Army Test and Evaluation Command

ATEC

Merging Operational Realism with DOE Methods in Operational Testing
NDIA Presentation on 13 March 2012

Nancy Dunn, DA Civilian
Chief, Editorial & Statistics/DOE Division,
US Army Evaluation Center
nancy.dunn@us.army.mil
(443)861-9638

Jonathan Fowler, DA Civilian
Chief, Mounted Maneuver Division,
Maneuver-Ground Evaluation Directorate
US Army Evaluation Center
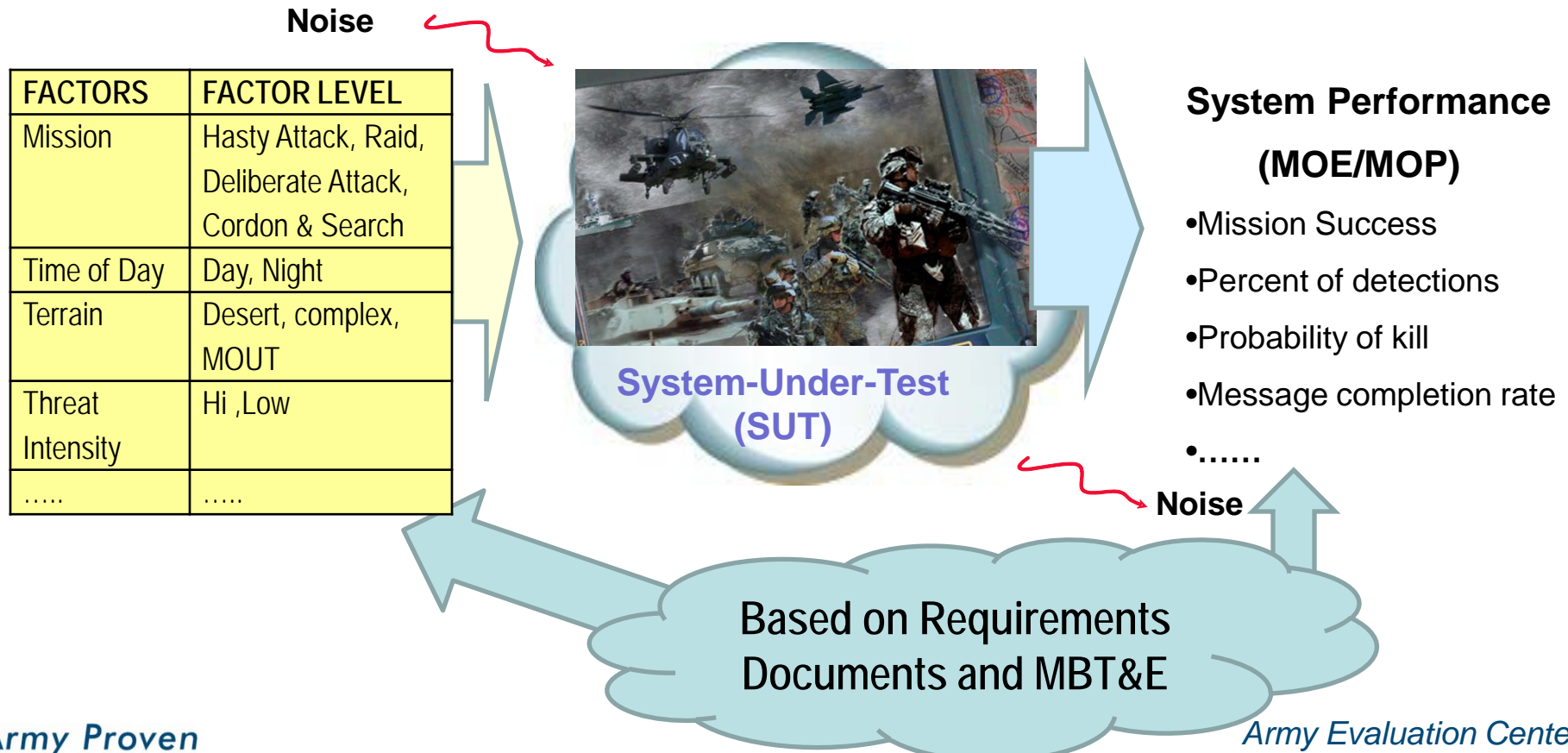jonathan.fowler@us.army.mil
(443)861-9624

Army Proven
Battle Ready

Problem

The desire for realistic "free play" in operational testing traditionally precludes or limits the use of scientific, quantifiable test and analysis techniques.

Operational Testing Fundamental Objective

Determine/demonstrate the impact of new equipment on mission accomplishment of a tactical unit in an operationally realistic environment (a.k.a. factors and conditions)

Noise

FACTORS	FACTOR LEVEL
Mission	Hasty Attack, Raid, Deliberate Attack, Cordon & Search
Time of Day	Day, Night
Terrain	Desert, complex, MOUT
Threat Intensity	Hi ,Low
…..	…..

System-Under-Test (SUT)

System Performance (MOE/MOP)

- Mission Success
- Percent of detections
- Probability of kill
- Message completion rate
- ……

Noise

Based on Requirements Documents and MBT&E

Army Evaluation Center

DOE Fundamental Objective
Scientific Answers to <u>Four</u> Test Event Challenges

Four Challenges faced by each test event

1. ***Which Points?*** A: span/populate the battle-space

2. ***How many?*** A: sufficient samples to control our twin errors – false positives & negatives

3. ***How To Execute?*** A: Randomize and Block runs to exclude effects of the unknown-unknowns

4. ***What Conclusions?*** A: build math-models of input/output relations, quantifying noise, controlling error

DOE effectively addresses all these challenges!

Noise

Inputs
Factors/
Levels
(x's)

DoE

Outputs
Responses
(Y's)

Noise
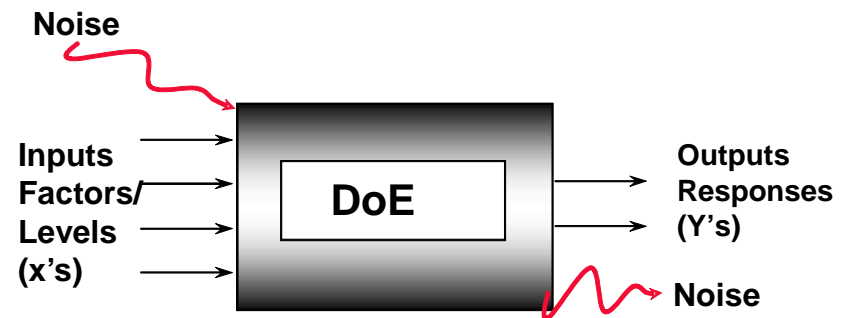
Many design choices: Full factorials, Fractional Factorials, D-Optimal, Split Plot, etc.

Army Proven
Battle Ready

Why The Merger is so Challenging

	Operational Testing	DOE	Paradigm Shift Needed
Goals	Usually very general	Usually very focused	Need more specific goals Specific questions
Responses	More Qualitative, Soldier Surveys, Subject Matter, Expert observations	More Quantitative	Keep qualitative add more quantitative responses, when possible
Factors/ Factor levels	Many factors not controllable. Test Constraints limit testable factor levels.	Controlled by Tester	Accept that OT data will have more variability, due to required BLUFOR "free play"
Multiple Responses with different factors/ levels.	One Large Event, sometimes with specific "excursions".	Each response can have its own Test Design	Create multiple nested Test Designs using DOE within the one larger OT

Army Proven Battle Ready

Army Evaluation Center

Creating Multiple Nested DOEs for OT Based on Specific Test Goals

Noise

Inputs Factors/ Levels (x's) → **Mission Success DOE** → **Outputs Responses (Y's)**

Noise

Noise

Inputs Factors/ Levels (x's) → **MISSION TASK 1 DOE** → **Outputs Responses (Y's)**

Noise

Noise

Inputs Factors/ Levels (x's) → **MISSION TASK 2 DOE** → **Outputs Responses (Y's)**

Noise

Noise

Inputs Factors/ Levels (x's) → **SPECIFIC TASK DOE** → **Outputs Responses (Y's)**

Noise

Requires more controls
- OPFOR
- BLUFOR higher HQ "White Cell"

A Way This Can Done
Example: Stryker ICVV-Scout T&E for Modifications

Remote Weapon
Station for firing
under armor

LRAS3
sensor stored
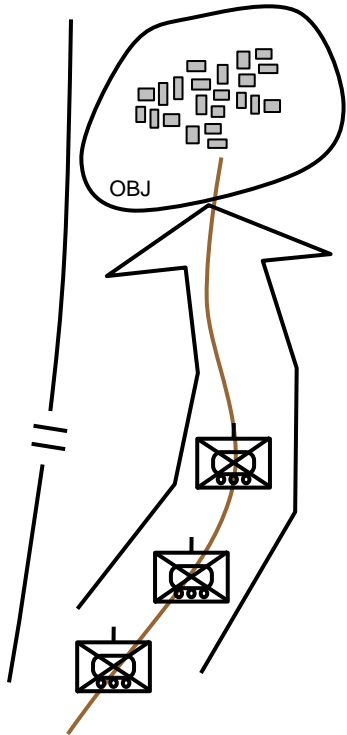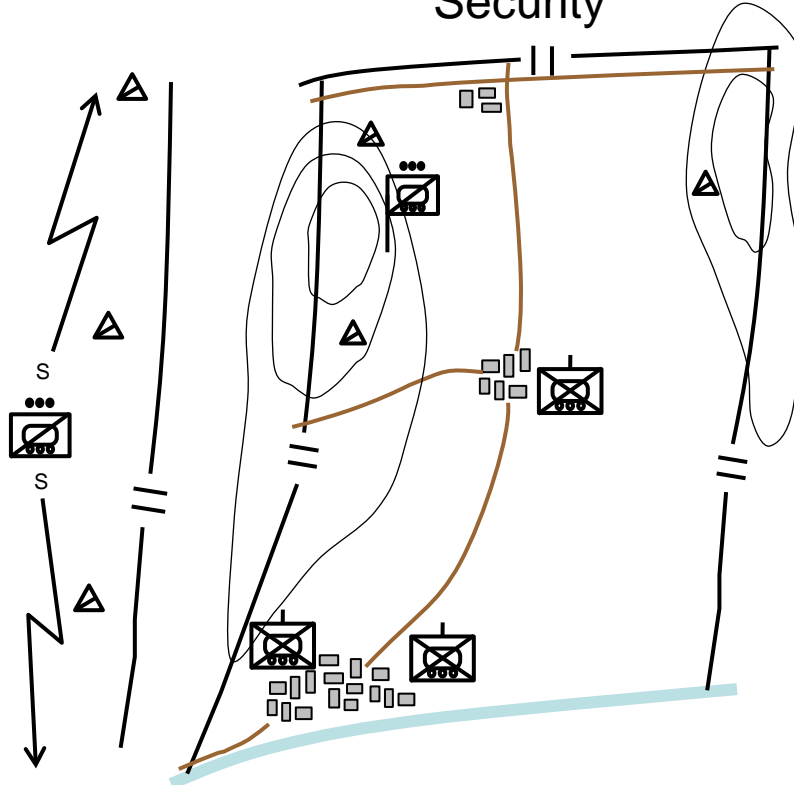inside no longer
mounted on top

Double-V Hull
for additional
survivability

Army Proven
Battle Ready

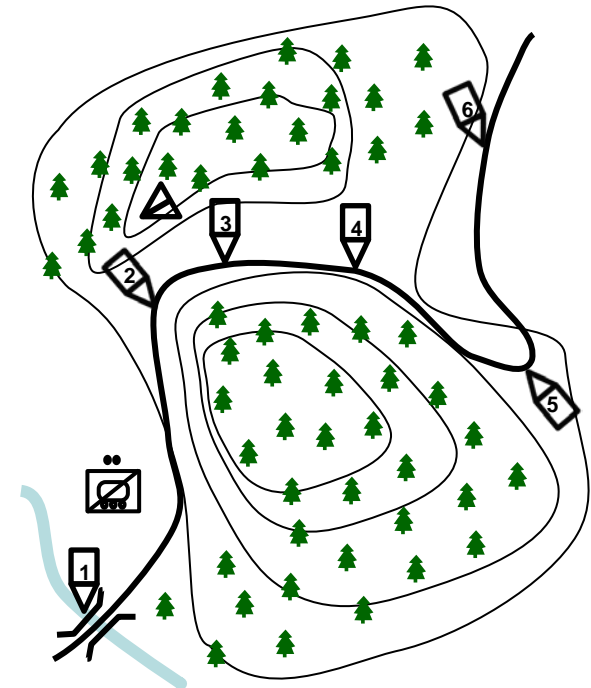# First Review and Understand Platoon/Section Missions

Screen

Area Security

Route Reconnaissance

Focus on a select three of all possible reconnaissance and security missions to give a broad, representative range of context for task accomplishment

Next Review and Understand Section/Team Collective Tasks

- Ingress without detection
- Establish and maintain communication
- **Establish an Observation Post (OP)**
    - Prepare LRAS3 for dismounted operations
- Observe a Named Area of interest (NAI)
- Report on enemy and noncombatant activity
- Conduct local security
- Conduct maintenance
- React to contact
    - **Hand over targets from LRAS3 to Remote Weapon Station (RWS)**
    - Call for fires or air support
    - **Conduct vehicle ingress/egress under duress**
    - **Conduct casualty evacuation**
- **Recover an Observation Post (OP)**
    - Recover and stow LRAS3 from dismounted operation
- Egress/displace without detection

The test is designed around creating opportunities to observe **select tasks** within the mission context.
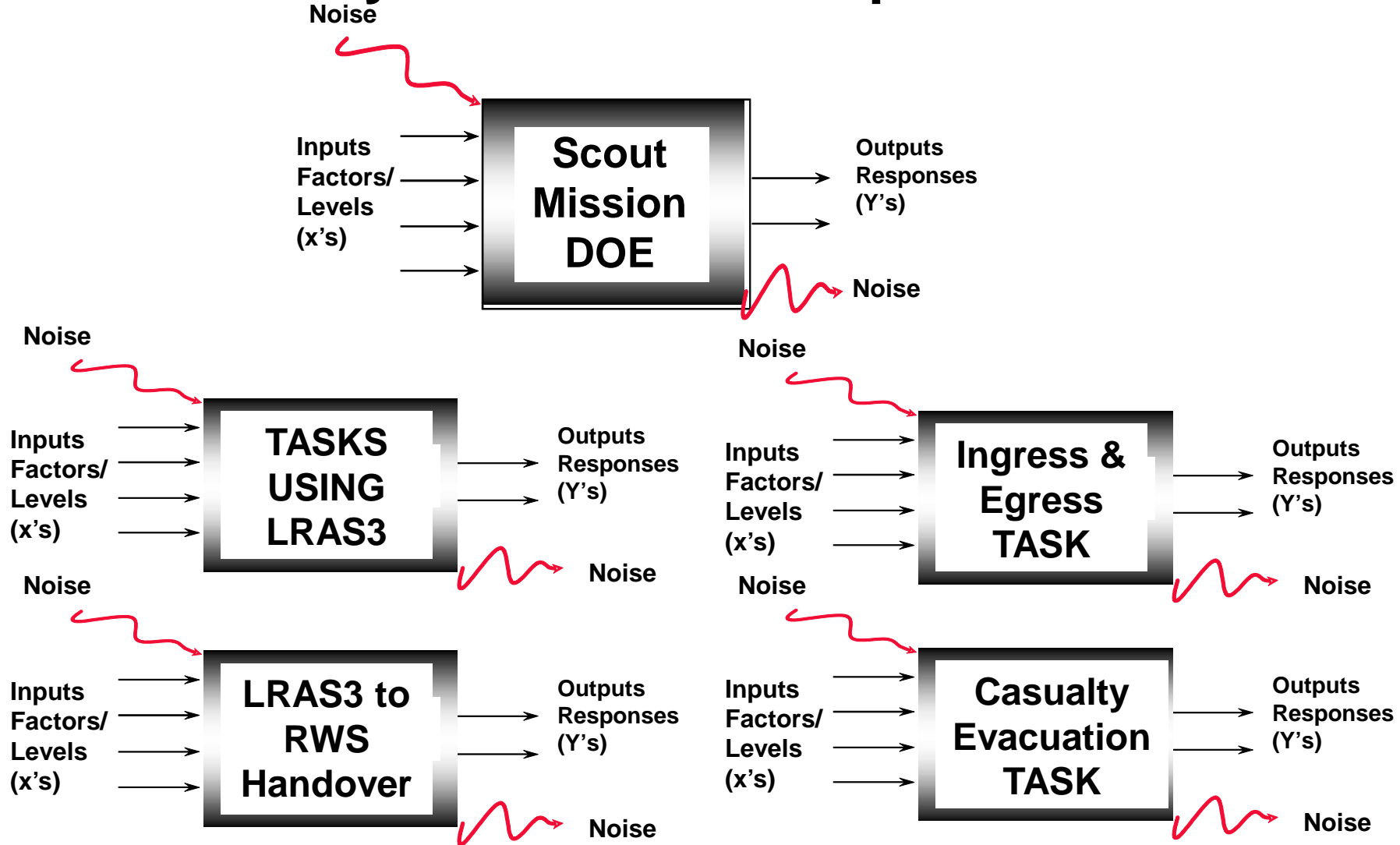
Create Specific Goals
Based on Missions, Collective Task, and Modifications to the Stryker ICVV-Scout

- **Determined that the test will consist of Scout Section Missions.**

- Can a scout section equipped with the ICVV-S accomplish its recon and security missions?
  - Can the scout section dismount, put into operation, shut down, and redeploy the LRAS3 from the ICVV-S?
    - Can the scout section effectively hand over targets from the dismount LRAS3 to the RWS on the ICVV-S?
  - Can a casualty be evacuated from the ICVV-S with the LRAS3 stowed?
  - Can the crew ingress and egress the vehicle quickly enough?

Multiple Nested DOEs
Based on Stryker ICVV-Scout Specific Test Goals

Noise

Inputs Factors/ Levels (x's) → **Scout Mission DOE** → **Outputs Responses (Y's)**

Noise

Noise

Inputs Factors/ Levels (x's) → **TASKS USING LRAS3** → **Outputs Responses (Y's)**

Noise

Noise

Inputs Factors/ Levels (x's) → **Ingress & Egress TASK** → **Outputs Responses (Y's)**

Noise

Noise

Inputs Factors/ Levels (x's) → **LRAS3 to RWS Handover** → **Outputs Responses (Y's)**

Noise

Noise

Inputs Factors/ Levels (x's) → **Casualty Evacuation TASK** → **Outputs Responses (Y's)**

Noise

Army Proven
Battle Ready

Scout Section Missions
(Information for DOE test matrices)

Responses: SME ratings and Unit survey of ability of unit equipped with the ICVV-S to conduct its reconnaissance and security missions		
Factor	**Control**	**Factor Levels**
Type mission	SV*	Area Security, Route Recon, & Screen Missions
Light	SV*	Day, Night
LRAS3	SV*	Deploying LRAS3, Not deploying LRAS3
Test assets & crew	Held constant	Two ICVV-Ss manned by a scout section
LRAS3 Stowage	Held constant	ICVV-S with all equipment required to complete a mission (LRAS3)
Terrain	Held Constant	Primary. Secondary, cross country, trails (All Terrains Covered in Each Mission)

* SV – Systematically varied using DOE principals

Test Design Matrix for
Scout Section Mission Success

Test Design: The power is >= 91% for α =0.1 and S:N ratio of 1.0 for a full factorial completely randomized design with one or two repetitions per cell. We will be able to analyze all main effects and all interactions.

Mission	Day		Night	
	Deploying LRAS3	Without deploying LRAS3	Deploying LRAS3	Without deploying LRAS3
Area Security	2	1	1	2
Route Recon	1	2	2	1
Screen	1	1	1	1
Total	4	4	4	4

Establish an Observation Post
(Information for DOE test matrices)

Responses: **Time to set up OP, SME ratings & Unit survey of ability of scout team to establish an Observation Post.**		
Factor	**Control**	**Factor Levels**
Light	SV	Day, Night
Team	SV	Team #1 & Team #2
LRAS3	SV	Deploying LRAS3, Not deploying LRAS3
Responses: Time (seconds), SME ratings & Unit surveys of ability of scout crew to effectively hand over targets from the dismount LRAS3 to the RWS on the ICVV-S?		
Light	SV	Day, Night
Team	SV	Team #1, #2

* SV – Systematically varied using DOE principals

Test Design Matrix for Observation Post

Test Design: The power is 98% for α =0.1 and S:N ratio of 1.0 for a completely randomized full factorial design with 2 replications.

	Day		Night	
	Deploy LRAS3	No LRAS3 Deployed	Deploy LRAS3	No LRAS3 Deployed
Crew #1 (Crew #2 in over watch)	2	2	2	2
Crew #2 (Crew #1 in over watch)	2	2	2	2
Total	4	4	4	4

Test Design Matrix for Ability to Hand Over Targets From Dismounted LRAS3 to the RWS

Test Design: The power is 78% for α =0.1 and S:N ratio of 1.0 for a completely randomized full factorial design, which can analyze main effects and interactions.

	Day	Night
Crew #1 (Crew#2 in over watch)	2	2
Crew #2 (Crew#1 in over watch)	2	2
Total	4	4

Army Proven
Battle Ready

- **"Free play" applies to the BLUFOR test unit**
  - The unit whose Soldiers will conduct the supporting collective and individual tasks with the new equipment
- **Everything and everyone else in the operational test "box" is an enabler to create opportunities for data collection from the test unit in a realistic environment**
  - **Operational Test Team (working with T&E IPT)**
    - Crafts an operationally realistic environment in the box
    - Applies DOE factors and conditions into a series of vignettes
    - Writes Operations Orders (OPORDs) from BLUFOR higher headquarters
    - Designs event-driven, realistic triggers to elicit desired BLUFOR tasks
  - **BLUFOR higher headquarters "White Cell"**
    - Reinforces constraints & triggers realistically during the test event
  - **OPFOR**
    - Understands role as challenging enabler, not competitor
    - "Free play" within the constraints of each encounter; led by operational test

Free play within realistic constraints in a challenging environment

Shaping the Conditions - Executing

- **The BLUFOR conducts doctrinally realistic missions in "free play"**
  - Conducts unit-level troop leading procedures & develops own operations orders
  - Responds to battlefield stimuli as they see fit, given their training
- **Everyone else monitors the test & adjusts as necessary**
  - **Operational Test Team**
    - Orchestrates all actions in the box to maintain realistic environment and recognizes approaching conditions to execute triggers
    - Directs execution of triggered actions by White Cell and OPFOR
    - Monitors successful collection of data points using matrices as checklists
    - Develops changes to the test schedule in reaction to "missed" data points and builds consensus with Evaluator, User, Developer, OSD for the changes
  - **BLUFOR higher headquarters "White Cell"**
    - Simulates the existence of higher headquarters and adjacent units
    - Sends triggers: Fragmentary Orders (FRAGOs), intel updates, and reports from adjacent units as directed by OT Team
  - **OPFOR**
    - Acts as triggers: executes specific contacts as directed by OT Team
    - Fights hard within the realistic constraints of the scenario and encounter

Summary
Merging Operational Testing with DOE

Determine the BLUFOR Mission Echelon for Test

- **Determine which missions and supporting tasks are most important/most affected by the SUT**

- **Determine the main Goals of the OT to guide the test planning**

- **Determine the number of Test Matrices using DOE that need to be created and how they fit together within the BLUFOR missions**

- **Using the White Cell and the OPFOR determine how each condition in the test matrix can be forced to occur**

- **Plan time at the end of test to run select additional missions for opportunities to collect data points missed due to "free play"**

- **Analyze the responses using techniques that look over the factors**

Focus on the right thing….

Doing things right….

Army Proven
Battle Ready

Conclusion

Test designs based on DOE that include free play for the test unit in operationally realistic environment can be created and executed.

The resulting data allow the use of scientific, quantifiable test and analysis techniques which provide for more meaningful evaluations that better inform senior acquisition decision makers and Warfighter commanders.

Army Proven
Battle Ready

Army Evaluation Center