International Test & Evaluation LVC Conference

Capability Test Design and Analysis

January 13, 2009

Mark J. Fiebrandt
Lead Operations Research Analyst, M&P Div, JTEM
Mark.Fiebrandt@jte.osd.mil
757.638.6055

Dr. David A. Dryer
Lead Systems Engineer, M&P Div, JTEM
David.Dryer@jte.osd.mil
757.638.6091

jtem@jte.osd.mil

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
13 JAN 2009		**00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Capability Test Design and Analysis	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Joint Test and Evaluation Methodology (JTEM),Washington,DC,20301**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Live-Virtual Constructive Conference, 12-15 Jan 2009, El Paso, TX

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**24**	

Capability Test Design & Analysis Objectives

1. Discuss the concept of conducting a capability evaluation strategy refinement process for testing in a joint environment (TIJE)
2. Review the methods and processes for an evaluation strategy refinement process
3. Review potential design of experiment techniques for large number of factors
4. Review tools and techniques for an evaluation strategy refinement process
5. Step through a “case study” example of an evaluation strategy refinement process
6. Review potential issues and insights

JTEM Capability Test Methodology (CTM) v2.0

6 Steps
14 JTEM Processes

0. Develop T&E Strategy

- T&E Strategy (TES)
- T&E Master Plan (TEMP)

- Develop Capability/SoS Description
- Develop Joint Operational Context for Test (JOC-T)
- Develop Evaluation Strategy
- Develop/Refine Capability Crosswalk

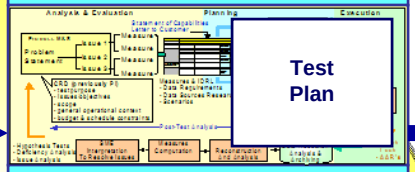
1. Characterize Test

- Program Introduction Document (PID)
- Statement Of Capabilities (SOC)

- Develop Test Concept
- Refine Evaluation Strategy
- Technical Assessment

Capability Subset Focus

2. Plan Test



- Develop Test Design
- Perform LVC Distributed Environment Analysis
- Develop Test Plan

Test Concept

Integrated Vignettes

Joint Operational Context for Test

LVC Distributed Environment Design

3. Implement LVC-DE

- Design LVC Distributed Environment Configuration
- Integrate LVC Distributed Environment

Joint Mission Environment

Test Data

Test Control & Monitoring

5. Evaluate Capability

- Analyze Data
- Evaluate SoS Performance & Joint Mission Effectiveness

Capability Set Focus

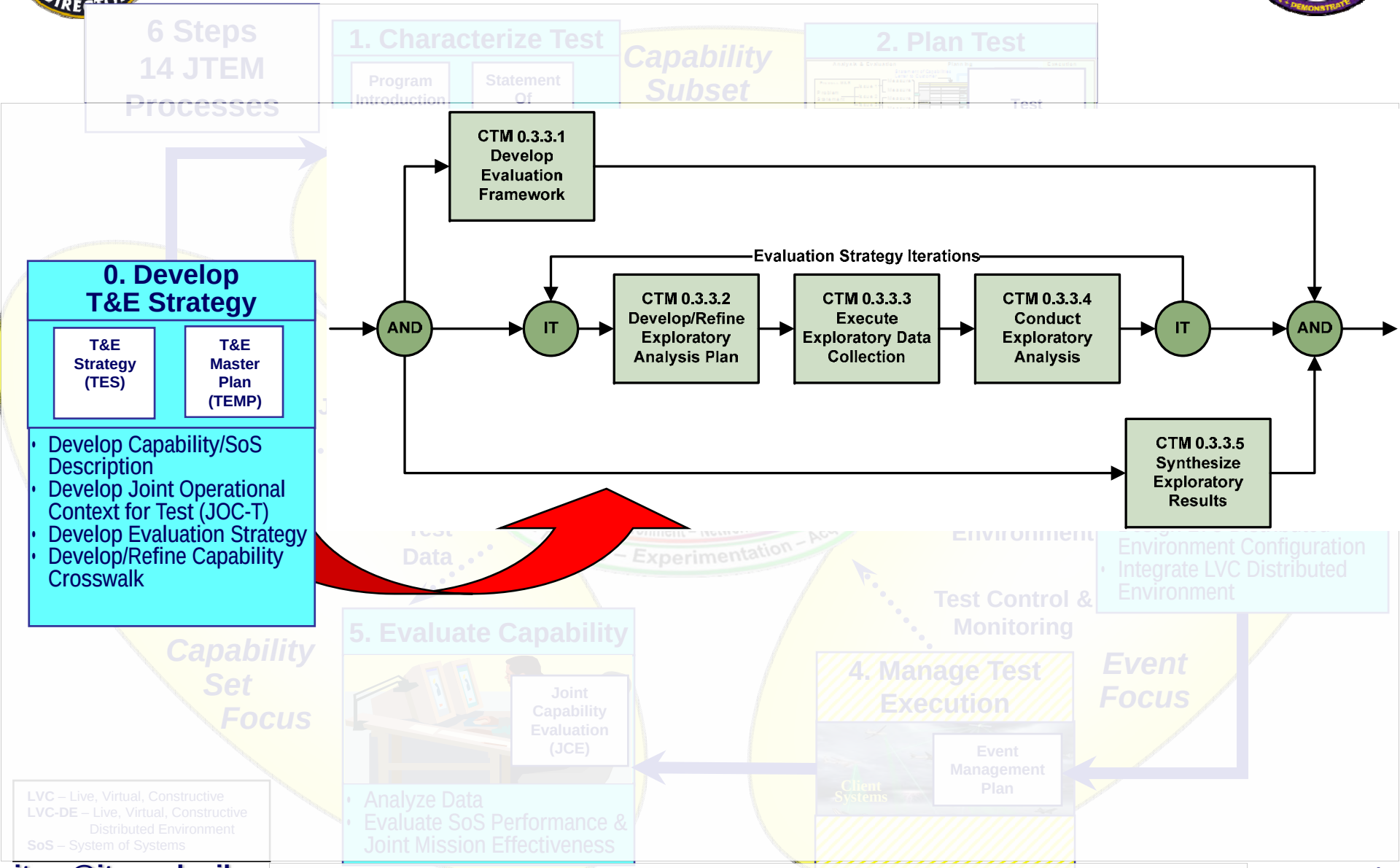
4. Manage Test Execution

Event Focus

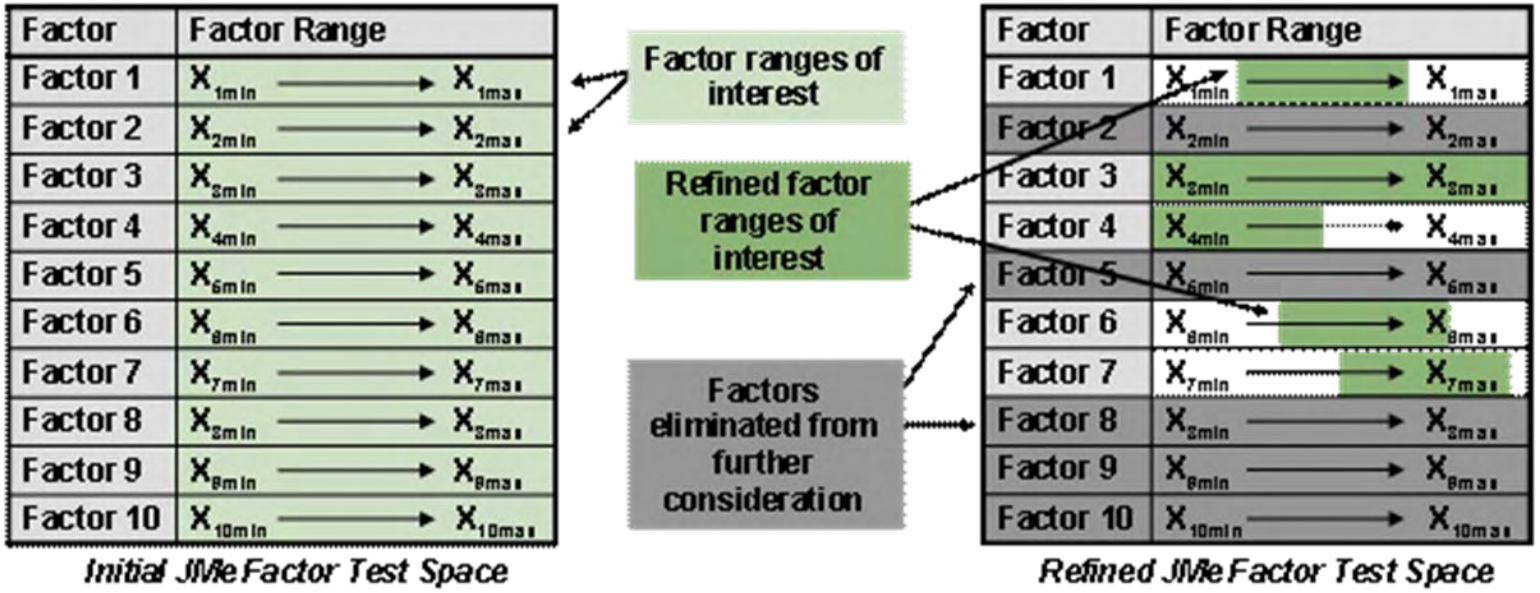
LVC – Live, Virtual, Constructive
LVC-DE – Live, Virtual, Constructive Distributed Environment
SoS – System of Systems

jtem@jte.osd.mil

JTEM CTM 0.3.3 (Develop Evaluation Strategy)



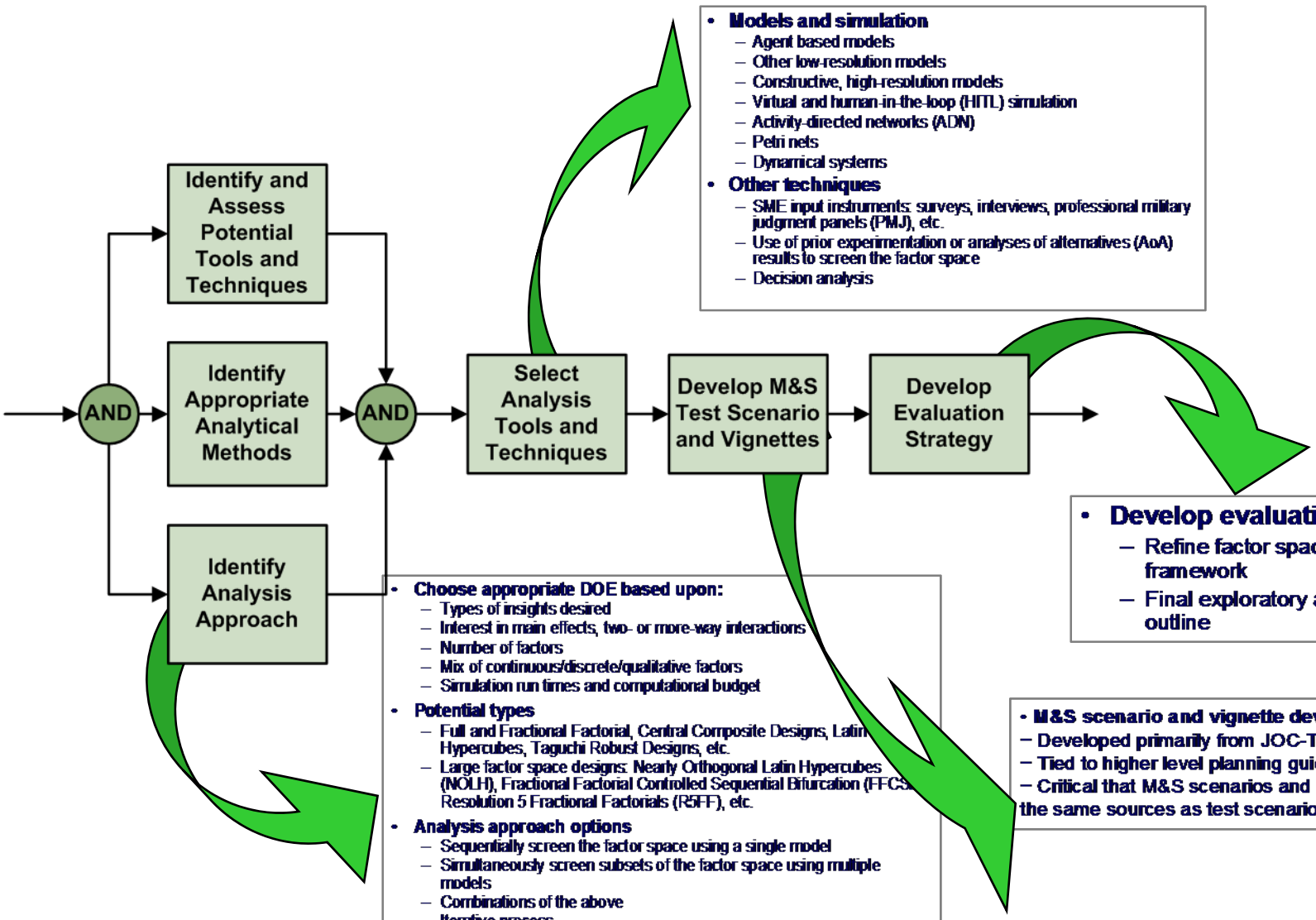
Exploratory Analysis Purpose



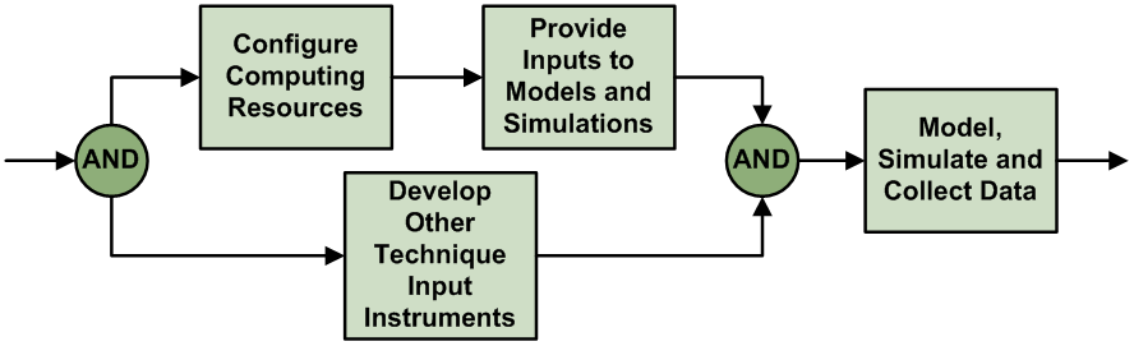
Purpose is three-fold:

- To explore a wide range of possible factors and levels that might affect joint mission effectiveness (JMe), referred to as the initial JMe factor test space;
- To identify those combinations of factors that have the greatest impact on JMe, referred to as the refined JMe factor test space; and
- To recommend potential factor combinations of interest from the refined JMe factor test space for subsequent test events, referred to as potential test trial sets.

CTM 0.3.3.2: Develop/Refine Exploratory Analysis Plan

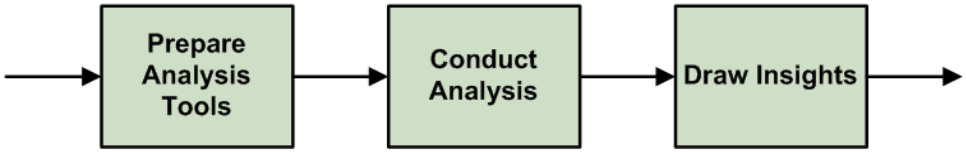


CTM 0.3.3.3: Execute Exploratory Data Collection

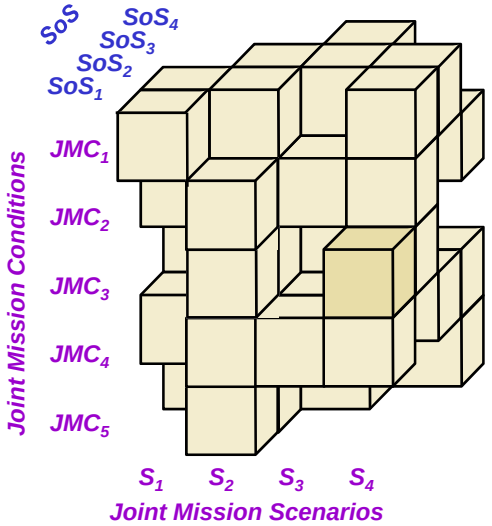


- Depending upon tools and methods chosen, configuration of computing resources may be quite involved.
- Other instruments of data collection are also developed here (surveys, interview scripts, PMJ panel planning, etc.).
- Analysts develop model inputs (scenario files, DOE input files, data output specifications, etc.) to execute the data collection.
- Model runs and data collection are executed according to exploratory analysis strategy.

CTM 0.3.3.4: Conduct Exploratory Analysis

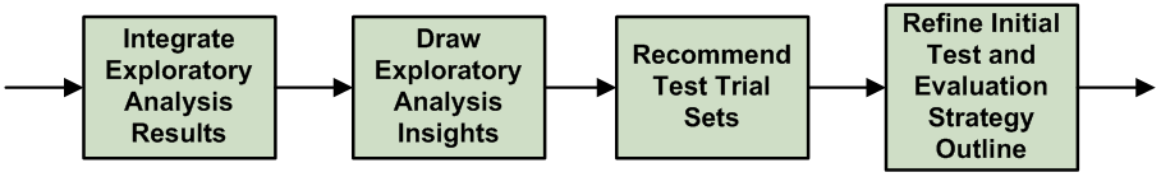


- Identify and select appropriate tools and techniques to conduct the analysis.
- Evaluate measure responses from exploratory runs to refine the JMe factor test space.
- Identify factors to explore during next iteration, as required.
- Potential analyses involved in these steps will be discussed in more detail as part of CTM 5, Evaluate Capability, since similar analytic methods will be used as part of both processes.

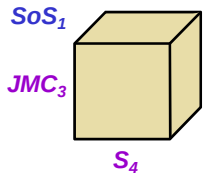


Refined JMe Factor Test Space

CTM 0.3.3.5: Synthesize Exploratory Analysis Results

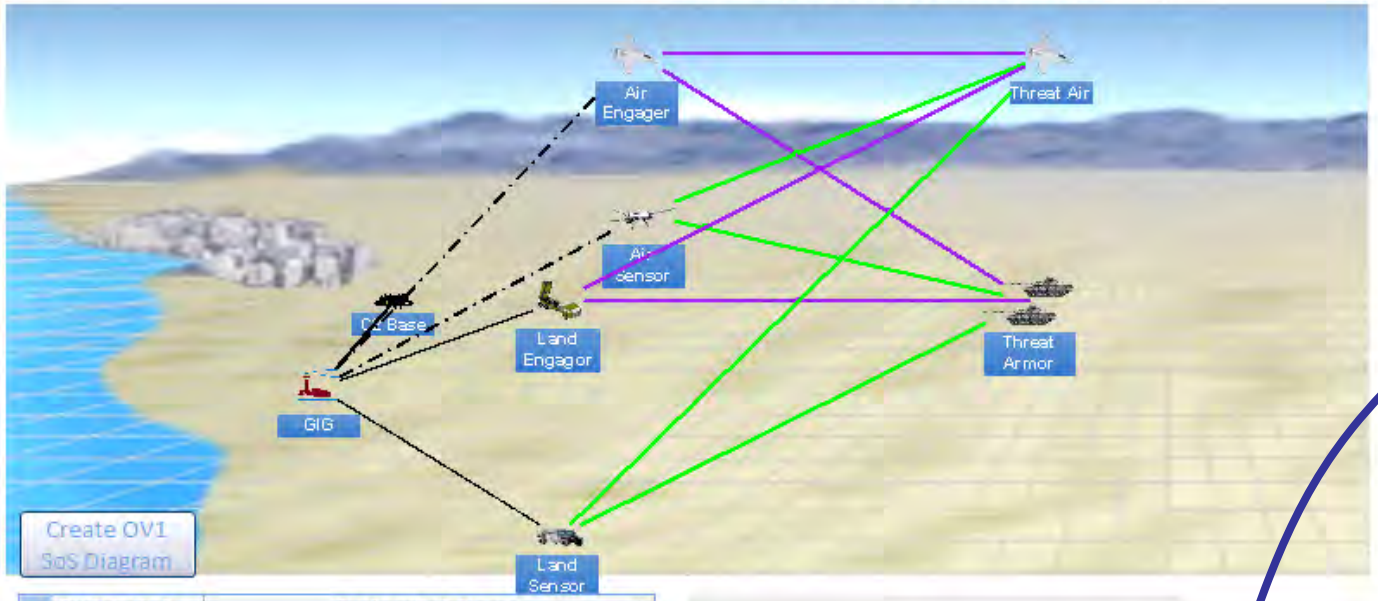


- This set of processes integrates the analyses conducted during the multiple exploratory iterations to draw insights about the "probable" factor space and the measure framework.
- The result should be a final refined factor space consisting of potential test trial sets of interest for subsequent testing.
- Must integrate model related data and qualitative data obtained from SMEs.
- Insights from the analysis will help inform the risk assessment conducted in the next step of the CTM.
  - Potential contributors to risk include the assumptions made during modeling, the capabilities of the models, the measures chosen, etc.
  - Subsequent tests can help validate assumptions made.



Potential Test Trial Set

Case Study Example
High Level Operational View (OV-1)



System of Systems (SoS)

SystemSoS	SystemSoS Name
IDFW SoS 1	C2 Capability System of Systems

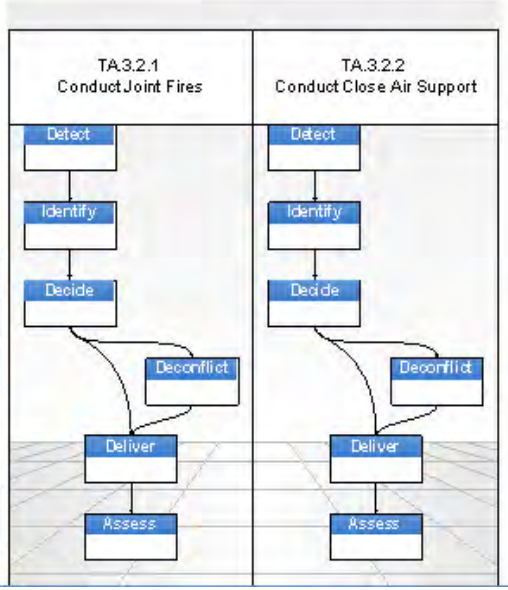
Record: 1 of 1 | No Filter | Search

Task

to perform

UJTL Numbe	Task Name
TA.3.2.1	Conduct Joint Fires
TA.3.2.2	Conduct Close Air Support

Record: 1 of 2 | No Filter | Search



Create OV1 Task Diagram

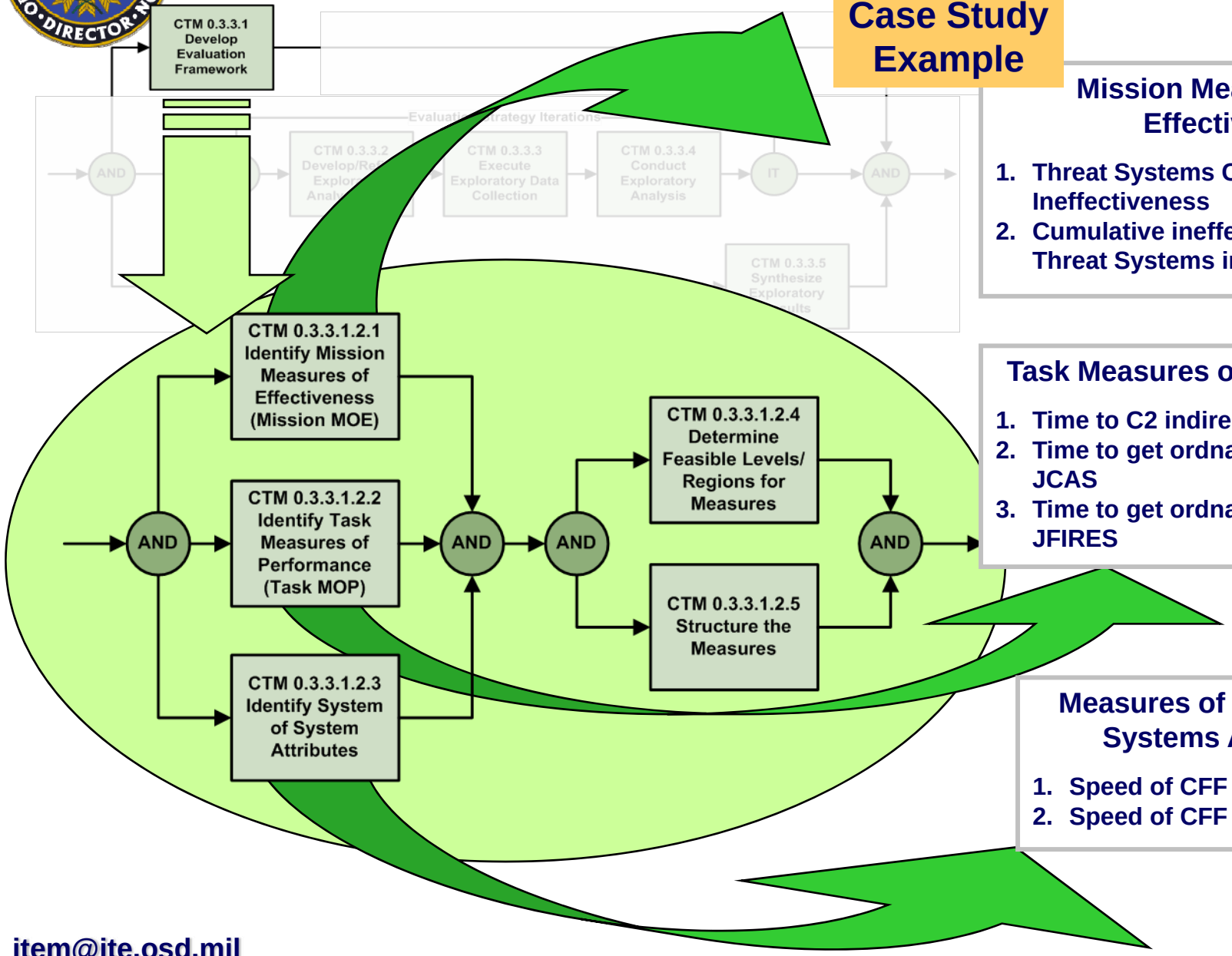
to achieve

Desired Effect	Desired Effect Name
IDFW MDE 1	Threat Platform Ineffectiveness

Record: 1 of 1 | No Filter | Search

Mission Desired Effect

Develop Evaluation Framework



Case Study Example

Mission Measures of Effectiveness

1. Threat Systems Combat Ineffectiveness
2. Cumulative ineffectiveness time Threat Systems in JOA

Task Measures of Performance

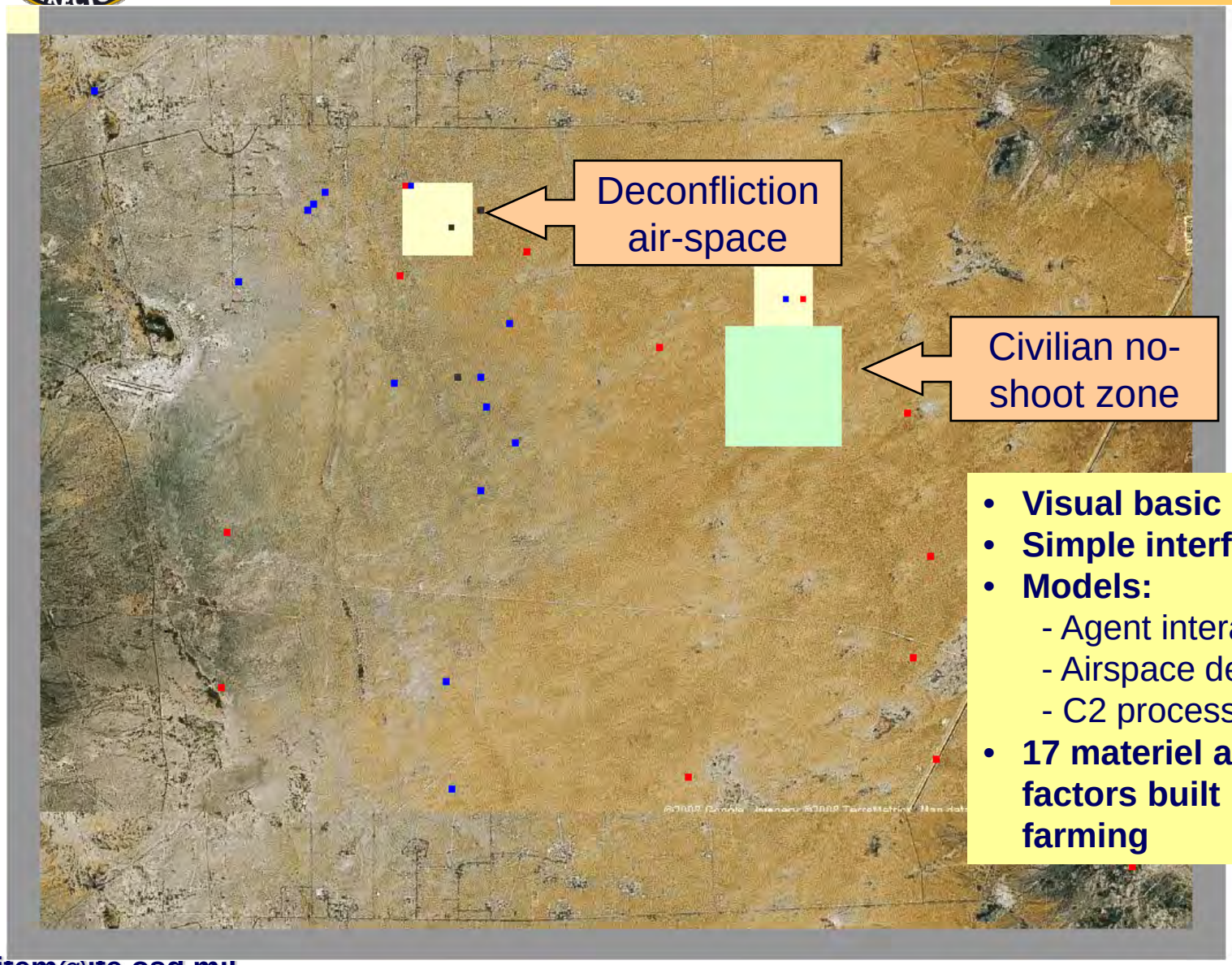
1. Time to C2 indirect fires (IF)
2. Time to get ordnance on target for JCAS
3. Time to get ordnance on target for JFIRES

Measures of System of Systems Attributes

1. Speed of CFF Decisions
2. Speed of CFF Deconfliction

Agent Based Model

Case Study Example



- **Visual basic code**
- **Simple interface**
- **Models:**
  - Agent interactions
  - Airspace deconflictions
  - C2 processes
- **17 materiel and non-materiel factors built in for data farming**

Factor Capability Crosswalk SME Estimates

Crosswalk Dimension	Crosswalk Sub Dimension	Factor	Levels	Factor Type
System of Systems	Materiel	1. Global Information Grid (GIG)?	Yes/No	Categorical
		2. Blue Speed	1/2 (Multiplier)	Continuous
		3. Blue Monitor	1/2 (Multiplier)	Continuous
		4. Blue Fires	1/2 (Multiplier)	Continuous
	Non-materiel: Doctrine	5. Multiple Trackers?	Yes/No	Categorical
		6. Expedited Call For Fire?	Yes/No	Categorical
		7. Call For Fire type Decision	A/C	Categorical
		8. Expedite Move	Yes/No	Categorical
		9. Restricted Op. Zone Type	Restrictive/Permissive	Categorical
		10. ROZ Size	1/2 (Multiplier)	Continuous
		11. ROZ Expiration	Yes/No	Categorical
		12. ROZ Slack Time	1/2 (Add. Time Increments)	Continuous
		13. Multi-Service Wait time	1/2 (Add. Time Increments)	Continuous
Condition	Environmental	14. Adverse Weather?	Yes/No	Categorical
		15. Civilian Zone?	Yes/No	Categorical
		16. Civilian Zone Size	20/40	Continuous
		17. Civilian Zone Location	1/2	Categorical
		17 Factors with 3 dependencies		*7 Continuous/ 10 Categorical*

Case Study Example

Step 1: "Quick Look" Analysis

Dependent Variable: Number Threat Kills
Independent Variables: 17 factors (decision & conditional)
DOE: Resolution III Fractional Factorial (80 trials, 20 runs each)
Analysis tools: Stepwise Regression Model

SME Estimate

X

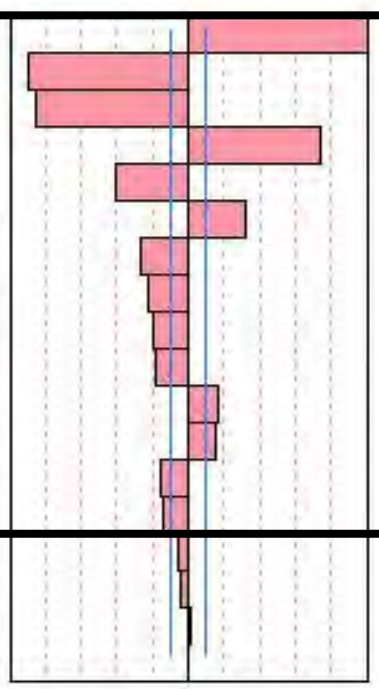
X

X

X

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Blue Speed	29.0625	1.409628	20.62	<.0001*
ROZ Size	-26.1625	1.409628	-18.56	<.0001*
CFF Type of Decision[A]	-12.39375	0.704814	-17.58	<.0001*
ROZ Type[P]	10.71875	0.704814	15.21	<.0001*
Multiple Trackers?[No]	-5.81875	0.704814	-8.26	<.0001*
ROZ Slack Time	9.3875	1.409628	6.66	<.0001*
GIG?[Off]	-3.81875	0.704814	-5.42	<.0001*
Civilian Zone?[3]	-6.58125	1.409628	-4.67	<.0001*
ROZ Expiration[No]	-2.79375	0.704814	-3.96	0.0001*
Expedite Move[No]	-2.63125	0.704814	-3.73	0.0003*
Blue Fires	4.8375	1.409628	3.43	0.0008*
Blue Monitor	4.4375	1.409628	3.15	0.0020*
Expedited CFF?[No]	-2.21875	0.704814	-3.15	0.0020*
Adverse Weather?[No]	-2.11875	0.704814	-3.01	0.0031*
Multi-service Wait time	-1.5875	1.409628	-1.13	0.2620
Civilian Zone?[4]	-1.175	1.409628	-0.83	0.4059
Civilian Zone?[1]	0.41875	1.409628	0.30	0.7669
Civilian Zone?[2]	0.0125	1.409628	0.01	0.9929



Significant factors

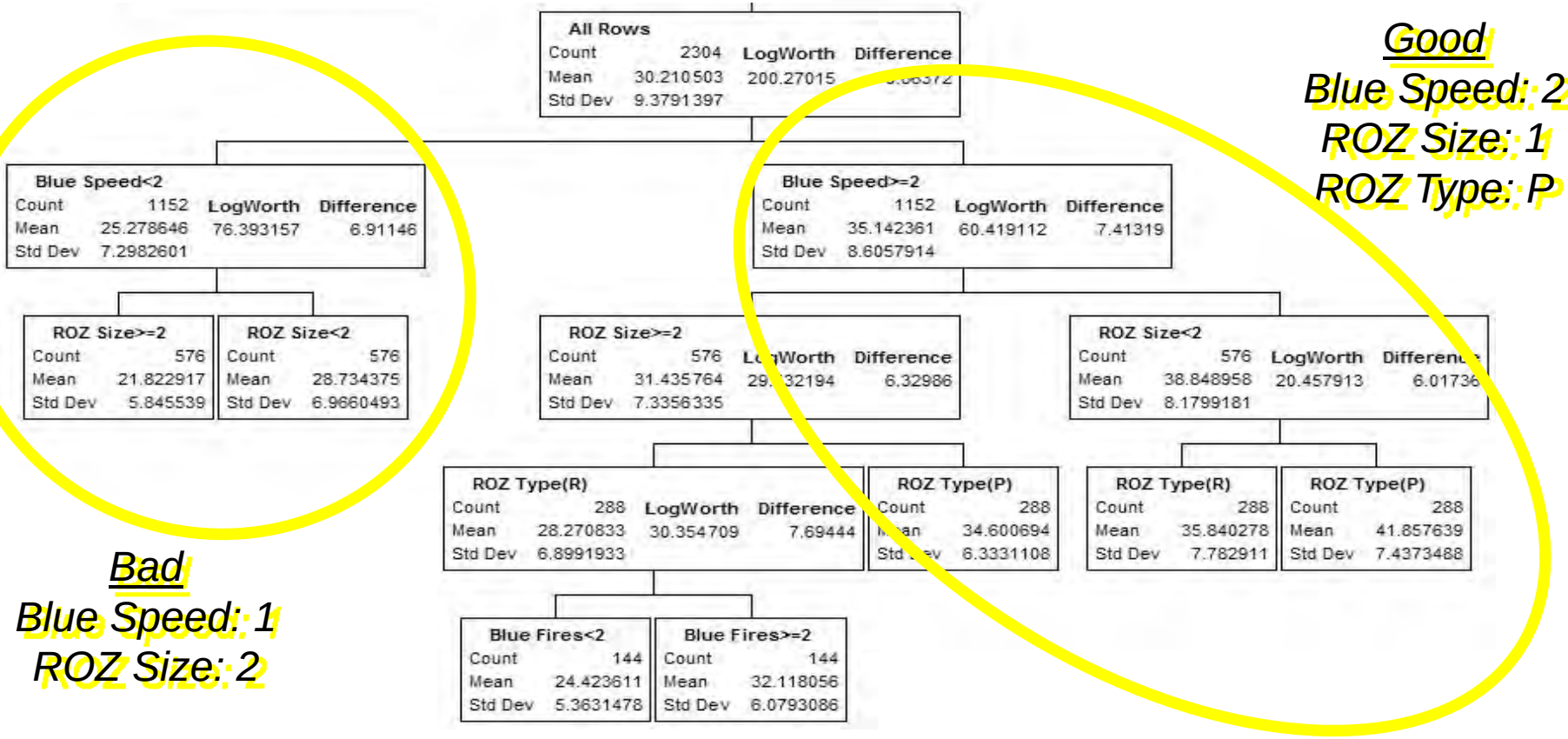
Non-Significant factors

Findings:
- 14 factors significant, 3 factors not significant
- Adverse weather factor non-intuitive

jtem@jte.osd.mil

Step 2: Main Effects Analysis

Dependent Variable: Number Threat Kills
Independent Variables: 14 factors (decision & conditional)
DOE: Resolution V Fractional Factorial (2304 trials, 3 runs each)
Analysis tools: Classification and Regression Tree (CART) partitioning



Good
Blue Speed: 2
ROZ Size: 1
ROZ Type: P

Bad
Blue Speed: 1
ROZ Size: 2

Findings:
- Stressing factors (Blue speed, ROZ size, ROZ type)

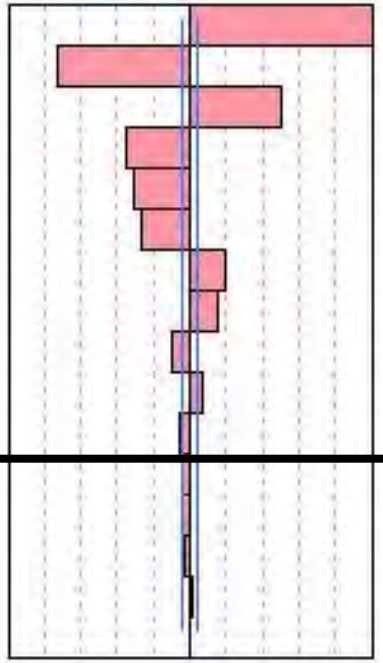
Step 2: Main Effects Analysis

Dependent Variable: Number Threat Kills
Independent Variables: 14 factors (decision & conditional)
DOE: Resolution V Fractional Factorial (2304 trials, 3 runs each)
Analysis tools: Stepwise Regression Model

SME Estimate

Sorted Parameter Estimates

SME Estimate	Term	Estimate	Std Error	t Ratio	Prob>|t|
X	Blue Speed	9.8637153	0.243801	40.46	<.0001*
	ROZ Size	-7.162326	0.243801	-29.38	<.0001*
X	ROZ Type[P]	2.4657118	0.1219	20.23	<.0001*
X	GIG?[Off]	-1.751302	0.1219	-14.37	<.0001*
	CFF Type of Decision[A]	-1.508247	0.1219	-12.37	<.0001*
	Multiple Trackers?[No]	-1.313802	0.1219	-10.78	<.0001*
	Civilian Zone?[No]	1.3897569	0.172393	8.06	<.0001*
	Civilian Zone?[Yes20]	1.0837674	0.172393	6.29	<.0001*
	Expedite Move[No]	-0.455295	0.1219	-3.73	0.0002*
	Adverse Weather?[No]	0.3658854	0.1219	3.00	0.0027*
	Expedited CFF?[No]	-0.30599	0.1219	-2.51	0.0121*
	ROZ Expiration[No]	-0.325087	0.172393	-1.89	0.0595
	Blue Monitor	-0.426215	0.243801	-1.75	0.0806
	Multi-service Wait time	-0.243924	0.243801	-1.00	0.3172
X	Blue Fires	0.0894097	0.243801	0.37	0.7139
	ROZ Expiration[Yes1]	-0.03342	0.172393	-0.19	0.8463



Significant factors

Non-Significant factors

Findings:
- 10 factors significant
- Blue fires no longer significant
- Adverse weather factor intuitive

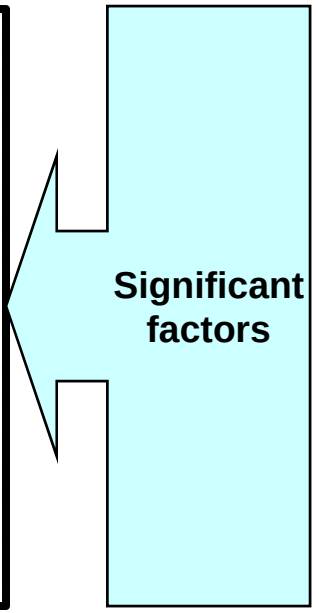
Step 2: Main Effects & Two-way Interaction

Dependent Variable: Number Threat Kills
Independent Variables: 14 factors (decision & conditional)
DOE: Resolution V Fractional Factorial (2304 trials, 3 runs each)
Analysis tools: Stepwise Regression Model

SME Estimate

Sorted Parameter Estimates

	Term	Estimate	Std Error	t Ratio	Prob>|t|
X	Blue Speed	9.8637153	0.202018	48.83	0.0000*
	ROZ Size	-7.162326	0.202018	-35.45	<.0001*
X	ROZ Type[P]	2.4657118	0.101009	24.41	<.0001*
X	(Blue Fires-1.5)*ROZ Type[P]	-4.063368	0.202018	-20.11	<.0001*
X	GIG?[Off]	-1.751302	0.101009	-17.34	<.0001*
	Civilian Zone?{Yes40-Yes20&No}	-1.855143	0.107136	-17.32	<.0001*
	CFF Type of Decision[A]	-1.508247	0.101009	-14.93	<.0001*
	Multiple Trackers?[No]	-1.313802	0.101009	-13.01	<.0001*
	Adverse Weather?[No]*ROZ Type[P]	1.2773438	0.101009	12.65	<.0001*
	(Blue Monitor-1.5)*ROZ Type[P]	1.7161458	0.202018	8.49	<.0001*
	(Blue Speed-1.5)*ROZ Type[P]	1.2421875	0.202018	6.15	<.0001*
	GIG?[Off]*ROZ Expiration{No-Yes1}	-0.6875	0.123711	-5.56	<.0001*
	(Civilian Zone?{Yes40-Yes20&No}+0.33333)*(ROZ Size-1.5)	1.1803385	0.214273	5.51	<.0001*
	(Civilian Zone?{Yes40-Yes20&No}+0.33333)*ROZ Type[P]	-0.57194	0.107136	-5.34	<.0001*
	CFF Type of Decision[A]*ROZ Type[P]	0.531684	0.101009	5.26	<.0001*
	Adverse Weather?[No]*CFF Type of Decision[A]	0.5099826	0.101009	5.05	<.0001*



Findings:
- 6 main effect factors significant
- 5 additional factors significant in two-way interactions
- Blue fires & adverse weather part of two-way interactions

Step 3: Aggregated Conditional Factors

Dependent Variable: Number Threat Kills

Independent Variables: 10 factors (9 decision, 1 conditional)

DOE: Resolution V Fractional Factorial (128 trials, 3 runs each)

Analysis tools: Classification and Regression Tree (CART) partitioning

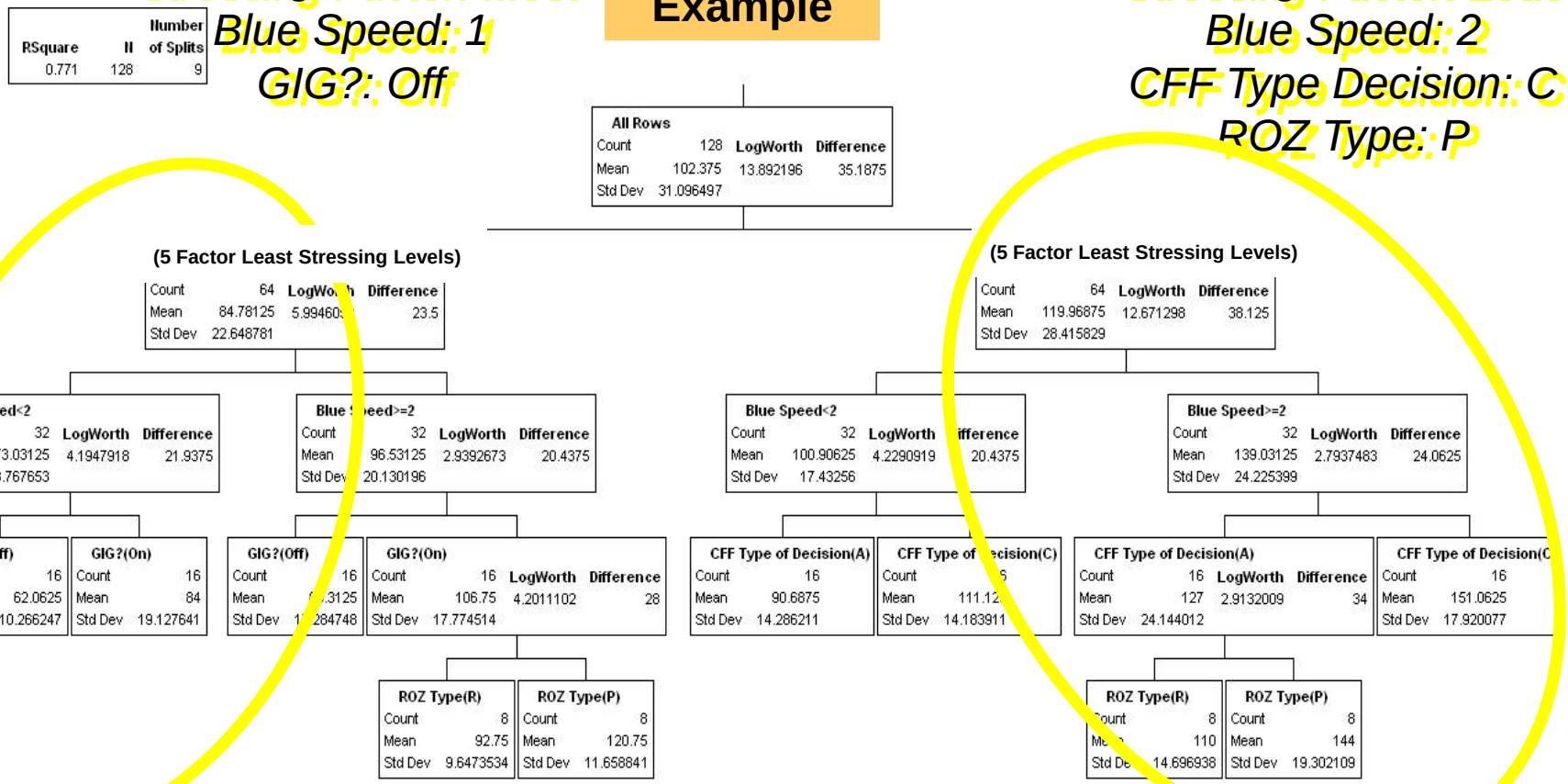
Bad

Stressing Factor: Most
Blue Speed: 1
GIG?: Off

Case Study Example

Good

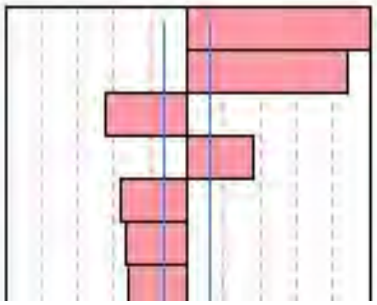
Stressing Factor: Least
Blue Speed: 2
CFF Type Decision: C
ROZ Type: P



Step 3: Aggregated Conditional Factors Results

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
(5 Factor Least Stressing Levels)	17.59375	1.122227	15.68	<.0001*
Blue Speed	30.8125	2.244455	13.73	<.0001*
(Blue Fires-1.5)*ROZ Type[P]	-15.8125	2.244455	-7.05	<.0001*
ROZ Type[P]	6.375	1.122227	5.68	<.0001*
CFF Type of Decision[A]	-6.296875	1.122227	-5.61	<.0001*
GIG?[Off]	-5.90625	1.122227	-5.26	<.0001*
Multiple Trackers?[No]	-5.59375	1.122227	-4.98	<.0001*



Dimension	Sub Dimension	Priority/Factor	Levels	Factor Type
System of Systems	Materiel	1. Blue Speed	1/2 (Multiplier)	Continuous
		2. Blue Fires	1/2 (Multiplier)	Continuous
		5. Global Information Grid (GIG)?	Yes/No	Categorical
	Non-materiel: Doctrine	3. Restricted Op. Zone Type	Restrictive/Permissive	Categorical
		4. Call For Fire type Decision	A/C	Categorical
		6. Multiple Trackers?	Yes/No	Categorical
		Most Stress 1. ROZ Size	1 (Multiplier)	Continuous
		Most Stress 2. ROZ Expiration	No	Categorical
		Most Stress 3. Multi-Service Wait time	2 (Additional Time Increment)	Continuous
Condition	Environmental	Most Stress 4. Adverse Weather?	Yes	Categorical
		Most Stress 5. Civilian Zone?	Yes (Size = 40)	Categorical

Findings:

- One aggregated condition factor significant
- Six decisional factors significant

jtem@jte.osd.mil

Step 1: Measures Relationship Table

Case Study Example

Green: Direct relationship
Red: Indirect relationship
No color: No relationship

Factors	Values	MMOE1	MMOE2	TMOP1	TMOP2	TMOP3	MOSA1	MOSA2	M1T1	M1T2	M1T3	M2T1	M2T2	M2T3	M1S1	M1S2	M2S1	M2S2	T1S1	T1S2	T2S1	T2S2	T3S1	T3S2
GIG	Off or On	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	0	1	0	1	0	1
Blue Speed	1 or 2	1	1	1	0	0	1	0	1	0	0	1	0	0	1	0	1	0	1	0	0	0	0	0
Blue Monitor	1 or 2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Blue Fires	1 or 2	1	1	0	-1	0	1	-1	0	-1	0	0	-1	0	1	-1	1	-1	0	0	-1	1	0	0
Civilian Zone	Off or On	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1
Civilian Zone Size	Small or Large	-1	-1	-1	0	0	-1	0	1	0	0	1	0	0	1	0	1	0	1	0	0	0	0	0
Multiple Trackers	No or Yes	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Adverse Weather	No or Yes	1	1	1	1	0	0	1	1	1	0	1	1	0	0	1	0	1	0	1	0	1	0	0
Expedited CFF	No or Yes	1	1	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
CFF Type of Decision	Closest or Available	-1	-1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1
Expedite Move	No or Yes	1	1	1	-1	0	1	0	1	-1	0	1	-1	0	1	0	1	0	1	0	-1	0	0	0
ROZ Type	Restrictive or Permissive	1	1	1	1	1	1	-1	1	1	1	1	1	1	1	-1	1	-1	1	-1	1	-1	1	-1
ROZ Size	1 or 2	-1	-1	-1	0	0	-1	1	1	0	0	1	0	0	1	-1	1	-1	1	-1	0	0	0	0
Multi-service Wait time	1 or 2	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
ROZ Expiration	No or Yes	1	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1
ROZ Slack Time	1 or 2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Civilian Zone Location	1 or 2	0	0	-1	0	-1	-1	-1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1

Findings (Direct and Indirect relationships):

- Direct relationship for all measures: Blue Monitor, Multiple Trackers, ROZ Slack Time
- Indirect relationship for MMOEs and TMOPs/MOSAs: CFF Type of Decision, ROZ Expiration
- Direct relationship across both MMOEs

MMOE – Mission Measure of Effectiveness
TMOP – Task Measure of Performance
MOSA – Measure of System of Systems Attributes

jtem@jte.osd.mil

Insights into Exploratory Analysis

- Resolution V DOE needed for assessing two-way interactions
  - Resolution III does not confound main effects with one another, but does confound main effects with 2-factor interactions
  - Resolution V does not confound main effects and 2-factor interactions, but confounds main with 4-factor and 2-factor with 3-factor
- Factor prioritization is an iterative process
  - Initial DOE and data farming may provide first insights into significant measures
  - May require further exploration to validate initial findings
  - May differ across multiple measures and require retaining uncertain factors in the second design
  - Iterative farming can provide additional prioritization of factors
- Factors with more than two discrete levels requires additional farming to assess their impact
  - Requires crossing with additional factors
  - May wish to assume two levels for initial design
- Multiple measures (dependent variables) adds significant complexity to determining factors with highest impact
  - Requires evaluation of factors across measures
  - Constructing relationship tables provides insights on measure impacts

Summary

- Exploratory analysis requires an iterative process for prioritizing factors
- Factors can be analyzed across multiple dependent variables (measures)
- Automated tools for DOE and modeling can help to simplify the exploratory analysis process
- Non-materiel factors can be equally important to testing a System of Systems

Max Lorenzo
Joint Test Director
757.638.6079
max.lorenzo@jte.osd.mil

Mr. Tim Beach
Chief, Analysis & Methods & Processes
757.638.6088
timothy.beach@jte.osd.mil

Dr. David Dryer
Lead Systems Engineer
757.638.6091
david.dryer@jte.osd.mil

Mark Fiebrandt
Lead Operations Research Analyst
757.638.6055
mark.fiebrandt@jte.osd.mil

Serving the testing, acquisition, and warfighting communities

Abstract

The Joint Test and Evaluation Methodology (JTEM) project has been collaborating with various government organizations and academia to develop enhanced Design of Experiment (DOE) modeling and analysis approaches for Testing in a Joint Environment (TIJE). This paper discusses the applied research that has been conducted in this area over the past three years, as well as its application to JTEM test events. Discoveries involving enhanced data farming techniques and technology applications have proven to be catalysts for test and evaluation of complex adaptive systems. Hybrid DOE models for large factor test designs (e.g., Fractional Factorial Controlled Sequential Bifurcation, Resolution Five Fractional Factorial, Nearly Orthogonal Latin Hypercube) have demonstrated success in refining robust Joint test spaces. Innovative application of analytical models and methodologies (e.g., Advanced Response Surface Methodology, Classification and Regression Tree) have improved our ability to analyze Critical Capability Issues (CCI) involving multiple responses. Agent based model simulation prototypes (e.g., Tester, MANA, Pythagoras) have been modified and/or developed by our academic and government partners to enable enhanced test design and evaluation of capabilities in a Joint environment. Proof of concept efforts in this collaboration has included International Data Farming Workshop (IDFW) events, where various techniques and tools have been explored for use in Testing in a Joint Environment (TIJE). Key research techniques and selected results are presented in the context of a use case that is based upon JTEM test events.