# TEMP and Test Plan Review Examples

Dr. Catherine Warner

Science Advisor
Director Operational Test and Evaluation (DOT&E)

# DOE Guidance by Milestone

| Milestone Supported | Information Content |
|---|---|
| **A** | Identify responsibilities of T&E WIPT for test design purposes<br>The goal(s) to be addressed at each stage of testing<br>Metrics for each goal/question<br>Initial listing of factors<br>Language for the overall testing strategy, including:<br>• Screening experiments to ensure important factors are considered in operational testing<br>• Sequential experimentation |
| **B** | Identify responsibilities of T&E WIPT for test design purposes<br>The goal(s) to be addressed at each stage of testing<br>Metrics for each goal/question<br>Refined listing of factors and levels<br>Test designs to support resourcing for limited user tests (LUT) and operational assessments (OA)<br>Language for the overall testing strategy, including:<br>• Screening experiments to ensure important factors are considered in operational testing<br>• Sequential experimentation |
| **C** | Identify responsibilities of T&E WIPT for test design purposes<br>The goal(s) to be addressed at each stage of testing, focusing on IOT&E<br>Metrics for each goal/question<br>Refined listing of factors and levels, based on prior testing and the operational mission.<br>Details on how the factors and levels will be varied and controlled during each stage of testing<br>Complete test designs to support resourcing for IOT&E<br>Language for the overall testing strategy, including:<br>• How previous knowledge is being used to inform IOT&E test planning.<br>• Analysis plans to support power calculations |

# Specific example: Q-53
# Army Counter-Fire Radar Test Plan
# (Test completed – report in progress)

- **Response variables clearly identified: circular error probable, point of origin error, point of impact error , probability to detect, probability to classify**
  - Includes mix of continuous and binary responses
- **Integrated testing –**
  - Shot performance characterization data will be obtained from DT.
  - Ten IOT&E missions will be randomly selected to confirm accuracy of DT missions
    - » Could we do better than random?
- **Clear identification of factors and levels**
  - Table 1-2 and 1-3
- **Problems discovered in DT prohibited the OT use of DT data**
  - No test design for OT was acceptable because of large OT
    - » 278 Fire missions, 3649 projectiles fired
    - » Allowed for characterization of performance despite no DOE basis for the design

# Specific example: SDB II
# Air Force/Navy Air Delivered Weapon TEMP
# (In OSD routing for approval)

- **Operational evaluation is based on modeling and 55 flight test weapon releases**
  - Test design includes 50 of 55 shots to characterize performance across three attack modes (normal*, laser, coordinate)
  - An additional 28 Government Confidence Test (GCT) weapon releases will be conducted before OT for normal attack mode only; some OT shots may be reduced if GCT eliminates factors that do not affect performance
  - Table 3-6 identifies factors for DT, GCT, and OT

- **Response variables**
  - Primary response variables are Probability of Single Shot Kill and miss distance

- **Overarching Design**
  - Primary factors airspeed, altitude, range represent launch acceptability region (LAR); GCT will define the levels; will be incorporated into all 55 shots
  - Sub-designs for each of the attack modes; laser and coordinate are legacy modes so fewer shots are dedicated for this
  - Power shown for overarching and subdesigns



*** Normal attack mode is a new capability developed for moving targets**

# Specific example: Triton
# Navy UAV Integrated Evaluation Framework (in-progress)

- **IEF serves as a source document for OT description in TEMP and basis for Test Plans**
  - Triton IEF identifies six mission oriented COIs and associated tasks
  - Response variables (RV) are identified for each COI, continuous and binary
  - Only RV that drive test design are associated with COIs, other measures are listed but not assigned to a COI (e.g., mission planning, cybersecurity)

- **Mission tasks are organized into Vignettes**
  - Test design for each vignette vice per COI or RV
  - Tables including RV, Factors, Levels for each vignette are provided that adequately span the operational envelope and can be reasonable controlled in OT
  - Split plot designs are used with whole plots for hard to change factors (e.g., light, altitude); however, test execution most likely will result in a more randomized test design

- **Power of tests documented (at 80% confidence level)**
  - Per RV for main effects, 2-factor, and 3-factor interactions
  - Standard deviation and operationally significant effect sizes based on historical performance of similar systems