demo A, B, C) Min, Max Short, Intermittent, Long Min,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



# EFs Defines IT Data Needs



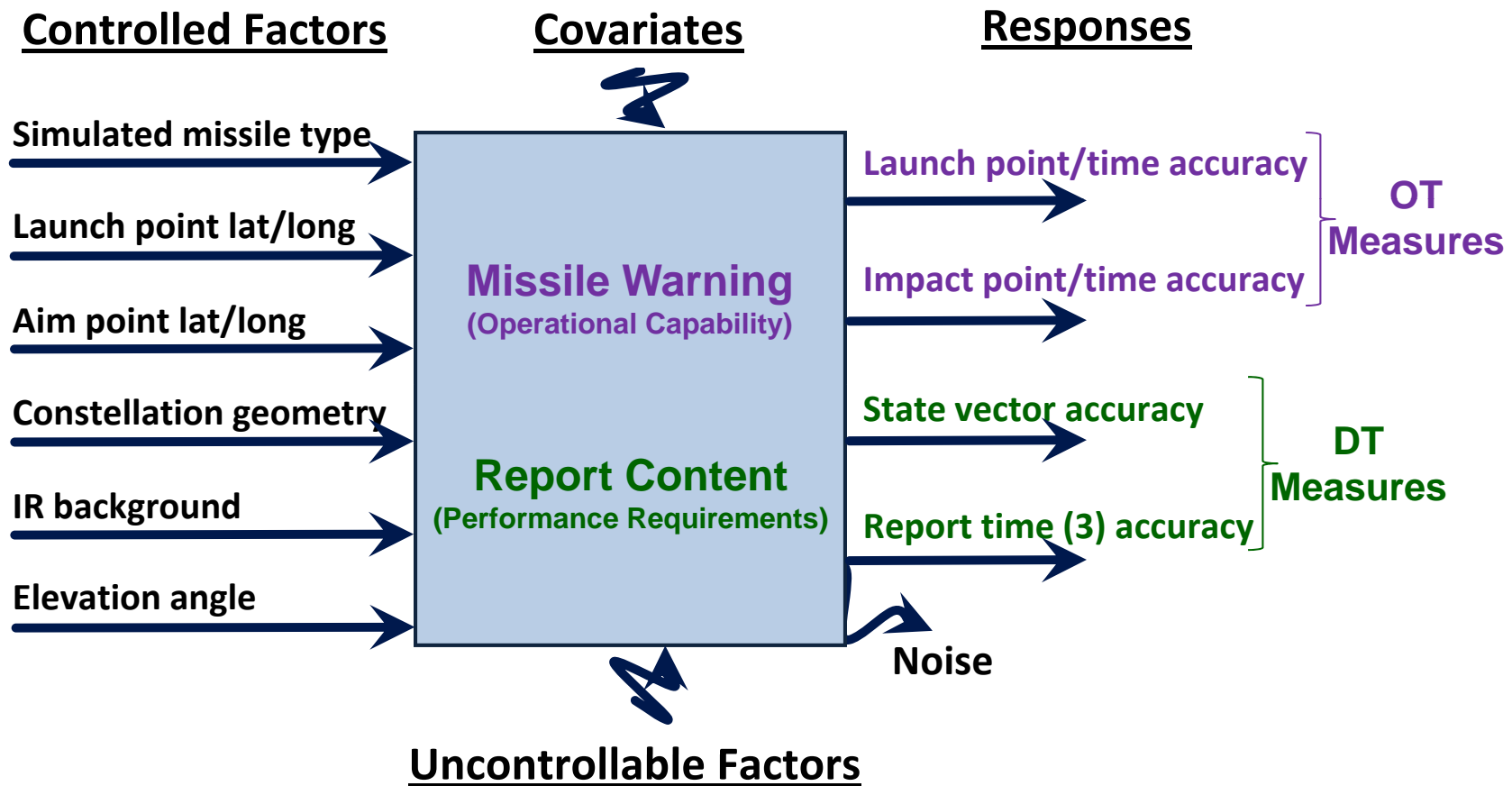
*DT EF* | *OT EF*



Potential common data for IT



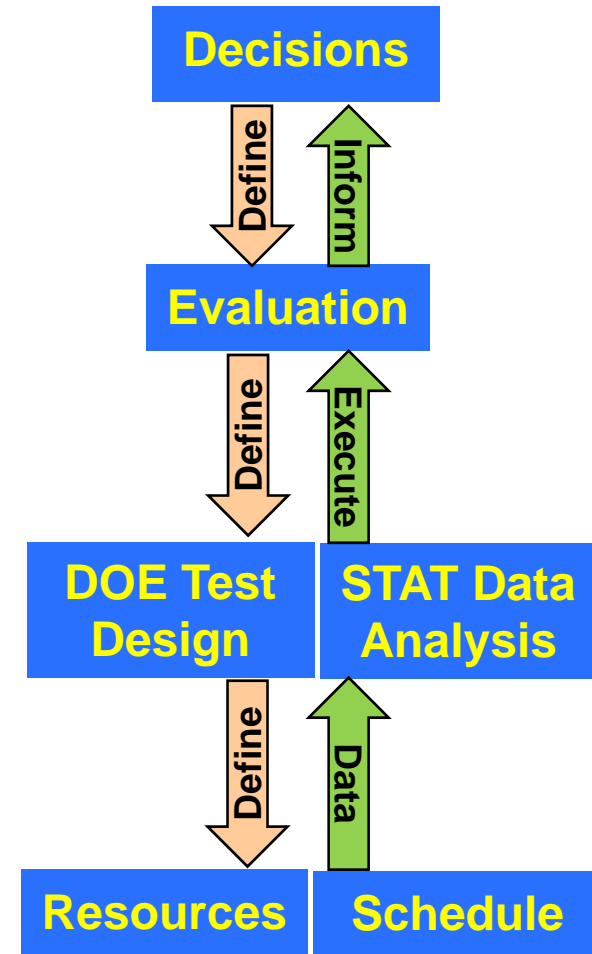
# IT Design Example – IPO Diagram





# Summary & Way Ahead

- “E”: DEF focuses system evaluation to inform decisions
  - DSQ (decision) → DEO (capability) → TM (measure)
- “T”: DOE & STAT adds rigor to the T&E strategy and feeds the DEF
  - Plan (measures, factors, test design) → Execute (test points) → Data (statistical analysis) → Inform



*decision-quality information*

....all others must bring data