Marine Corps Operational Test and Evaluation Activity

Operational Test & Evaluation Manual

Third Edition

This page intentionally left blank for two-sided printing.

UNITED STATES MARINE CORPS
MARINE CORPS OPERATIONAL TEST AND EVALUATION ACTIVITY
2032 BARNETT AVENUE
QUANTICO, VIRGINIA 22134-5014

3980
Dir/0058
22 Feb 13

From: Director, Marine Corps Operational Test and Evaluation Activity (MCOTEA)
To: MCOTEA All Hands

Subj: MCOTEA OPERATIONAL TEST AND EVALUATION MANUAL THIRD EDITION

1. The *MCOTEA Operational Test and Evaluation Manual* presents a process rooted in both the scientific method and Marine Corps operations. The Manual combines elements of Marine Corps missions and tasks with systems engineering, decision analysis, and design of experiments providing a process supporting all test and evaluation (T&E) activities MCOTEA performs. The execution of the steps outlined in this Manual must result in a balance between objective data gathering with operational judgment rooted in realistic scenarios and hard-learned lessons through combat experiences.

2. The overall T&E, as established in the first edition continuing throughout the second edition, has not changed apart from adjustments and clarifications based on lessons learned. The third edition of the Manual, which supersedes the second edition, incorporates policy revisions to Department, Component, and Service-level instructions and orders and to internal MCOTEA policy. Changes to the third edition also include a new in-depth example running throughout the six-steps of the MCOTEA process. Division and Staff responsibilities have been updated within and external to the MCOTEA six-step process reflecting organizational structure changes rooted in the establishment of the technical skills necessary to support operations analysis for Marine Corps capabilities. The third edition brings about increased guidance on test execution and an update to the MCOTEA T&E templates. Finally, the third edition splits into two volumes. Volume I is the core of the MCOTEA process, while Volume II presents the unique elements incorporated in the second edition like Reliability, Availability, and Maintainability, Live Fire, and Modeling and Simulation.

3. This Manual is a living document and will be updated regularly with additional material. All hands are encouraged to submit comments or recommendations to the Scientific Advisor for the improvement of this Manual.

4. Use of this Manual for performing Marine Corps operational test and evaluation is mandatory and effective immediately.

5. Thank you for all your continued professionalism and cooperation.

A. J. PASAGIAN

MCOTEA Mission

MCOTEA provides operational testing and evaluation for the Marine Corps and conducts additional testing and evaluation as required to support the Marine Corps mission to man, train, equip, and sustain a force in readiness.

MCOTEA Vision

MCOTEA is the Marine Corps leader in all aspects of realistic operational test and evaluation of materiel system capabilities throughout a materiel system's life cycle. Our highly trained, professionals are a voice for the Operating Force Marine, enabling informed decision-making, and ensuring always that our test reports accurately and objectively describe what we know and don't know about the Operational Effectiveness, Suitability, and Survivability of the materiel solution we evaluate. MCOTEA is a source for objectivity in the Marine Corps and, where appropriate, DOD's acquisition process. Our expertise, professionalism, and integrity makes us a sought-after partner within the DOD acquisition community.

Table of Contents

Chapter 1. Organization 1-2
General Philosophy 1-2
Executive Office 1-2
Director 1-2
Deputy Director 1-2
Scientific Advisor 1-3
Chief of Staff 1-3
Divisions 1-3
Combat Service Support Division 1-4
Cyber Division 1-4
Expeditionary Division 1-4
Live Fire Division 1-4
MAGTF C4ISR 1-4
Operational Test and Analysis Division 1-4
Staff Functions 1-5
Business Management 1-5
S-1: Human Capital and Administration 1-5
S-3: Operations 1-6
S-4: Logistics 1-6
Fiscal 1-7
Chapter 2. Background & Paradigm 2-2
MCOTEA's Mandate and Purpose 2-2
MCOTEA's Working Relationships with Other Organizations 2-3
Working Partners 2-3
Deputy Commandant for Combat Development and Integration 2-3
Marine Corps Systems Command 2-3
Program Executive Officer Land Systems 2-3
Oversight/Non-Chain of Command 2-3
Director, Operational Test and Evaluation 2-3
Assistant Secretary of the Navy (Research, Development, and Acquisition) 2-4
Gate Reviews 2-4
Fellow Operational Test Agencies 2-5
Joint Interoperability Test Command 2-5

Communication/Information Sharing 2-5
Navy Enterprise T&E Board of Directors 2-5
N084 2-6
OSD Test Investment Coordinating Committee 2-6
Test Resources Management Center 2-6
Acquisition Life Cycle Overview 2-6
ACAT Designation 2-6
Evolutionary Acquisition 2-6
Incremental Testing Requirements 2-7
Test and Evaluation Paradigm 2-7
Test Relationship to Evaluation 2-8
The Evaluation Continuum 2-9
Advantages of MCOTEA's Early Involvement 2-9
Continuous Evaluation 2-9
Early Identification of Deficiencies 2-10
Collaboration Along the Acquisition Time Line 2-11
Pre-Milestone A 2-11
Milestone A to Milestone B 2-12
Post Milestone B 2-13
Obtaining and Using Developmental Test Data 2-13
MCOTEA Test Team Billets and Best Practices 2-14
The Operational Test 2-14
Before Testing 2-14
During Testing 2-15
After Testing 2-15
Test Manager 2-15
Before Testing 2-15
During Testing 2-15
After Testing 2-15
Operations Research/Systems Analyst 2-15
Before Testing 2-16
During Testing 2-16
After Testing 2-16
Mathematical Statistician 2-16
Before Testing 2-16
During Testing 2-16
After Testing 2-16

Data Manager 2-16
Before Testing 2-16
During Testing 2-16
After Testing 2-17
Supplementary Team Members 2-17
MCOTEA's 6-Step Test and Evaluation Process 2-17
Step 1. System Evaluation Plan and FD/SC Charter 2-17
Step 2. Test Concept and TEMP Input 2-17
Step 3. Test Planning 2-17
Step 4. Operational Test Execution 2-17
Step 5. Operational Test Reporting 2-18
Step 6. System Evaluation and Reporting 2-18
Records Management and Lessons Learned 2-18
Types of MCOTEA Tests 2-18
Operational Testing 2-18
Initial Operational Test and Evaluation 2-19
Follow-on Operational Test and Evaluation 2-19
Multi-Service Operational Test and Evaluation 2-19
Marine Corps Lead Service 2-19
Other Service OTA Lead 2-20
MCOTEA Assessments 2-20
Top-Level MCOTEA Functions 2-21
Chapter 3-0. The 6-Step Process 3-2
Entry of New Work into MCOTEA 3-2
Program SharePoint Support 3-2
Staff Responsibilities 3-3
Plan-Test-Report 3-4
Integrated Testing Within the 6-Step Process 3-4
MCOTEA Assessments 3-4
System Assessments 3-5
Quick Reaction Assessment 3-6
AAPs and ACAT IV(M) 3-6
Intermediate Assessments 3-6
DT Observation 3-6
Operational Assessment 3-7

Early Operational Assessment 3-8

Chapter 3-1. Step 1: System Evaluation Plan 3-1-2

Evaluation Purpose 3-1-2

Evaluation Paradigm:

The Importance and Benefits of Continuous Evaluation 3-1-2

Review Program Documentation 3-1-3

System Evaluation Plan (SEP) 3-1-3

Background 3-1-3

Scope 3-1-4

Objectives 3-1-4

Operational Task Analysis 3-1-5

Identify Missions 3-1-5

Identify Tasks 3-1-5

Identify Lower-Level Subtasks 3-1-6

Develop Evaluation Questions 3-1-8

Evaluation Questions 3-1-8

Characteristics of Evaluation Questions 3-1-8

Top-Down & Mission-Based 3-1-9

OE/OS/OSur Interrelationship 3-1-9

Operational Effectiveness 3-1-9

Operational Suitability 3-1-9

Operational Survivability 3-1-10

Critical Operational Issues 3-1-10

Issues 3-1-10

Tasks and Subtasks 3-1-10

Suitability and Survivability 3-1-11

Determine Level of M&S Support Required 3-1-11

Construct Evaluation Standards and Measures 3-1-11

Developing Standards 3-1-11

Attribute Variations 3-1-11

Mapping Attributes to the Evaluation Framework 3-1-13

Attributes Mapping Matrix 3-1-13

Establish Standards for Evaluation Questions 3-1-13

Standards for Performance and Conditions 3-1-16

Finalize Evaluation Questions 3-1-17

Types of Measures 3-1-17

Properties of Measures 3-1-17

Measures of Effectiveness 3-1-18

Measures of Performance 3-1-18
Measures of Suitability 3-1-18
Measures of Survivability 3-1-18
Preferential Measures 3-1-18
Measure Scales 3-1-18
Measure Alignment with Objectives 3-1-18
Measure Data Type 3-1-19
Establishing Dominant Measures 3-1-19
Properties of the Evaluation Model 3-1-20
Screening Criteria 3-1-20
Aggregation Method 3-1-20
Properties Necessary for Aggregation 3-1-20
Collectively Exhaustive 3-1-21
Mutual Exclusivity 3-1-21
Evaluation Framework Operability 3-1-21
Building an Evaluation Model 3-1-22
Identifying Screening Criteria 3-1-22
Constructing the Analytic Model 3-1-23
Constructing an Example Decision Model 3-1-24
Weighting Elements of the Decision Model 3-1-27
Combining Screening Criteria, Analytic Model, and Decision Model 3-1-28

Additional SEP Elements 3-1-28
Preparation 3-1-28

Chapter 3-2. Step 2: Test Concept, TEMP Input, and Charter Development 3-2-2

Technical Information about Test Concept Development 3-2-2
Identify the Test Objective 3-2-2
Define the System 3-2-2
Define the Trial 3-2-2
Identify Cause-Effect Relationships 3-2-3
Factors and Cause and Effect Relationships 3-2-3
Link Inputs, Process Flows, and Outputs to Cause-Effect Relationships 3-2-4
Select Sample Size, Design Type, and Analysis Method 3-2-5
Test Limitations 3-2-5
Integrated Data Requirements List 3-2-6
Preparing TEMP Input 3-2-6
TEMP Background and Structure 3-2-6
Summary of TEMP Part III Development 3-2-8
Preparing TEMP Part IV Inputs 3-2-8

Identifying Resource Requirements 3-2-8
Identify Test Locations 3-2-8
Estimate Time Requirements 3-2-8
Identify M & S, Cyber, and LFT&E Needs and Required Resources 3-2-9
Identify Key Resources 3-2-9
Estimate Costs 3-2-9
Inputs and the Operational Test Plan 3-2-9
Integrated Testing 3-3-2
Writing DT Observation Plans 3-3-2
Chapter 3-3. Step 3: Test Planning 3-3-2
Observing Test Events 3-3-3
DT Observation Reporting 3-3-6
DT Report Review Process and IAR Preparation 3-3-6
Early Operational Test Planning Activities 3-3-6
Human Research Protection Program 3-3-7
Creating the Feasibility of Support Message 3-3-7
Test Planning Timeline 3-3-8
Test Planning Process 3-3-8
Check Lessons Learned Database 3-3-8
Writing the Test Plan 3-3-8
Initial Sections of the Test Plan 3-3-8
Refining the Schedule 3-3-8
Operational Test Readiness Board/Review Process 3-3-9
Conducting the Pre-OTRR 3-3-9
Conducting the OTRB 3-3-10
Conducting the OTRR 3-3-10
Coordination of Personnel 3-3-11
Marine Officer in Charge and Test Unit 3-3-11
Data Collectors 3-3-11
Contractors 3-3-11
Active Duty Marines 3-3-11
Reservists 3-3-12
Detailed Test Planning 3-3-12
Insert Sample Size and Test Design 3-3-13
Finalize Data Requirements 3-3-13
Data Requirements for Measures 3-3-13

Determine Data Requirements for Analysis 3-3-13
Develop Data Collection Methods 3-3-13
Automated Data Collection 3-3-13
Manual Data Collection 3-3-14
Develop Data Reduction Methods 3-3-16
Adding Details to Test Trials 3-3-16
Test Plan Annexes 3-3-16
Requirements for Other-Service Assets 3-3-20
Navy Assets 3-3-20
Army Air Assets 3-3-20
Marine Air Assets 3-3-20
Air Force Assets 3-3-20
Equipment Requirements and Test Site Coordination 3-3-21
Identifying Required Facilities and Logistics Support 3-3-21
Data Collection Forms 3-3-22
Types of Forms 3-3-22
Quantitative Forms 3-3-22
Qualitative Forms 3-3-22
Verification Forms 3-3-22
Survey Questions 3-3-22
Data Collector Training 3-3-22
Environmental Considerations 3-3-23
Data Collection Based on Data Requirements 3-3-23
Building a Data Repository 3-3-23
Maintaining Data Integrity and Security 3-3-24
Data Collection Verification and Validation 3-3-24
Test Site Visit 3-3-24
Equipment Check 3-3-24
Instrumentation 3-3-24
Transporting the Test Team and Test Equipment 3-3-25
Travel Orders 3-3-25
Transporting Test Equipment 3-3-25
Site-Specific Restrictions 3-3-25
Marine Corps Equipment 3-3-25

Test Funding 3-3-26

Chapter 3-4. Step 4: Operational Test Execution 3-4-2

Pretest Activities 3-4-2

New Equipment Training 3-4-2

DC Training 3-4-2

Test Execution 3-4-2

Pilot Test 3-4-2

Operator Learning Curve During Pilot Test 3-4-3

Pilot Test Data Collection 3-4-4

Record Test 3-4-4

Hold a Morning Brief 3-4-4

Manage Test Deviations 3-4-4

Manage Site Visitors 3-4-5

Validate Data Daily 3-4-7

Transfer Data to Analysts 3-4-7

Conduct Daily Hot Wash 3-4-7

Write Daily Situation Reports 3-4-7

Release of Test Data 3-4-7

Data Reduction 3-4-7

Posttest Activities 3-4-7

FD/SC Scoring Conference 3-4-7

Test Site Closeout 3-4-8

Data Review 3-4-9

Chapter 3-5. Step 5: Operational Test Data Reporting 3-5-2

Background 3-5-2

Test Data Report 3-5-2

Test Data Report Examples 3-5-2

Chapter 3-6. Step 6: System Evaluation and Reporting 3-6-2

Purpose of Evaluation 3-6-2

Evaluation and Reporting Requirements for OT 3-6-2

Determine Threshold Satisfaction 3-6-2

Determine Operational Effectiveness 3-6-2

Determine Operational Suitability 3-6-2

Determine Operational Survivability 3-6-3

Assess Impact to Combat Operations 3-6-3

Major System Deficiencies 3-6-3

Operational Deficiencies 3-6-3
Types of Evaluation Reports 3-6-3
Evaluation Process Basics 3-6-4
Populations and Samples 3-6-4
Parameters and Statistics 3-6-4
Considering Uncertainty with Confidence Bounds 3-6-4
Evaluation at Early Stages of System Development 3-6-4
Evaluation at Later Stages of System Development 3-6-5

Evaluating COIs 3-6-8
Example of Analysis of COI Results 3-6-8
OE/OS/OSur Conclusions 3-6-13
Determining OE 3-6-13
Transferring Decision Model Answers to the OER 3-6-14
Determining OS and OSur 3-6-14
Performing Sensitivity Analysis 3-6-14
Transparency of Evaluation 3-6-14
Why Depict? 3-6-15
Exploring the Data 3-6-15
How Often to Depict? 3-6-15
Which Depictions to Use? 3-6-15
Types of Depiction 3-6-15
In General 3-6-16
Working with Graphics in MCOTEA Documents 3-6-17
System Evaluation Plan 3-6-17
Formulas 3-6-17
Mission Capability Level Value Function 3-6-17
Test Plans 3-6-17
Reports 3-6-17
Chapter 4. Documentation 4-2
MCOTEA's Standardized Approach to Documentation 4-2
Early Test and Evaluation Planning 4-2
DT Observation 4-2
Operational Test Plans 4-2

Operational Test Reports 4-2
Joint or Multi-Service Test Events 4-3
M&S Accreditation Process 4-3
Base Templates 4-3
Using Citations in Text 4-3
CRB Document Approval Process 4-4
Document Change Process 4-4
General Guidance for Using Templates 4-4
Cover Pages 4-4
Executive Summaries 4-4
Graphics 4-4
Annexes 4-4
Editorial References 4-6
Annex A: Marking Cover Pages 4-7
Distribution Statements 4-7
Method 4-7
Proper Marking of Documents 4-7
Defense Technical Information Center 4-7
DTIC Submission Process 4-7
Cover Page 4-7
Volume II, Chapter 1. Live Fire Test and Evaluation VII-1-10
Background VII-1-10
Objective VII-1-10
Requirement for LFT&E VII-1-11
Vulnerability LFT&E VII-1-12
Penetration VII-1-12
Kill VII-1-13
Lethality LFT&E VII-1-13
LFT&E Management VII-1-13
MCOTEA Vulnerability Process VII-1-15
Building Block Approach VII-1-15
Armor Validation VII-1-16
Armor Characterization VII-1-16
Armor Exploitation VII-1-16
Ballistic Hull and Turret VII-1-16

Component Ballistic Testing VII-1-16
System-Level Testing VII-1-17
Controlled Damage Testing VII-1-17
Full-Up System-Level Testing VII-1-17
Marine Corps Lethality Process VII-1-17
Building Block Approach VII-1-17
Qualification Testing VII-1-17
Munition Terminal Effects Characterization VII-1-17
System-Level Testing VII-1-18
End-to-End FUSL Testing VII-1-18
LFT&E Key Elements VII-1-18
Scope of LFT&E VII-1-18
LFT&E Strategy VII-1-18
Live Fire System Evaluation Plan VII-1-20
Threats VII-1-20
Shotlines VII-1-21
LFT&E Reporting VII-1-21

Volume II, Chapter 2. Reliability, Availability, and Maintainability (RAM) VII-2-2

Availability (A) VII-2-2
Material Availability (A_m) VII-2-2
Achieved Availability (A_a) VII-2-4
Inherent Availability (A_i) VII-2-4
Reliability VII-2-4
Maintainability VII-2-6
Maintenance VII-2-6
Preventive (Scheduled) Maintenance VII-2-6
Corrective (Unscheduled) Maintenance VII-2-6
Operational Mission Failures VII-2-7
Crew-Correctable Maintenance Action VII-2-7
Essential Maintenance Actions VII-2-7
Unscheduled Maintenance Actions VII-2-7
Maintenance Metrics VII-2-7
Mean Time to Repair VII-2-8
Maximum Time to Repair VII-2-8
Maintenance Ratio VII-2-8
Administrative and Logistics Down Time VII-2-8
Linking Metrics to

Time-Based Models VII-2-8
Operational Mode Summary and Mission Profile VII-2-11
Operational Mode Summary VII-2-11
Mission Profile VII-2-12
Linking Time-Based Models to Failure Definition/Scoring Criteria VII-2-13
Incident Classification VII-2-14
No Test VII-2-14
Crew Correctable Maintenance Action (Optional) VII-2-14
Operational Mission Failure VII-2-15
Essential Maintenance Action VII-2-15
Unscheduled Maintenance Actions VII-2-15
Incident Chargeability VII-2-15
Hazard Severity Assessment VII-2-16
Hazard Probability Assessment VII-2-16

Chapter VII-3. Modeling and Simulation and the Verification, Validation, and Accreditation Process VII-3-2

Deciding to Use M&S VII-3-3
Expanded Definitions of M&S VII-3-3
Using M&S in the Test Process VII-3-4
Requirement for VV&A VII-3-4
Developing Specific Intended Uses and Accreditation Criteria VII-3-5
Locating the Right M&S VII-3-5
Existing Models VII-3-5
New/Modified Models VII-3-6
M&S Funding and Schedule Requirements VII-3-6
Verifying, Validating, and Accrediting M&S VII-3-6
Accreditation Agent VII-3-8
Other Important V&V Roles VII-3-8
VV&A Process in Total VII-3-9
The Accreditation Process VII-3-9
Accreditation Schedule VII-3-9
Overview of the Accreditation Assessment VII-3-9
Simulation Control Panel VII-3-12
Overall Assessment VII-3-15
VV&A Documentation Process VII-3-15

Acronyms Acro-1

Bibliography Ref-1

1 Organization

Executive Office
Divisions
Staff Functions

Chapter 1. Organization

General Philosophy

The Marine Corps Operational Test and Evaluation Activity (MCOTEA) is organized into an Executive Office, Divisions, and Staff sections (figure 1-1). These components support the Director in accomplishing all of the functions assigned to MCOTEA to ensure realistic, rigorous, independent, and unbiased Operational Test and Evaluation (OT&E) for the Marine Corps.

This section briefly describes each component of the MCOTEA organization. These descriptions are an overview and do not attempt to cover all of the functions associated with each component.

Executive Office

Director

The Director, MCOTEA, with support from the divisions and staff, ensures the effective performance of all the top-level functions discussed in Chapter 2 and the following additional responsibilities (Secretary of the Navy 2011):

- Director, MCOTEA, shall prepare the operational test content, with the exception of Live-Fire Test and Evaluation (LFT&E), and a listing of resources required to execute operational test for input into the Test and Evaluation Master Plan (TEMP).
- Director, MCOTEA, shall request, from the office of Assistant Commandant of the Marine Corps (ACMC), the assignment of a test director (TD) for acquisition category (ACAT) I and certain ACAT II programs and shall coordinate with the Marine operating forces and other commands in matters related to OT&E by publishing a test planning document (TPD).
- Director, MCOTEA, shall manage those joint Office of the Secretary of Defense (OSD)-directed multi-service OT&Es for which the Marine Corps is tasked and coordinate Marine Corps support for other military Services' OT&Es.
- Director, MCOTEA, shall prepare and provide directly to the ACMC, within 90 days (or as stipulated in the TEMP) after completion of OT&E, an Operational test Agency (OTA) evaluation report for the system under test.
- Director, MCOTEA, shall advise the ACMC on OT&E matters. When significant limitations are identified during system evaluation, the Director, MCOTEA, shall advise the milestone decision authority (MDA) of risk associated in the procurement decision.
- Director, MCOTEA, shall maintain direct liaison with OSD's Director, Operational Test and Evaluation (DOT&E), the Marine operating forces for OT&E matters, and other military activities and commands, as required.
- Director, MCOTEA shall represent the Marine Corps in all multi-service OT&E matters.
- Director, MCOTEA shall be the primary interface with Joint Interoperability Test Command (JITC) on joint interoperability testing conducted during OT.
- For USMC programs not required by statute to conduct LFT&E, but where LFT&E is appropriate, the Director, MCOTEA shall concur with the LFT&E strategy as approved by the MDA in the Test and Evaluation Strategy (TES) or TEMP.

Deputy Director

The Deputy Director, MCOTEA, assists the Director in performing his responsibilities and directs the staff in supporting the Director and executing MCOTEA functions. In addition, the Deputy assists in determining the future direction of and vision for the Activity and represents MCOTEA by interfacing with external organizations. Finally, the Deputy is the organization's security officer and

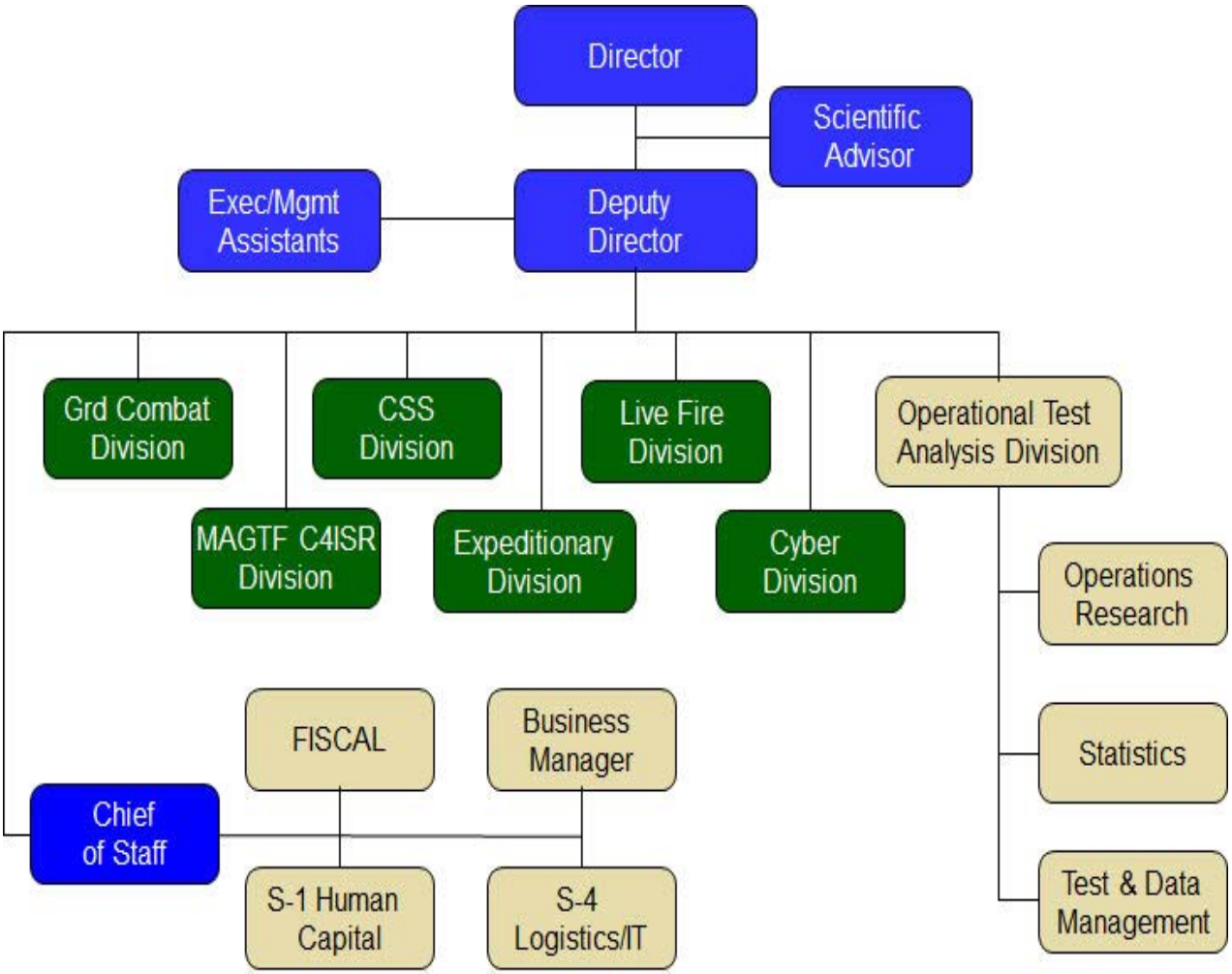


Figure 1-1.
MCOTEA's Organization

handles all security-related issues.

Scientific Advisor

The Scientific Advisor (SA) provides technical advice on evaluation strategies, test planning, and test execution and provides quality assurance for MCOTEA products. The SA tracks Department of Defense (DOD) and Department of the Navy (DON) policies and interprets their effect on MCOTEA. In addition, the SA assists the Director and the Deputy in determining MCOTEA's future direction. The SA investigates new testing and evaluation methodologies and instrumentation of use to MCOTEA. The SA also interfaces with external organizations in various forums.

Finally, the SA leads MCOTEA's efforts in process improvement and recommends any changes to the Director.

Chief of Staff

The Chief of Staff (COS) serves as the overall staff lead under the cognizance of the Deputy Director. The COS ensures that the staff executes the Director's guidance in a coordinated and integrated manner. The COS also ensures timely, efficient, and effective coordination of staff efforts in support of the divisions. The COS is responsible for implementing the MCOTEA Safety Program.

Divisions

Testing and evaluation is accomplished in MCOTEA's functionally aligned divisions. The divisions ensure that sufficient and qualified personnel are assigned to each test program and that MCOTEA testing is well planned, well coordinated, and has

sufficient materiel support. Each division is run by a Division Head, who may be assigned as a technical authority providing oversight of other Government agency personnel or contractors supporting their programs.

The divisions provide services to the Marine Corps, Multi-Service, and Joint Service organizations and perform various levels of testing depending on system complexities and the decision maker's needs. The divisions work in close coordination with the lead OTA for programs requiring MOT&E. The divisions are supported by the Operational Test and Analysis Division (OTAD).

Combat Service Support Division

Combat Service Support Division (CSSD) is responsible for monitoring and testing programs associated with personnel combat survivability, motor transport, and medical assets (Combat Service Support Section); combat engineering equipment and robotics (Combat Engineer Section); and Chemical, Biological, Radiological, and Nuclear (CBRN) equipment.

Cyber Division

The Cyber Division (CD) is responsible for evaluating all programs entering MCOTEA to ensure that they have an integrated, realistic Cyber Test and Evaluation program that provides operationally relevant data to determine Operational Survivability (OSur). The Cyber Division is also responsible for conducting Information Assurance (IA) and Interoperability (IOP) Assessments of the Marine Expeditionary Forces under the DOT&E IA & IOP Program.

Expeditionary Division

Expeditionary Division (ED) is responsible for monitoring and testing programs associated with USMC amphibious vehicles (Amphibious Vehicle Section); Navy ship and ship-to-shore connector programs (Naval Section); and expeditionary initiatives (Expeditionary Initiatives Section).

Ground Combat Division

Ground Combat Division (GCD) is responsible for monitoring and testing programs associated with infantry weapon systems, infantry combat equipment, and non-lethal systems (Infantry Section); artillery and artillery support equipment (Fires Section); and combat vehicle and anti-armor systems (Combat Vehicles Section).

Live Fire Division

The Live Fire Division (LFD) is responsible for evaluating all programs entering MCOTEA to ensure that they have an integrated, realistic Live Fire Test and Evaluation program where applicable, to include the assessment of ballistic and lethality requirements for non-LFT&E programs.

MAGTF C4ISR

The Marine Air-Ground Task Force (MAGTF) Command, Control, Communications, Computers, Intelligence, Surveillance , and Reconnaissance Division (MC4ISRD) is responsible for monitoring and testing programs associated with Marine Corps information, command, control, and intelligence systems (C4ISR Section); command and control systems (MAGTF Command and Control (C2) Section); information systems, communications and networking systems, and simulators (Information Systems Section).

Operational Test and Analysis Division

The Operational Test and Analysis Division (OTAD) provides support directly to each division's Operational Test Project Officers (OTPO). This support includes decision science capabilities in evaluation strategy, analytical test design, test concept

development, test and data management, and evaluation reporting. The OTAD is also responsible for providing specialty services, including:

- Design of Experiments (DOE) Development
- Implementation of Modeling & Simulation (M&S) and Accreditation for MCOTEA use in evaluations
- Development of techniques for determining Reliability, Availability, and Maintainability (RAM)
- Consideration of Human Factors in test planning and system evaluation
- Development, auditing, improvement, and enforcement of MCOTEA processes

Staff Functions

Staff Sections support the Director in executing all MCOTEA functions. In particular, the staff supports the divisions by ensuring that testing and evaluation is well planned and coordinated, adequately staffed, and has sufficient materiel support. The staff also helps ensure that internal processes:

- Are efficient and consistent with higher-level directives
- Contribute to the delivery of high-quality products
- Support effective communication and coordination with external agencies

Staff Section numbering and functions reflect common MAGTF usage where possible to facilitate communication with Marine Corps organizations. Each Staff Section is run by a Staff Lead, who may be assigned as a technical authority providing oversight of other Government agency personnel or contractors supporting program tasks.

Business Management

The Business Management section coordinates business activities and processes across MCOTEA. The section consists of a Business Manager and a Business Administration Specialist, with augmentation by technical experts as required. The Business Management Section focuses on consistency and efficiencies in sustained work efforts as well as achievement of MCOTEA initiatives. Section interfaces are directed both internally and externally in support of the Activity. The Business Management Section helps to maintain an organization responsive to changes in the acquisition environment as well as to internal and external leadership direction. Specific activities include:

- Initiatives as directed by the Director
- Supplemental resourcing to support the Activity through Services Contract development and execution
- Supplemental resourcing to support the Activity through other Government agency resourcing
- Coordination of supply and other contracting efforts
- Cost estimation process improvement
- Development and application of Activity-level Quality Control and Process Improvement Metrics and Configuration Management and Change Control processes
- Development and application of resource analysis and allocation tools and processes, scheduling tools and processes, and risk management practices
- Coordination and integration with MCOTEA's Fiscal Section
- Support of Information Sharing and Management processes
- Best practices identification and implementation

S-1: Human Capital and Administration

The S-1 is the primary advisor to the Director, Deputy Director, divisions, and staff on all civilian and military personnel matters, and maintains accountability of all personnel. The S-1 is responsible for:

- Developing plans, policies, procedures, and

programs related to civilian and military human capital administration

- Overseeing the Unit Table of Organization
- Recommending manpower allocation in collaboration with the Staff Leads and Division Heads
- Performing administrative support functions including correspondence, mail, Freedom of Information Act requests, travel authorization, etc.
- Maintaining the MCOTEA Records Management system
- Coordinating Commanders' Conferences, ceremonies, change of command, and other events
- Processing Digital Message Service traffic
- Tracking and coordinating training for MCOTEA military and civilian personnel
- Handling protocol issues and public affairs
- Managing the civilian Performance Management System

S-3: Operations

The S-3 coordinates and manages MCOTEA's external logistics support for current and future operations. This includes, but is not limited to, the coordination of ranges and all facets of external logistics support expected from the host unit. The S-3 coordinates MCOTEA's attendance and participation at the Force Synchronization Conference to ensure that training and support requirements for the OTPOs in support of OT/DT are identified and deconflicted with the MARFORCOM G-3/5/7. The S-3 serves as MCOTEA's central point of contact when it comes to coordinating test schedules and test range usage.

In order to provide support, the S-3 takes an active role in the staffing of test documents and plans, and provides input from an S-3 perspective.

The S-3 is also the entry for programs at MCOTEA. As part of the civilian force development perspective, the S-3 coordinates with the Competency Leads from MARCORSYSCOM, or their designates, to give MCOTEA personnel the opportunity to get the same training that their Acquisition professional counterparts receive.

S-4: Logistics

S-4 Logistics/Information Technology Staff Section supports the organization in all areas of information technology (IT), logistics, facilities maintenance, and operational test. S-4 coordinates with test teams during test planning, analyzing IT and logistics requirements, developing solutions, and making recommendations for data collection and data management in support of test events. S-4 is responsible for the MCOTEA Test Support Facilities at Twentynine Palms Marine Air Ground Combat Center (MAGCC) and Camp Pendleton, providing site coordination for all support requirements during operational test.

The S-4 logistics section is responsible for:

- Ensuring the adequacy and sufficiency of test sites
- Coordinating use of test assets, ranges, and other facilities with other USMC, DOD, and Joint Services test agencies and logistics commands
- Coordinating the transportation of all equipment to and from test sites
- Serving as a single point of contact with Base Comptrollers for coordination of test site ranges, fuel, vehicles, food, ServMart and inventory management of test equipment
- Maintaining facilities and supply inventories

The S-4 IT section is responsible for:

- Managing MCOTEA data collection equipment and databases for testing world-wide
- Managing NIPR/SIPR web portals
- Navy Marine Corps Intranet (NMCI) contracting services
- Managing telecommunications, including Blackberry service, base telephones, VTC services
- Maintaining classified IT equipment and

network connectivity

- Reporting Information Assurance (IA) workforce compliancy
- Providing help desk support
- Providing SharePoint administration functions

Fiscal

The Fiscal Section manages all funds received throughout the year for Operations and Maintenance Marine Corps (O&MMC); Research, Development, Test, and Evaluation (RDT&E); and other programs. The Fiscal Section also develops Program Objective Memorandum (POM) briefs for consideration in the overall RDT&E and O&MMC POM submissions, and submits POM and budget exhibits justifying the request for resources. In addition, the Fiscal Section

- Manages and monitors transaction source documents
- Oversees the development of civilian labor cost projections
- Approves all credit card purchases and training requests
- Manages Procurement Request builder
- Oversees the Defense Travel System program
- Accepts invoices in Wide Area Work Flow

This page intentionally left blank for two-sided printing.

2

T&E Basics Background and Paradigm

MCOTEA's Mandate and Purpose
MCOTEA's Working Relationships with Other Organizations
Acquisition Life Cycle Overview
Test Relationship to Evaluation
The Evaluation Continuum
Collaboration Along the Acquisition Time Line
MCOTEA Test Team Billets and Best Practices
MCOTEA's 6-Step Evaluation Process

Chapter 2. Background & Paradigm

MCOTEA's Mandate and Purpose

MCOTEA is the independent OTA for the United States Marine Corps (SECNAV 2011). In this capacity, MCOTEA provides information to the MDA as part of the decision-making process for acquiring solutions that satisfy validated user needs. MCOTEA serves the MDA, the USMC, and the DOD by objectively evaluating, under operational conditions, how well a solution meets required mission capabilities. MCOTEA's role is to ensure that deployed systems accomplish their missions effectively without imposing unreasonable requirements on field support infrastructure.

Integrated testing is "the collaborative planning and collaborative execution of test phases and events to provide shared data in support of independent analysis, evaluation, and reporting by all stakeholders, particularly the developmental (both contractor and government) and operational test and evaluation communities" (Office of the Secretary of Defense 2008).

The fundamental purpose of Initial Operational Test and Evaluation (IOT&E) is to assist in managing the risks involved in developing, producing, fielding, operating, and sustaining systems and capabilities. Initial Operational Test (IOT), preceded by the materiel developer's developmental testing, investigates the Operational Effectiveness, Operational Suitability, and Operational Survivability (OE/OS/OSur) of an acquisition system. MCOTEA assists program acquisition by collaboratively planning and participating in integrated test events, observing developmental test events, and providing Observation and Assessment Reports throughout the acquisition cycle.

Evaluation of test data from integrated testing and IOT provides a basis for assessing system performance. System evaluation is typically an overarching strategy that gathers information from multiple developmental and operational test events.

MCOTEA strives to provide decision makers with timely information on program capabilities and limitations. To accomplish this, MCOTEA ensures that each system proposed for acquisition is tested adequately, evaluated objectively, and reported on independently. Integrated testing and system evaluation allow the acquisition community to learn about and correct or mitigate a system's operational limitations before full-rate production (FRP) and deployment. In turn, a fielded system's user community can apply knowledge gained from IOT&E to optimize system use.

To properly measure a system's capabilities, MCOTEA uses a Mission-Based Testing approach and custom designs each evaluation strategy. Test planning focuses on the missions the system is designed to support. Top-level requirements for adequate operational testing are as follows:

- Employ a production-representative system in realistic operating conditions with typical Marine operators and maintainers
- Collect data that accurately describes the test conditions and system performance results
- Analyze the data independently and without bias for use in system evaluation

Top-level requirements for objective system evaluation are as follows:

- Collect and evaluate information from a variety of developmental and operational test events
- Determine if thresholds in the approved capabilities documentation and Critical Operational Issues (COI) have been satisfied
- Determine the system's OE/OS/OSur
- Assess system effects on combat operations

OE/OS/OSur

⇨ OE is based on mission success

⇨ OS is based on factors that affect mission accomplishment

⇨ OSur is based on the degree to which the system puts operators at risk

- Provide any additional information on the system's operational capabilities

MCOTEA's Working Relationships with Other Organizations

MCOTEA reports directly to the ACMC and interacts with other organizations at various levels and to varying degrees (fig. 2-1).

Working Partners

MCOTEA's closest working partners, Deputy Commandant for Combat Development and Integration (DC, CD&I) and Marine Corps Systems Command (MCSC)/Program Executive Officer Land Systems (PEO LS), form the acquisition "triad" with MCOTEA.

Deputy Commandant for Combat Development and Integration

The DC, CD&I is responsible for identifying gaps in combat capabilities and for generating the Joint Capabilities Integration Development System (JCIDS) documents to address these gaps, including the

- Initial Capabilities Document (ICD)
- Capability Development Document (CDD)
- Capability Production Document (CPD)
- Concept of Operations (CONOPS)
- Concept of Employment (COE)

MCOTEA works closely with the DC, CD&I organization, primarily the Capabilities Development Directorate, very early in the system's acquisition cycle to ensure that requirements are testable and that MCOTEA understands the context in which the requirements were generated.

All inquiries or discussions about requirements should occur in the context of a Capabilities Development Integrated Product Team (IPT) and/or the Test and Evaluation Working-level IPT (T&E WIPT). These two chartered activities are essential and official tools of integrated testing (USMC 2010).

Marine Corps Systems Command

MCSC is the Commandant's agent for acquiring and sustaining systems and equipment used to accomplish the warfighting mission. MCSC addresses system capabilities and requirements generated by DC, CD&I. MCOTEA works closely with MCSC from early in the acquisition cycle to after IOT to help mitigate program risk. The Commander, MCSC is the Marine Corps Executive Agent for DT.

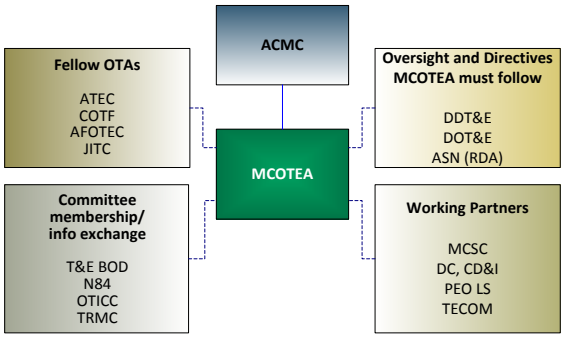


Figure 2-1. MCOTEA's Relationship with Other Organizations

Program Executive Officer Land Systems

PEO LS partners with MCSC to develop, deliver, and provide lifecycle planning for assigned programs. As with MCSC, MCOTEA works closely with PEO LS from early in the acquisition cycle to after IOT to help mitigate program risk. As with MCSC, MCOTEA observes developmental testing and conducts assessments on systems with PEO LS and conducts Initial Operational Test (IOT) on selected systems, as required.

Oversight/Non-Chain of Command

Director, Operational Test and Evaluation

The Director, Operational Test and Evaluation (DOT&E) in the Office of the Secretary of Defense (OSD) is the principal OT&E official within the DOD. DOT&E's job is to help ensure that a system is operationally effective and suitable before going beyond Low-Rate Initial Production (LRIP). Stated another way,

DOT&E's primary interest is to ensure that OT&E and Live-Fire Test and Evaluation (LFT&E) are adequate before FRP or deployment, and that tests and evaluations are properly executed according to statute and DOD policy.

Although not in the MCOTEA chain of command, DOT&E has significant oversight over any MCOTEA programs on the DOD T&E Oversight List, and MCOTEA is required to conform to DOT&E guidance for these programs.

Any program, regardless of Acquisition Category level, can be included on the T&E Oversight List. Selection criteria include ACAT level, Congressional and/or DOD interest, programmatic risk level, technical complexity, and relationship with other systems. All "oversight" programs require additional briefings, reports, and supporting documentation and often require additional testing. The DOT&E website at http://www.dote.osd.mil contains the Annual T&E Oversight List. The Defense Acquisition Guidebook (DAU 2012) contains additional details.

DOT&E's primary responsibilities for Oversight List programs are to provide final approval for the TEMP before milestone decision reviews and to approve operational test plans before those tests may commence. No operational testing may occur for a program on the Oversight List until DOT&E has provided written approval of the operational test plans. Early involvement of DOT&E personnel in drafting the T&E strategy, the TEMP, and operational test plans for programs on the Oversight List will help ensure smooth approval.

Assistant Secretary of the Navy (Research, Development, and Acquisition)

Assistant Secretary of the Navy (Research, Development, and Acquisition) (ASN (RDA)) is the DOD's Component Acquisition Executive for acquisition activity, including test and evaluation. ASN (RDA) provides DON-level acquisition and T&E guidance to supplement guidance from DOD. Although ASN (RDA) is not in the MCOTEA chain of command, MCOTEA is required to conform to ASN (RDA) T&E guidance.

DON uses the Gate Review process to help monitor programs of interest. The Gate Review process provides a framework for engaging senior naval leadership on certain acquisition programs to improve decision making through better understanding of program risks and costs (SECNAV 2011).

Gate Reviews

The Gate Review process helps ensure alignment between capability requirements and acquisition while improving senior leadership visibility into program risks and costs throughout the development cycle. DON has adopted the Probability of Program Success (PoPS) approach, used in conjunction with Gate Reviews, to assess and monitor the health of naval acquisition programs. Program health is subdivided into 17 metrics, one of which is T&E.

Six Gate Reviews are distributed over two "passes." Figure 2-2 shows where the Gate Reviews fall in the acquisition process. The first three gates constitute the "requirements" gates while the last three constitute the "acquisition" gates. The Gate Reviews are conducted at the 3-star level and above, and attendance is by invitation only. Table E1T3 of SECNAVINST 5000.2E (SECNAV 2011) contains more detail about participants and topics for each Gate Review.

MCOTEA is periodically called upon to contribute to or attend a Gate Review

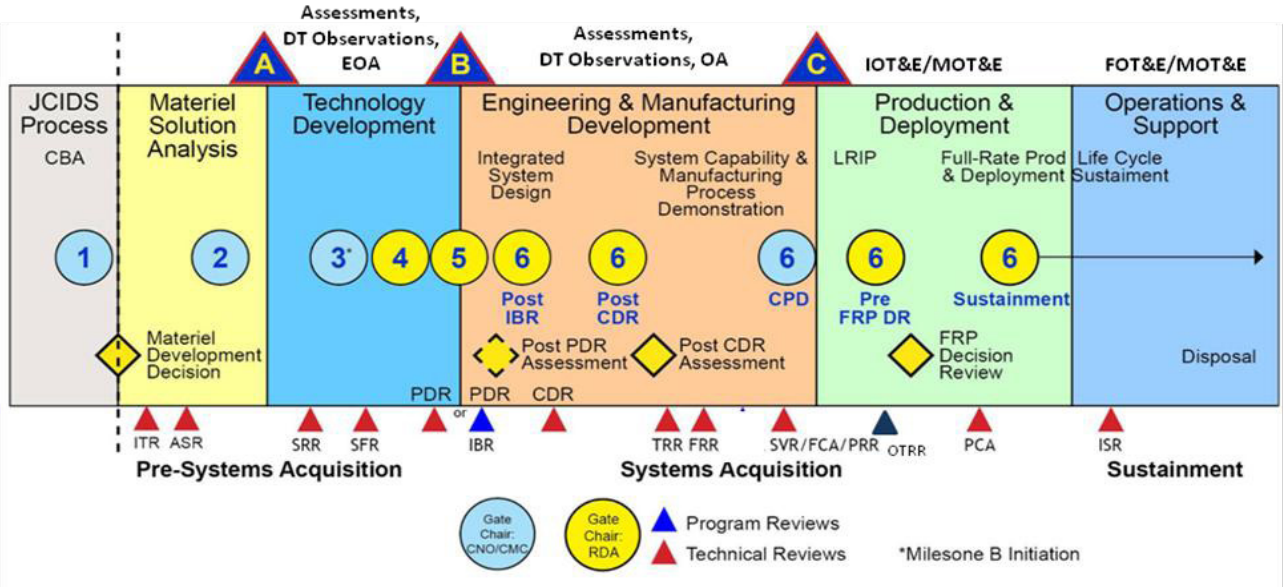


Figure 2-2. Gate Reviews in the Acquisition Process

pertaining to the T&E metric. Although this typically happens at Gate 6 (there are usually multiple Gate 6s), MCOTEA could be involved in earlier Gate Reviews as well.

The key to success during a Gate Review is to coordinate with the materiel developer's Program Manager (PM) ahead of time so the PM understands MCOTEA's concerns and MCOTEA understands the PM's position and proposed courses of action.

Fellow Operational Test Agencies

Each Service conducts OT&E through its respective OTA. MCOTEA periodically meets with fellow OTAs in various forums to discuss DOD-wide issues relating to OT&E. MCOTEA also participates with one or more of these OTAs in conducting MOT&E (ATEC 2012).

Joint Interoperability Test Command

For information technology systems (including National Security Systems) with interoperability requirements, the Joint Interoperability Test Command (JITC) is required to provide system Net-Ready certification memoranda to the Director, Joint Staff J-6 throughout the system life cycle, regardless of ACAT. JITC's philosophy is to leverage other planned test events to generate required data for the OSD-directed Net-Ready Key Performance Parameter (KPP) certification. A special test will be necessary only if other events do not provide the appropriate data.

Communication/ Information Sharing

Navy Enterprise T&E Board of Directors

The T&E Board of Directors (BOD) primarily addresses issues of concern to the Navy. The Director, MCOTEA is a member of the T&E BOD. SECNAVINST 3900.44 states "The

MCOTEA's Fellow OTAs

Navy: Operational Test and Evaluation Force (OPTEVFOR), headquartered in Norfolk, VA.

Air Force: Air Force Operational Test and Evaluation Center (AFOTEC), headquartered at Kirtland AFB, NM.

Army: Army Test and Evaluation Command (ATEC), headquartered in Aberdeen, MD.

Joint Command: Joint Interoperability Test Command (JITC) headquartered at Fort Huachuca, AZ.

Marine Corps members of this board will participate on a limited basis, pending corporate decisions on the applicability of the Enterprise concept of operations for the Marine Corps" (SECNAV 2009). Involvement in this Board helps MCOTEA stay abreast of new instructions, issues, and direction from SECNAV and opens a line of communication between MCOTEA and OPTEVFOR. SECNAVINST 3900.44 contains a list of all Board members (SECNAV 2009).

N084

N084 is the OPNAV Director, Test and Evaluation and Technology Requirements; N084 establishes T&E requirements and issues policy, regulations, and procedures governing Navy T&E. Historically, N084 has served as a conduit for MCOTEA to ASN (RDA) by promulgating directives from ASN (RDA) to MCOTEA and including MCOTEA in the review of key pending SECNAV documentation.

OSD Test Investment Coordinating Committee

The OSD Test Investment Coordinating Committee (OTICC) is the primary coordinating structure for test and evaluation investment matters within OSD. The OTICC advises the Director, Test Resources Management Center (TRMC) in oversight of the development of test technology and Joint test capabilities. MCOTEA is a primary member of the OTICC.

Test Resources Management Center

The Test Resource Management Center (TRMC) coordinates DOD test and evaluation resources and implements the annual DOD Strategic Plan for DOD T&E Resources. The primary program for execution oversight is the Central Test and Evaluation Investment Program (CTEIP) and the DOD T&E and Science and Technology (S&T) Programs. CTEIP includes the Joint Improvement and Modernization Program, the Resource Enhancement Project, Threat Simulators, and Target Management Investment projects. TRMC, in conjunction with the OTICC, coordinates with the T&E Executive Agents for each Service on the review and submission of T&E/S&T projects to ensure that Service/Agency Improvement and Modernization projects are addressed. MCOTEA participates as a primary member on all of these program/project working groups.

Acquisition Life Cycle Overview

ACAT Designation

One of the earliest steps in an acquisition system's lifecycle is ACAT designation. A program's ACAT is based on cost and/or MDA designation as a special interest. The ACAT level determines both the level of review required by law and the MDA's level within DOD. All ACAT programs except ACAT IV(M) and Abbreviated Acquisition Programs (AAP) require operational testing. MCOTEA participates in the ACAT determination process when the MDA requests MCOTEA's written concurrence with ACAT IV(M) or AAP designation.

Evolutionary Acquisition

Evolutionary acquisition delivers system capabilities in increments. A program executing an evolutionary acquisition strategy incorporates time-phased requirements into the system. Block upgrades, planned product improvements, and other efforts that provide a significant increase in operational capability and meet an ACAT threshold are managed as a separate increment (DOD 2008).

The evolutionary approach recognizes the need for incremental improvements at the beginning of a program. The idea is to balance technological maturity with evolving threats, cost, and the need to

get a capability to the user quickly. This allows the fielding of an initial, well-defined, and significant core operational capability quickly in response to validated requirements. This strategy results in fielding increased capability in succeeding increments as technology matures.

Incremental Testing Requirements

Figure 2-4 shows that, aside from developing the initial capability, each increment starts at the technology development phase and has its own milestones and operational testing requirements (SECNAV 2011; DAU 2012). The CDD defines the KPPs and Key System Attributes (KSA) that apply to each increment of Engineering and Manufacturing Development. Each increment will complete DT&E, OT&E, and LFT&E as required. An independent phase of OT&E must be completed for each increment before release to the user for programs requiring OT&E. As suggested by the figure, each increment is treated individually and will be at a different phase in the OT&E process at any particular time. This will involve concurrent test planning and execution activity for the different increments and may result in a higher degree of complexity, requiring each increment to be carefully tracked. The evolutionary strategy for each increment will be described in the TEMP.

In general, T&E that has confirmed the mission capabilities of an increment need not be repeated in its entirety to confirm that the subsequent increment continues to provide those mission capabilities. "However, regression testing to reconfirm previously tested operational capabilities and/or suitability might be required if the subsequent increment introduces a significantly changed hardware or software configuration, or introduces new functions, components, or interfaces that could reasonably be expected to alter previously confirmed capabilities" (DAU 2012).

Test and Evaluation Paradigm

The MCOTEA approach to testing and evaluating is designed to maximize synergy with the rest of the Marine Corps acquisition process consistent with federal law and DOD, DON, CJCS, and Marine Corps guidance. This approach reduces program risk and overall cost, thereby maximizing value to the Marine Corps and DOD. In accordance with DODI 5000.02 (DOD 2008), MCOTEA must accomplish the following before or during IOT&E:

- Determine if thresholds in the approved

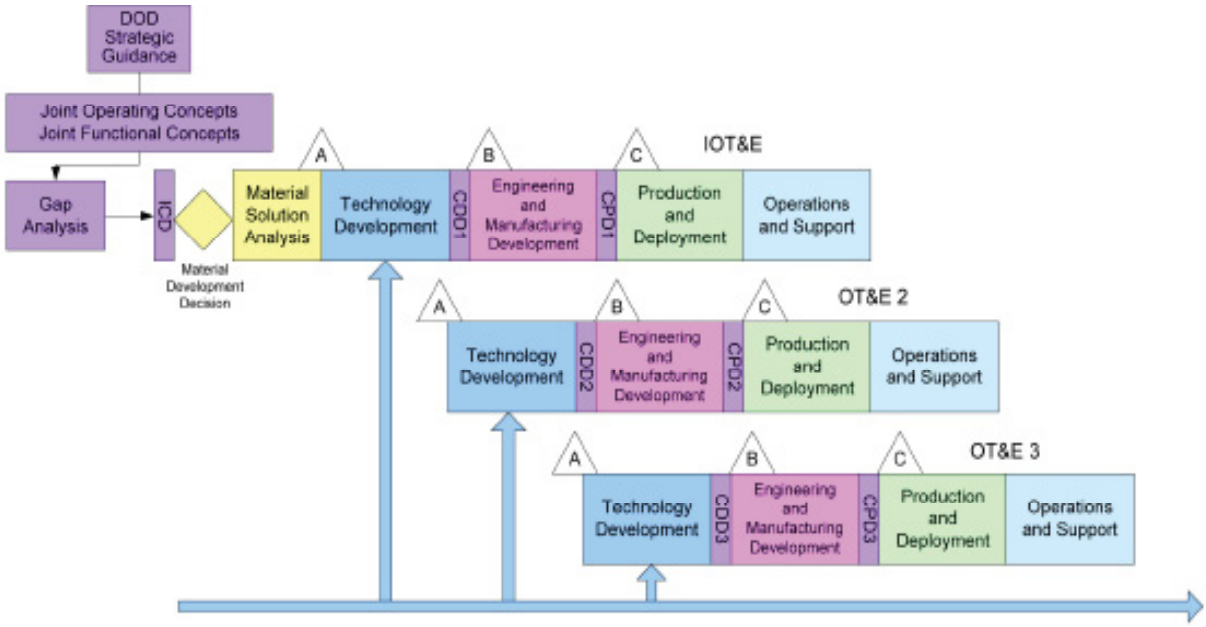


Figure 2-4. Incremental Technology Development

capabilities documents and COIs have been satisfied

- Determine OE/OS/OSur of the system under realistic operational conditions, including Joint combat operations
- Assess the effect to combat operations
- Provide additional information on the system's operational capabilities and limitations

The evaluation associated with accomplishing these tasks is rooted in a process that takes place throughout the life of a program. MCOTEA uses the results of non-MCOTEA developmental testing, when appropriate, as well as the results of MCOTEA's assessments and operational testing. MCOTEA accomplishes these tasks using a combination of integrated planning and frequent testing in conjunction with continuous evaluation. MCOTEA employs the DOD definition of integrated testing: "Integrated testing is the collaborative planning and collaborative execution of test phases and events to provide shared data in support of independent analysis, evaluation, and reporting by all stakeholders, particularly the developmental (both contractor and government) and operational test and evaluation communities" (OSD 2008). MCOTEA does not call out individual tests as being "integrated"; instead, MCOTEA collaboratively plans all test phases with the Materiel Developer throughout the life of a program while maintaining the independence of IOT. Although test events are collaboratively planned to ensure all needed data will be available, and some may be collaboratively executed (excluding IOT/Follow-on Operational Test (FOT)/Multi-Service Operational Test (MOT), which is executed only by MCOTEA), both the DT and the OT evaluations must be done separately and independently.

Test Relationship to Evaluation

Test and evaluation are often thought of as a single process, while in reality they are two related but distinct processes. Testing involves the physical exercising, by trial or examination, of a component, system, concept, or approach for the sole purpose of gathering data and information regarding the item under test. Evaluation seeks to ascertain the worth of, or to fix the value of, a component, system, concept, or approach. Testing provides a source of data for the evaluation process that uses the data to derive useful information about what has been tested. The relationship of testing to evaluation is many-to-one; that is, several tests may be required to support a single evaluation.

MCOTEA's System Evaluation Plan (SEP) creates a framework and methodology for evaluating the entirety of program data obtained from assessments and IOT. The SEP is intended to provide a transparent, repeatable, and defensible approach to evaluation with the added benefit of minimizing the overall cost of program testing. In addition, MCOTEA welcomes DOT&E, Program Office, and DC, CD&I suggestions pertaining to the SEP; however, in the final analysis, the evaluation process belongs to MCOTEA and MCOTEA is under no obligation to accept these suggestions.

As the OTA for the Marine Corps, MCOTEA is charged with both the operational testing and evaluation of systems. The purpose of operational testing is to determine how the system performs under test using production-representative components, operated and maintained by typical users, under realistic operational conditions. An Operational Test (OT) is a discrete event that provides invaluable information about the system under test and its expected capabilities and limitations during combat operations. It is a major input to the evaluation of the system, but

not the only input.

Other tests and assessments increase the knowledge about the system under test as the system matures during the acquisition cycle. These tests also provide input to the overall system evaluation. During developmental testing, system components are checked to ensure that they function as designed and the system is checked to ensure that it meets the requirements derived from the ICD/CDD/CPD. MCOTEA generally uses the data gathered during DT to determine if the thresholds in the approved capabilities documentation have been demonstrated. In addition, aggregating DT data over time can be useful in determining aspects of a system's OS. Furthermore, MCOTEA's assessments provide insight into the level of system maturity and overall system capabilities and limitations.

The purpose of the evaluation is to use all relevant information from DT, MCOTEA's assessments, operational testing, relevant M&S results, and the results of any Live Fire testing to determine OE/OS/OSur. Evaluation involves compilation and analysis of data gathered over the life of the program, with emphasis on system performance during operational testing.

The Evaluation Continuum

Advantages of MCOTEA's Early Involvement

According to DODI 5000.02, "T&E expertise must be brought to bear at the beginning of the system lifecycle to provide earlier learning about the strengths and weaknesses of the system under development" (DOD 2008). MCOTEA does not wait until a full-blown OT is needed to get involved in the program acquisition process. Ideally, MCOTEA involvement begins very early in the acquisition cycle. MCOTEA's goal is to become involved in a new program as early as the formation of the Requirements Transition Team (RTT), a team formed before the Materiel Development Decision (MDD) to facilitate the development and transition of potential requirements into the acquisition process (USMC 2010). This early involvement includes early program reviews, demonstrations, developmental working groups, M&S activities, and other technical development activities. According to SECNAVINST 5000.2E, "Early, active, and continuous participation by test agencies during the development of capabilities documents will support effective communication and common interpretation" (SECNAV 2011).

MCOTEA's goal is to develop draft COIs (see Chapter 3-1) prior to the Analysis of Alternatives (AoA). The AoA identifies potential options to an MDD, thereby guiding the Materiel Solution Analysis phase of acquisition.

Having draft COIs available allows the AoA to examine alternatives based on the same high-level Issues the system will be expected to address throughout its lifetime, including during operational testing. Although other Issues may also be examined during the AoA, the COIs should form the basis of the major areas of comparison addressed in the AoA.

In addition to MCOTEA's involvement in monitoring and analyzing developmental testing, use of assessments early in a system's development can help to identify technology risks and illuminate potential operational issues. Integrated testing and early OAs can be expected to emphasize the use of prototypes. Early MCOTEA involvement should benefit the entire Marine Corps acquisition process while minimizing the cost of the overall program.

Continuous Evaluation

The evaluation of a system is the result of the accumulation of data and facts about the system obtained during the entire acquisition cycle (SECNAV 2011). This accumulation of data starts with early research and developmental

MCOTEA's involvement at an early stage benefits both MCOTEA and the new program for the following reasons:

- ⇨Generates COIs at an early stage so system designers know the high-level issues their system is intended to address
- ⇨Lends operational test perspective that aids in developing unambiguous requirements that can be tested
- ⇨Helps MCOTEA gain better understanding of the context in which the capabilities and requirements were determined
- ⇨Provides insight into potential system and operational deficiencies early in the program when remedial action can easily be taken
- ⇨Provides insight into potential IOT requirements to ensure that range capabilities and technologies exist to meet those test requirements. If a shortfall is recognized early enough, initiation of a test technology development program may be in order.
- ⇨Provides independent insight to decision makers into the program's progress toward meeting the desired level of Operational Effectiveness, Suitability, and Survivability

testing and continues through IOT and FOT. Integrated testing and early assessments can contribute important contextual information, result in enhanced understanding of system capabilities, and make significant contributions to satisfying the requirement to examine the extent to which CDD/CPD thresholds have been satisfied. Of course, the events that will yield the most important information from the system evaluation perspective are the IOT and, if applicable, FOT.

Figure 2-5 illustrates how input from various assessments and testing events contribute to the aggregated evaluation of a system. As shown in the figure, in addition to OT results near the end of the acquisition cycle, the results of observations and assessments at earlier stages in the program are fed back to the program to help the PM identify program risks. Waiting until IOT to evaluate a system for the first time does little to affect the actual design of the system. Therefore, MCOTEA provides feedback to the PM and MDA periodically during the acquisition cycle. This feedback indicates if a program is progressing towards IOT and identifies potential concerns.

Early Identification of Deficiencies

Identification of system deficiencies is

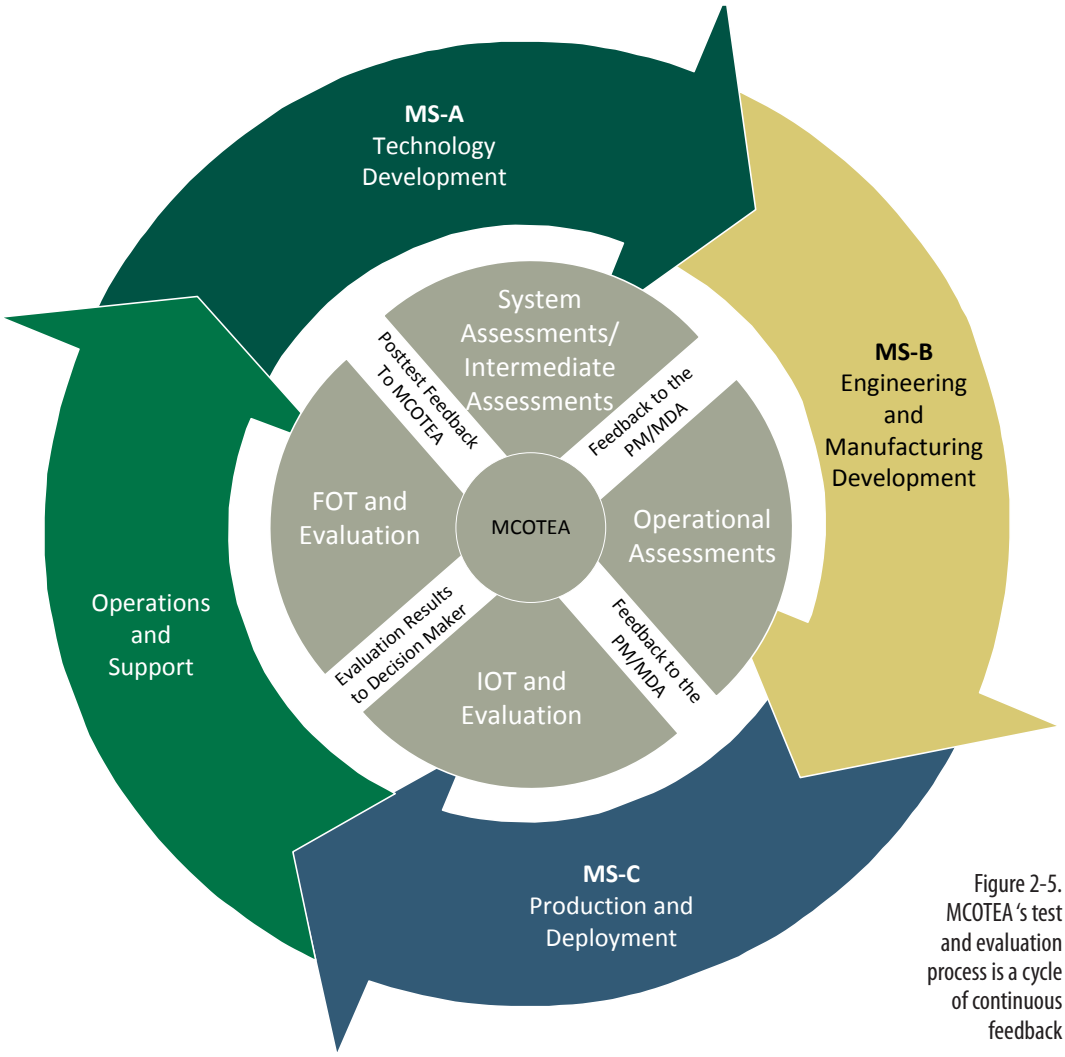


Figure 2-5. MCOTEA's test and evaluation process is a cycle of continuous feedback

most valuable early in the program. The value of identifying deficiencies diminishes as the program matures. Alterations to a more mature program are more difficult and expensive to make, whereas assessing program progress at early and intermediate stages enables the Marine Corps to adjust the program more effectively.

System assessment feedback that occurs early in a program is different in nature from the evaluation of a mature program. MCOTEA assesses a system's progress based on standards appropriate for its developmental stage. Early evaluation feedback tends to be limited in scope, but this feedback builds a history for the program that shows when issues were identified and how they were mitigated. This opens an additional window on how the program is maturing as a function of time.

Finally, obtaining Warfighter feedback after system fielding is important for optimizing the MCOTEA test process as well as the Marine Corps acquisition process.

Collaboration Along the Acquisition Time Line

Throughout the acquisition cycle, MCOTEA brings the operational testing perspective to all milestone assessment teams. In general, the Materiel Developer, MCOTEA, and DC, CD&I will participate in one another's Subject Matter Expert (SME) panels throughout the life of a program. The cognizant MCOTEA Division can expect to participate in various Gate Reviews to support the briefing requirements of PoPS program criteria pertaining to test and evaluation.

Pre-Milestone A

Figure 2-6 illustrates key points of MCOTEA's interaction with other agencies before MS A. Early in a program's life, the RTT stands up to facilitate the transition from desired capabilities to an actual system. Participation in this team may be MCOTEA's first official activity on a new program. MCOTEA reviews the draft Initial Capabilities Document (ICD) once it is written to ensure that the proposed capabilities are testable. MCOTEA also participates in working groups that generate the applicable CONOPS and COE. This early participation with DC, CD&I enhances MCOTEA's understanding of the context in which the capabilities were generated.

Before the AoA, MCOTEA establishes draft COIs, which help the AoA team determine the categories for comparing alternatives and by using the process

Continuous evaluation increases the efficiency of the Marine Corps acquisition cycle in the following ways:

- ⇨Gathers important data on most of the thresholds in the capabilities documents before operational testing
- ⇨Allows evaluation feedback throughout the program focused on the PM and decision maker's needs and based on standards appropriate for the program's developmental stage
- ⇨Identifies important issues and potential deficiencies early enough in the program to allow relatively inexpensive corrective action
- ⇨Enables an independent mechanism for tracking program progress over time
- ⇨Allows operational testing to focus on COIs and operational mission performance rather than specification and threshold compliance

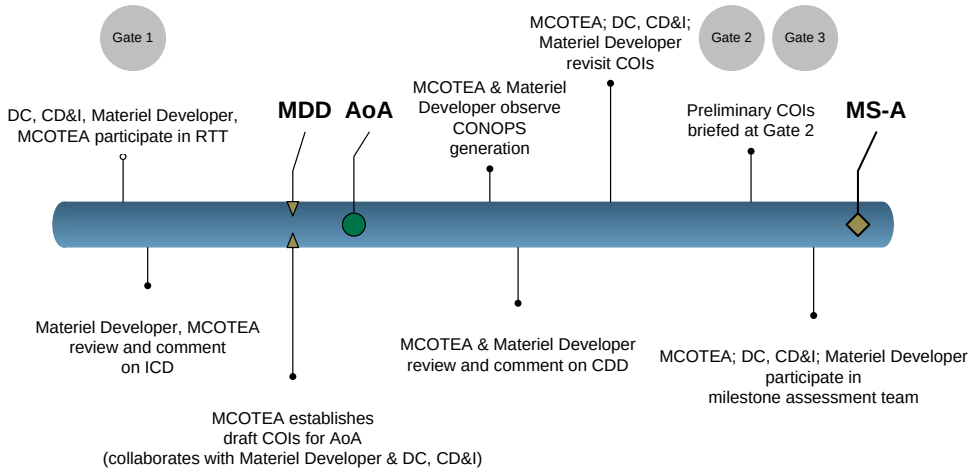


Figure 2-6. MCOTEA Interaction with System Acquisition Before MS A

introduced later in this chapter and described in detail in chapter 3-1. The goal is to use essentially the same COIs for system evaluation from the AoA through IOT. After the AoA, MCOTEA revisits the COIs and updates them based on information discovered during the AoA, an updated understanding of the system CONOPS, and any new information available in the capabilities documentation.

These preliminary COIs may be briefed at the Gate 2 review. The activity before MS A constitutes steps essential to the development of the program TEMP and the MCOTEA SEP.

Milestone A to Milestone B

The collaborative approach continues between MS A and MS B (fig. 2-7). At MS A and B, MCOTEA receives a copy of the Acquisition Decision Memorandum (ADM). After MS A, MCOTEA continues developing a framework for the evaluation by establishing test conditions, determining any implied Attributes, and tracing all Attributes to Subtasks, Tasks, and ultimately the COIs. The final COIs to be used in operational testing may be briefed at Gate 4.

The Materiel Developer and MCOTEA work together to efficiently assign Subtasks and Tasks (and their associated Issues) for examination, under specified conditions in Developmental Test (DT) and OT and assessments in accordance with the TEMP. Attributes with thresholds are also assigned to test events in the TEMP. The initial allocation of Subtasks and Tasks to specific tests may need to be modified based on test results themselves; however, the goal of allowing the IOT to focus on mission performance under realistic operational conditions remains unchanged. The goal is to examine all Attributes with thresholds to meet OT requirements before IOT, as well as to build a database to support the suitability determination.

Before issuing the Request for Proposal (RFP), the Materiel Developer and MCOTEA ensure that the RFP is consistent with the TEMP, and MCOTEA provides inputs to the Contract Deliverables Requirements List. In particular, MCOTEA input ensures that any contractor developmental test data and reports are available for inspection and possible inclusion in the overall evaluation plan. The Materiel Developer will consult with MCOTEA when determining the source selection criteria. MCOTEA may participate in prototype demonstrations

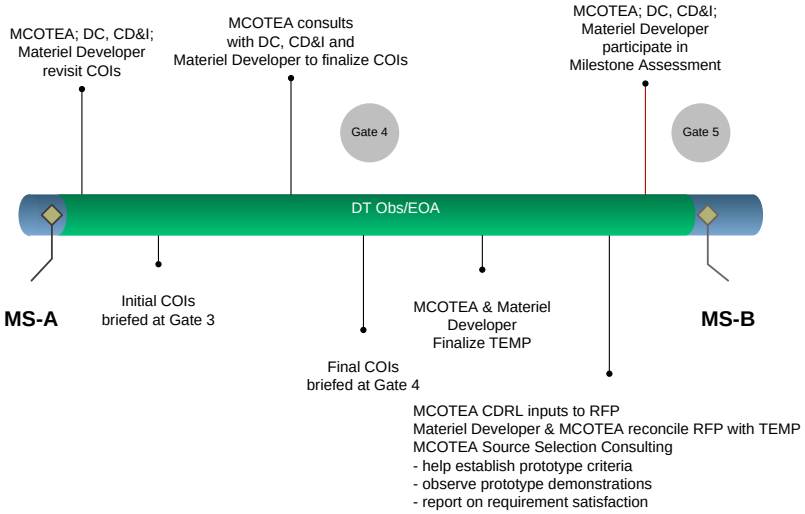


Figure 2-7. MCOTEA Interaction with System Acquisition Between MS A and MS B

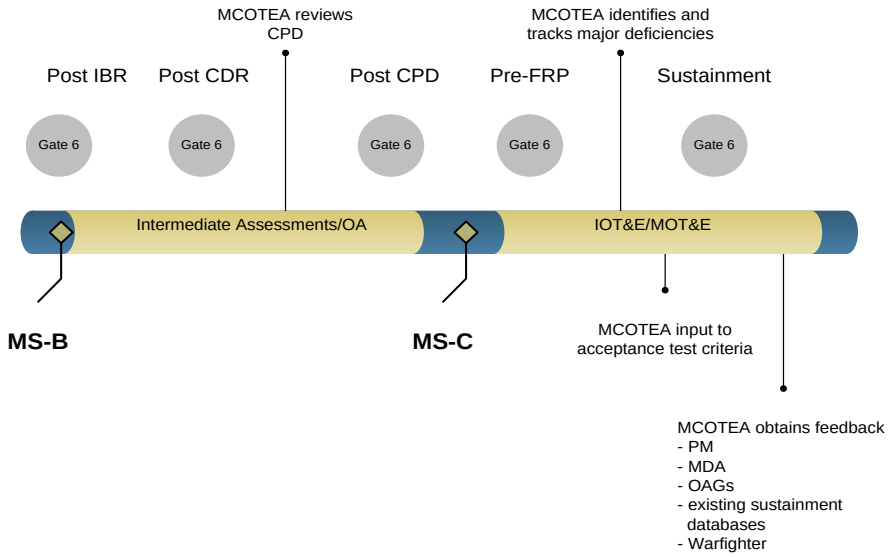


Figure 2-8.MCOTEA Interaction with System Acquisition Post MS B

associated with source selection and will have access to data obtained during prototype testing. After prototype testing, MCOTEA will provide input to the Materiel Developer from the operational test perspective; however, MCOTEA will not offer an opinion on relative candidate system performance.

Post Milestone B

After Milestone (MS) B, in addition to performing any assessments, MCOTEA reviews the CPD (fig. 2-8) and continues to plan for IOT.

After MS C, MCOTEA provides input to the Materiel Developer concerning the acceptance test criteria used for each early system purchased. MCOTEA also expeditiously alerts the PM and MDA of any major system or operational deficiencies discovered during integrated or operational testing. Finally, MCOTEA seeks feedback from multiple sources with an eye toward improving MCOTEA's processes. These sources include the PM, MDA, Operations Advisory Groups (OAG), databases designed to monitor suitability data of systems after fielding, and Warfighter feedback from deployed units.

Obtaining and Using Developmental Test Data

MCOTEA leverages early testing opportunities during DT to maximize available information for decision makers and to minimize the risk and expense of the entire testing program.

The Integrated Test and Evaluation approach is formulated before any developmental testing takes place. The T&E approach is described in the Program Manager's Test and Evaluation Strategy, while the plan is described in detail in the TEMP. MCOTEA participates in TEMP development to reflect the integrated test approach and constructs its own SEP (see chapter 3-1) that details how data will be aggregated and used in the final system evaluation.

MCOTEA is aware of planned DT events through participating in the T&E Working-level Integrated Product Team (WIPT) and can expect to participate in the collaborative planning of these events. For MCOTEA to participate in a DT event at any level, the draft DT Plan

must be available for MCOTEA's review in ample time for MCOTEA to comment and offer suggestions based on shared data needs. The DT team may or may not accept these suggestions, based on time and cost constraints. However, the PM should be aware that MCOTEA testing requirements will need to be satisfied at some point, and although incorporating them into a DT event may raise the cost of that particular test, it may decrease the overall program testing cost and reduce risk by satisfying MCOTEA's requirements early.

MCOTEA Test Team Billets and Best Practices

MCOTEA test teams are functionally matrixed. The Division Head, with input from the Operational Test Project Officer (OTPO), constructs the team based on the needs of a program. The typical team is composed of an Operational Test Project Officer (OTPO), a Test Manager (TM), an Operations Research/Systems Analyst (ORSA), a Mathematical Statistician (MS), and a Data Manager (DM) (fig. 2-9).

There a number of other team members that must be considered when forming the test team (Cyber Analyst, Live Fire Analyst, Human Factors Analyst, etc.). Following are some generalized responsibilities for the typical team members. For more detailed responsibilities refer to Chapter 3.

The Operational Test Project Officer

The OTPO is the program lead and is responsible for managing the entire operational test program. Test project management requires staff action in three areas: OT&E documentation; system user-developer coordination; and OT&E resource management (cost, schedule, performance). The OTPO performs the following functions:

Before Testing

- Develops the POA&M, FoS Message, and OTRB Briefing
- Participates in the Capabilities Development and T&E WIPTs
- Reviews and comments on all Plans

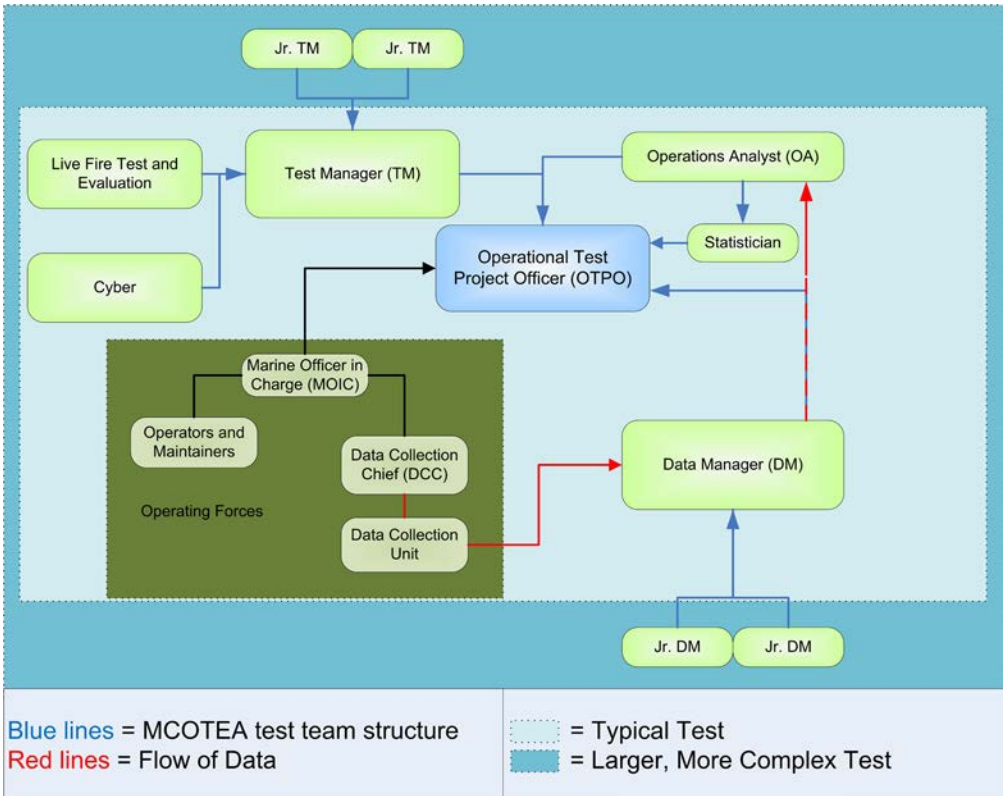


Figure 2-9. Test Team Organization

(Assessment, Evaluation, M&S V&V, Test Concept, etc.)

- Requests Developmental Test Plans from the Program Office
- Supports the OTRR meeting
- Leads the Data Collection V&V event
- Coordinates test range, logistics, instrumentation, transportation, & personnel resources

During Testing

- Supports movement to and from test site for QRA, Operational Assessment (OA), or OT
- Supports pretest setup and coordination
- Supports Pilot and Record Tests
- Performs posttest activities, data review, and teardown

After Testing

- Reviews and comments on Developmental Test, M&S V&V, and Observation Reports
- Supports the Data Review meeting
- Provides input to Test Data, Assessment, and Evaluation Reports
- Develops Lessons Learned
- Responsible for Records Management Closeout

Test Manager

The TM assists the OTPO in planning, executing, and reporting operational test events. In addition to writing the test plan, the TM helps coordinate the test team, coordinates with the S-3 to make logistical arrangements for the test site, and remains at the test site throughout test execution.

With each additional test managed, specific characteristics and lessons learned from previous tests can be applied to the new test. However, each test is unique and management rules will change to facilitate the particular requirements of the type of testing being managed. Flexibility and open-mindedness are critical to managing a test program well. The TM performs the following functions:

Before Testing

- Provides input to the Test and Evaluation Master Plan
- Develops Developmental Test Observation Plans
- Develops Test Plans
- Provides input to the Test Concept, FD/SC Charter, FoS Message, OTRB brief, M&S V&V Plans
- Reviews and comments on Developmental Test Plans
- Supports the Data Collection V&V event
- Coordinates test range, logistics, instrumentation, transportation, & personnel resources

During Testing

- Observes identified M&S V&V and developmental testing events
- Conducts movement to and from test site for QRA, OA, or OT
- Performs pretest setup and coordination
- Executes Pilot and Record Tests
- Performs posttest activities, data review, and teardown

After Testing

- Requests Developmental Test Reports from the Program Office
- Develops Developmental Test Observation Reports
- Supports the Data Review meeting
- Provides input to Test Data, Assessment, and Evaluation Reports
- Provides input to Lessons Learned

Operations Research/Systems Analyst

The ORSA plans for and conducts evaluation of test data. This is done by developing the SEP or the SAP. The ORSA also assists with the Test Concept development, test execution, and data

collection. The ORSA performs the following functions:

Before Testing

- Provides input into the POA&M
- Reviews and comments on capability documents
- Develops the Evaluation, Assessment, and M&S V&V Plans
- Provides input to the Test, and M&S V&V Plans
- Provides input to the Test Concept and TEMP
- Reviews and comments on Developmental Test Plans
- Supports the Data Collection V&V event

During Testing

- Supports pretest setup and coordination
- Supports the execution of the Pilot Test
- Supports posttest activities, data review, and teardown

After Testing

- Provides input to Test Data Report
- Supports the Data Review meeting
- Reviews and comments on M&S V&V Report
- Develops the M&S Accreditation, Assessment, and Evaluation Reports
- Provides input to Lessons Learned

Mathematical Statistician

The MS plans for and conducts analysis of test data. This is done by developing the Test Concept and the Test Data Report. The MS also assists with the development of the SEP/SAP and Test Plans, Developmental Test Plan reviews, data collection V&V, test execution, and data collection. The MS performs the following functions:

Before Testing

- Provides input to the SEP/SAP, Test Plans, and M&S V&V Plans
- Develops the Test Concept
- Provides input to the Test and Evaluation Master Plan
- Reviews and comments on Developmental Test Plans
- Supports the Data Collection V&V event

During Testing

- Supports pretest setup and coordination
- Supports the execution of the Pilot and Record Tests
- Supports posttest activities, data review, and teardown

After Testing

- Leads the Data Review meeting
- Reviews and comments on M&S V&V and Developmental Test Reports
- Analyzes data in accordance with the test plan
- Develops the Test Data Report
- Provides input to Test Data, M&S V&V, Assessment, and Evaluation Reports
- Provides input to Lessons Learned

Data Manager

Data Managers support the OTPO, TM, ORSA, and MS. The DM should establish a good working relationship with the test team and the support personnel (i.e., S-3, S-4) to ensure open communication, resulting in a positive working environment and a more efficient test. The DM performs the following functions:

Before Testing

- Provides input to the Test Concept and Assessment and Test Plans
- Supports the Data Collection V&V event

During Testing

- Leads the Data Collection V&V event
- Supports pretest setup and coordination
- Supports execution of the Pilot and Record Tests
- Supports posttest activities, data review, and

teardown

After Testing

- Provides input to Test Data Report
- Supports the Data Review meeting
- Provides input to Lessons Learned

Supplementary Team Members

The core test team (OTPO, TM, ORSA, DM, MS) may be supplemented with other specialty areas to include Cyber, Live Fire, Human Factors, M&S, etc.

MCOTEA's 6-Step Test and Evaluation Process

Step 1. System Evaluation Plan and FD/SC Charter

The SEP is MCOTEA's overarching plan for evaluating data that pertains to a system throughout the life of the program (DT as well as LFT&E and IOT). The SEP is the starting point of all IOT&E at MCOTEA and presents the methods and models by which MCOTEA will determine OE/OS/OSur. Annexes include Issues and Screening Criteria, the FD/SC Charter (which sets forth the failure definitions of mission essential functions), and the RAM Evaluation Plan.

Step 2. Test Concept and TEMP Input

During SEP development, the test team begins to develop details about the Test Concept, such as trial process flow, sample size, test limitations, test resources, required M&S support, etc., which become input to the TEMP. Careful and thorough development of the Test Concept leads to accurate and substantial TEMP input.

Step 3. Test Planning

The Test Plan is MCOTEA's detailed test execution and data collection plan. MCOTEA uses a mission-oriented context in operational testing to relate evaluation results to the Warfighter's ability to execute missions. Focusing on mission context during OT planning provides a robust OT environment and helps accomplish evaluation goals.

Test planning includes the following broad actions, all of which are explained in detail in chapter 3:

- Check Lessons Learned Database. The test team consults the Marine Corps Lessons Learned database (www.mccll.usmc.mil) for problems encountered and Lessons Learned during previous OTs.
- Establish the data collection plan. The plan includes Data Requirements as well as Methods for Data Collection, Reduction, and Analysis.
- Refine Test Trial Design. The test team refines the test trial design (initiated during Step 2) for collectiong test data formed around the missions the Marines will execute using the system under test.
- Refine Resource Requirements. The test team refines the resource requirements (initiated during Step 2) such as funding, required personnel from the Operating Forces, number of test articles, test site, instrumentation, etc.
- Confirm Readiness for Test. The Operational Test Readiness process ensures that the test team and system under test are ready to proceed to test.

Step 4. Operational Test Execution

With the approved Test Plan in hand and all preparations final, the test team arrives in the field to execute operational testing. Before the Record Test commences, however, four critical steps are taken:

- Observe New Equipment Training (NET). NET is required for all operators and maintainers participating in the OT. MCOTEA test team members observe NET because this is when Tactics, Techniques, and Procedures (TTP) are taught for the system under test. In addition, the OTPO and TM need to assess if the training has adequately prepared individuals to proceed to Pilot Test.

Characteristics of Operational Test

(IOT, FOT, and MOT)

Uses production or production-representative articles

Uses representative forces (friendly and opposing)

Employs realistic tactics, targets, and operational environments whenever possible

Determines OE/OS/OSur

May also support the decision to proceed beyond LRIP to FRP

- Execute the Pilot Test. The Pilot Test is used to validate the data collection plan and serves as a rehearsal and readiness check for the Record Test. The OTPO/TM must allow adequate time between the Pilot and Record Tests for careful examination of Pilot Test data results. If issues arise that are likely to affect the Record Test, MCOTEA leadership may decide to extend the Pilot Test.
- Execute the Record Test. The Record Test is the culmination of all IOT planning. Its essential purpose is to provide the data, collected under operational conditions, that is required to evaluate the system under test.
- Convene the FD/SC Scoring Conference. The scoring process examines the circumstances associated with each Test Incident Report (TIR).

Step 5. Operational Test Reporting

The MS ensures that the pedigree of the data taken is maintained and that all raw data taken during testing is saved and available for access well after testing is complete. In many cases data reduction, if required, depends on the analysis methodology in use. The raw data might be useful in future analyses and should be retained. Before leaving the test site, the test team writes the Test Data Report, which provides the reduced data on a CD and reports on test conduct, including any Test Limitations or Deviations.

Step 6. System Evaluation and Reporting

The test team produces the final Operational Test Agency Evaluation Report (OER), which includes a determination of OE, OS, and OSur as well as a report on the attainment of thresholds and an assessment of the system's impact to combat operations. The OER also includes a summary of all Major System and Operational Deficiencies noted throughout testing and evaluation.

Records Management and Lessons Learned

MCOTEA maintains all test data and other program records according to internal procedures as well as U. S. Government requirements. MCOTEA also records Lessons Learned using the Marine Corps Center for Lessons Learned Web site.

Process Feedback

MCOTEA continuously strives to improve its processes to ensure that MCOTEA tests and analyses are relevant, timely, accurate, unbiased, and operationally useful. To this end, MCOTEA solicits feedback from diverse sources as a means to improve existing processes and identify the need for potential new processes. Any suggestions for potential improvements to MCOTEA processes are forwarded to the Scientific Advisor for consideration. .

Potential sources of feedback include the following:

- MCOTEA test teams and test Operating Forces
- Databases on deployed systems
- PM and MDA
- OAG
- Warfighters themselves

Types of MCOTEA Tests

Operational Testing

This section refers to the steps required to execute individual operational tests: IOT, FOT, and MCOTEA-led MOT. Wherever IOT is mentioned, the concepts and procedures also apply to Follow-on Test and Evaluation (FOT&E) and MCOTEA-led MOT.

MCOTEA uses a mission-oriented context in operational testing to relate evaluation results to the impact on the Warfighter's ability to execute missions. Focusing on the mission context during operational test planning and execution provides a more robust operational test environment

and facilitates system evaluation goals.

Initial Operational Test and Evaluation

IOT&E consists of the test itself and the subsequent evaluation of test data. Initial Operational Test is a single but critical event, while evaluation is the result of a process, as explained in detail in later chapters. IOT is normally conducted during the Production and Deployment acquisition phase.

In general, IOT is the only OT phase required by Department of the Navy policy. In some cases, when the MS C decision and the FRP decision are planned concurrently, IOT may be performed during the Engineering and Manufacturing Development acquisition phase, prior to MS C. Note: No contractors developing the system under test may be involved in the operation or maintenance of the system during IOT unless the contractor will be involved in the same functions when the system is deployed in combat. If this is the case, contractor performance during IOT will be subject to review, analysis, and evaluation as part of the overall system evaluation.

After IOT, MCOTEA evaluates the data results along with other information from previous assessments and writes an Operational Test Agency Evaluation Repport (OER), which is forwarded to the ACMC. After ACMC approval, the OER is released to the PM and the MDA.

Follow-on Operational Test and Evaluation

Follow-on Operational Test and Evaluation (FOT&E) may be necessary after a successful MS C or FRP decision. The need for an FOT may be determined early by the MDA; if it is, it should be documented in the TEMP. Further reasons for an FOT&E include the following:

- To address a deficiency identified during system DT or OT
- To ensure that changes to the system since IOT have remedied previously recorded deficiencies and have not decreased system capability
- To refine the estimates, evaluate changes, and reevaluate the system to ensure that it continues to meet operational needs in a new environment or against a new threat

FOT&E employs the following:

- Production or production-representative articles
- Typical system users (Marines)
- Representative forces (friendly and opposing)
- Realistic tactics and targets when possible
- Operational conditions as close to actual as possible

MCOTEA evaluates the results of the FOT along with other relevant information and prepares an OFER, as described in chapter 4.

Multi-Service Operational Test and Evaluation

MOT&E is conducted jointly by two or more Services. When designated the Lead Service, MCOTEA prepares a single TEMP and MOT plan in coordination with all interested Services and defense agencies in accordance with the latest MOT&E Memorandum of Agreement (ATEC 2012).

Marine Corps Lead Service

When the Marine Corps functions as the Lead Service in an MOT&E, MCOTEA is responsible for accomplishing the following (not necessarily in this order):

- Conduct test planning, execution, and system evaluation in accordance with this manual
- Form the appropriate Multi-Service T&E WIPT
- Form a Test Management Council composed of one senior representative from each supporting Service to arbitrate disagreements not resolved at the T&E WIPT level
- Participate in early acquisition activities, including developmental testing, and invite other Service participation as they require
- Issue a call to the other interested Service

OT&E agencies for COIs and their Service-unique resource requirements

- Coordinate action on the TEMP to account for other Service issues and inputs
- Call a meeting of participating OTA Test Managers to assign responsibility for accomplishing evaluation and test objectives
- Formulate the test and evaluation strategy and portions of the TEMP in coordination with interested OTAs and the cognizant Joint Program Office (JPO)
- Report deficiencies identified in the system under test in accordance with this manual
- Coordinate FD/SC Charter development

MCOTEA evaluates the results of the MOT along with information from previous assessments in accordance with this manual and the MOT&E Memorandum of Agreement (ATEC 2012). MCOTEA coordinates the evaluation with the other Services and documents the results in an OER. The results are forwarded, as required, to the DOT&E, ACMC, MDA, and PM.

Other Service OTA Lead

When another Service OTA leads the MOT&E, Marine Corps input is either fully integrated within the TEMP or included as a Marine Corps appendix in the TEMP. MOT&E programs require the same rigor of planning, execution, and reporting as Marine Corps led programs. In order to provide the necessary rigor, teams will need to employ the MCOTEA 6-Step Process throughout the program. When the effort is redundant with the lead OTA, MCOTEA may leverage the lead OTAs efforts; however, all process steps must be addressed. For instance, an Operational Task Analysis is necessary to determine the Marine Corps missions and tasks. If the lead OTA conducts a comprehensive joint Operational Task Analysis then MCOTEA can leverage that effort, however, if they do not include the Marine Corps missions and tasks then MCOTEA may have to conduct a separate event to capture this information. Another area where MCOTEA will often have to provide a separate annex is when considering the evaluation method. It is not unusual for other OTAs to omit the evaluation method in the planning documents. In this case, MCOTEA may need to incorporate a separate annex with the Marine Corps evaluation method clearly defined. It is always in the best interest of all services to integrate the planning, execution, and reporting of MOT&E programs. However, the bottom line is that MOT&E programs require the same rigor as Marine Corps led programs. In order to support this effort, MCOTEA requires In-Process Reviews (IPRs) for all critical elements of the planning and reporting phases (i.e. Operational Task Analysis, Issues and Screening Criteria, Measures, Analytical Model, Decision Model, Integrated Data Requirements List, Data Reduction and Analysis Methods, Test Data and Evaluation Reporting). These IPRs will ensure that MCOTEA's input has been approved prior to submission into the MOT&E planning and reporting documents.

Types of Assessment and Testing Performed by MCOTEA

System Assessment
- ACAT IV (M)
- AAP
- Quick Reaction Assessment
- Non-programs of record

Intermediate Assessment
- DT Observation (ACAT IV (T) & above)

Operational Assessment
- Early Operational
- Operational

Operational Testing
- Initial
- Follow-on
- Multi-Service

MCOTEA Assessments

MCOTEA conducts three types of assessments: system, intermediate, and operational. A System Assessment is based on a SAP, while Intermediate and Operational Assessments stem from a SEP. An assessment provides a "progress report" on a system, not a "final grade," which would be OE/OS/OSur.

The following characteristics are common to all assessments:

- Contractors may be used to operate and maintain the system
- Use of production-representative articles is not required
- Technology demonstrators, prototypes, mock-ups, engineering development models, or simulations may be used
- OE/OS/OSur is not determined

- Assessment may have its own dedicated MCOTEA-led test event

Complete guidance about MCOTEA assessments is contained in chapter 3.

Top-Level MCOTEA Functions

The top-level functions performed by MCOTEA (SECNAV 2011) are as follows:

- Ensure that the OT of all ACAT I, IA, II, III, and IV(T) programs is effectively planned, conducted, evaluated, and reported
- Coordinate the scheduling of resources for operational testing requiring Marine Operating Forces support through Force Synchronization Conferences and the Two-Year Master Test Plan
- Provide input to the TEMP, Parts II–IV
- Prepare an OER within 45 days after completing IOT&E and provide directly to the ACMC
- Assist program acquisition by conducting Early Operational Assessments and Operational Assessments on request
- Assist program acquisition by collaboratively planning and participating in integrated test events, observing DT events and providing Observation Reports, and conducting Assessments throughout the acquisition cycle
- With the PM, decide the number of system articles to be procured for Initial Operational Testing for all Acquisition Programs not on the OSD T&E Oversight List
- Coordinate with Marine Operating Forces and other commands in matters related to OT&E by publishing a Feasibility of Support message
- Be the primary interface with JITC on Joint interoperability testing conducted during operational testing
- Manage those OSD-directed Multi-Service OT&Es for which the Marine Corps is tasked
- Coordinate Marine Corps support for other Services' OT&Es
- Effectively represent the Marine Corps in all Multi-Service OT&E matters

3 MCOTEA's 6-Step Test & Evaluation Process

Entry of New Work into MCOTEA
Program SharePoint Support
Plan-Test-Report
Integrated Testing Within the 6-Step Process
Step 1: System Evaluation Plan
Step 2: Test Concept and TEMP Input
Step 3: Test Planning
Step 4: Operational Test Execution
Step 5: Operational Test Data Reporting
Step 6: System Evaluation and Reporting

Chapter 3-0. The 6-Step Process

This chapter describes in detail the 6-step process MCOTEA uses to perform test and evaluation once a new program has entered the Activity. These introductory pages present an overview of the complete process. Each step is then presented in detail from the perspective of integrated testing. A section explaining how Assessments work within the 6-step process is presented at the beginning of the chapter.

Entry of New Work into MCOTEA

Any requests from external organizations for MCOTEA's assistance in developing new programs, including those that arrive before program funding is provided, are reported to the Deputy, who works with Division Heads to determine the appropriate level of MCOTEA's support and the Division that will execute the work.

Early requests typically come from CD&I, the Materiel Developer, or the RTT in support of early collaborative planning to include drafting of COIs and participation in the Capabilities Documentation IPT, where MCOTEA reviews capabilities documents and Concepts of Operation/ Employment.

Once the Deputy has reviewed the request, he will document the program acceptance and assignment in a Program Initiation Memorandum to the appropriate Division.

Early and periodic interaction between MCOTEA Test Divisions and Program Management Offices/Program Managers is expected and encouraged to help forge productive working relationships.

Program SharePoint Support

MCOTEA maintains a Microsoft SharePoint site that supports program activities from inception through closeout.

The MCOTEA SharePoint site comprises four functional parts: a Home Page (including access to all Executive, Staff, and Division areas), a data repository capability, an active program working area, and a Records Management document library.

The MCOTEA Home Page is the jumping off page for all other SharePoint capabilities. This page provides links to all other MCOTEA SharePoint sites, pages, and libraries.

The data repository capability is available to Divisions upon request. The creation of a separate data repository is only necessary when data sets are too large to manage through other means such as e-mailing spreadsheets. Teams must coordinate with the S-4 early in the test planning step in order to develop the data repository site.

The active program working area is a Division site used to manage active programs and the associated working documents. The Division coordinates site development with the S-4 upon the receipt of a Program Initiation Memorandum. Once developed, the Division manages the site with assistance from the S-4.

Once Division documents are staffed for signature and finalized, they will be elevated to the Records Management document library (a link may be left on the Division site for reference).

The Records Management document library is the single source of final documents (including final data results) at MCOTEA. Any document signed by the Director or by direction of the Director or a Division Head is an official record that MCOTEA must maintain. The S-1 maintains a Records Management File Plan that identifies the types of official (and certain unofficial) records maintained at MCOTEA as well as instructions for the proper storage and disposition of records.

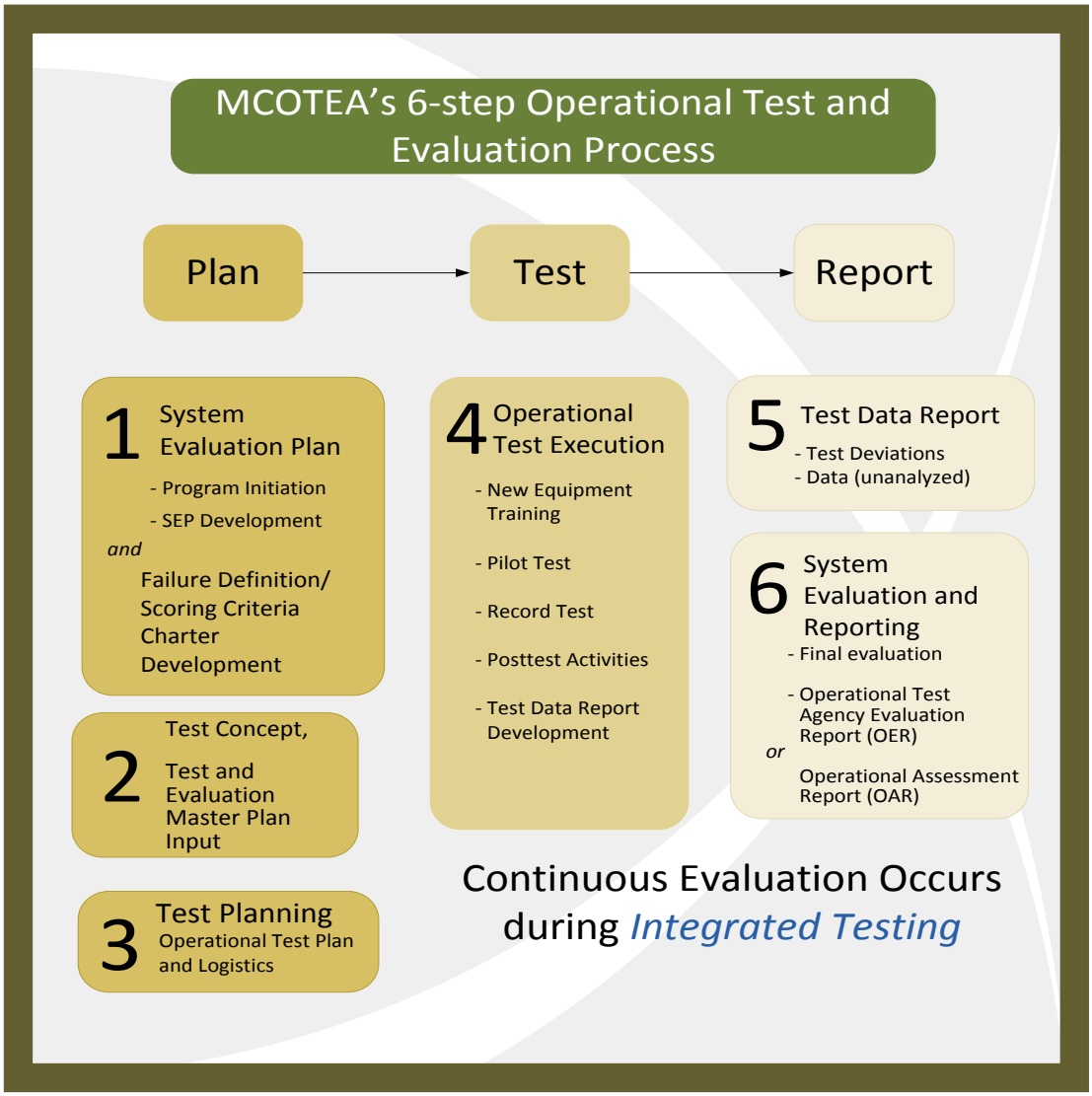


Figure 3-1. MCOTEA's 6-Step Process

Staff Responsibilities

The S-1 oversees the MCOTEA Records Management File Plan and document library. At a minimum, the S-1 Lead

- Maintains and archives MCOTEA records in accordance with the Marine Corps Records Management Program (MCO 5210.11)
- Maintains the official records of the Activity
- Conducts an annual review (usually in January) of MCOTEA's records

The S-4 oversees the MCOTEA SharePoint. At a minimum, the S-4 Lead

- Facilitates the development of the SharePoint site
- Maintains permissions and requests for access
- Provides configuration management across the organizational site (may assign as subtasks to users of SharePoint areas as appropriate)
- Conducts regular content reviews to ensure knowledge is useable and current

The Division Heads are responsible for following the MCOTEA staffing workflow to ensure that all official records are captured for records management purposes. In addition, Division Heads are responsible for

coordinating with the S-4 to develop the Division active program working area and data repository (data repositories are used during test execution only).

The OTPOs are responsible for tracking and submitting Milestone documentation in order to maintain a complete programmatic record. See the S-1 Records Management Lead for a complete list of required Milestone documents and the process for submitting.

Plan-Test-Report

MCOTEA organizes its test and evaluation process into six steps (fig. 3-1), grouped in a Plan-Test-Report arrangement. The evaluation process spans the entire arrangement.

Proper evaluation can only result from the accumulation of data and facts about a system over its acquisition life cycle, not from a single operational test. An overarching approach assures decision makers that MCOTEA's final report is wholly credible and defensible because it is based on evaluated test results spanning the program's history.

The System Evaluation Plan (SEP), developed in step 1, is MCOTEA's plan for evafor evaluating results from specific types of assessments and operational tests. The SEP also "feeds" the Test Concept, and the TEMP.

Details of test trials and test logistical needs are accounted for in step 3, Operational Test Planning, leading directly to Test Execution in step 4. By this time, all assessments performed as part of integrated testing are concluded.

Steps 5 and 6 produce the test and assessment/evaluation reports. The TDR provides an early look at test data, the assessment provides a system progress review, while the evaluation analyzes the results in depth and provides decision makers with an OE/OS/OSur determination.

These six steps, grounded in the scientific method and applied consistently across all programs, ensure a substantial and thorough test and evaluation process.

Integrated Testing Within the 6-Step Process

MCOTEA's primary mission is OT&E, but considerable effort is also devoted to integrated testing, discussed in detail in chapter 2. In terms of MCOTEA's 6-step process, integrated testing occurs primarily between steps 2 and 3, before IOT commences (fig. 3-2, facing page). MCOTEA may use or perform various assessments to provide information about a system's progress towards IOT or to gather data to fulfill evaluation requirements established in the SEP.

MCOTEA Assessments

MCOTEA's process is heavily dependent on performing assessments and analyzing their results to support a system's overall evaluation. This section provides an overview of the different types of MCOTEA assessments. MCOTEA conducts three types of assessments: System, Intermediate, and Operational, as defined in the following sections. Assessments occur either as standalone events (System Assessment) or as pre-IOT events (Operational Assessment.) Common to all assessments are the following characteristics:

- Contractors may be used to operate and maintain the system
- Use of production-representative articles is not required
- Technology demonstrators, prototypes, mock-ups, engineering development models, or simulations may be used
- OE/OS/OSur is not determined

The results of any assessment are sent to the PM and MDA and may be distributed further at the discretion of the Director, MCOTEA. See chapter 4 for reporting

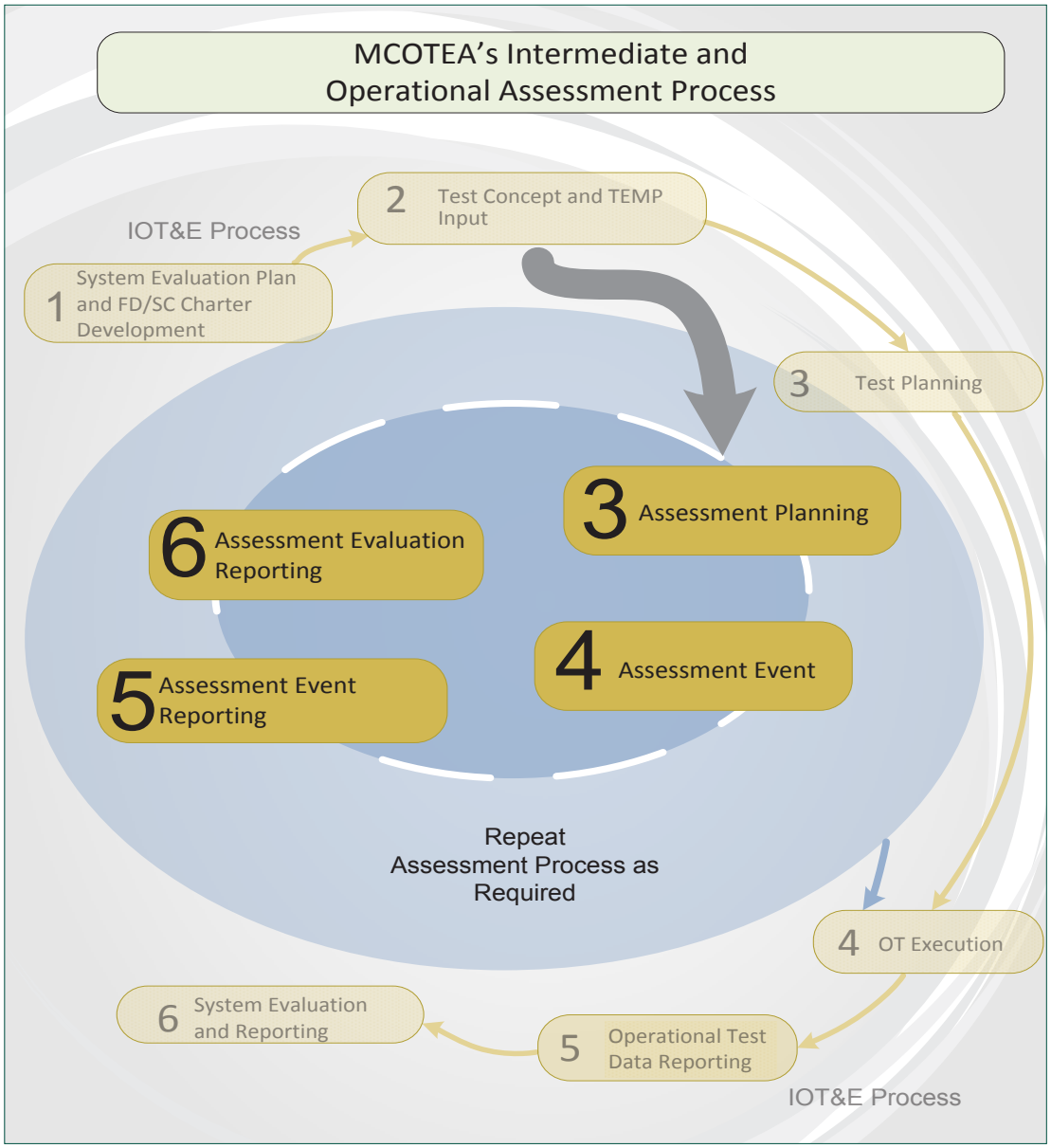


Figure 3-2. Intermediate and Operational Assessment Process

requirements and deadlines.

System Assessments

As noted in the introduction to this chapter, System Assessments pertain to programs being tested or examined that do not require operational test, such as Quick Reaction Assessments (QRA), Abbreviated Acquisition Programs (AAP), ACAT IV(M) programs, and other ACAT programs. MCOTEA uses this type of assessment to answer specific questions to address risk areas, as written in the SAP.

To begin the System Assessment process, MCOTEA writes a SAP, which serves as a framework and methodology for performing the assessment and provides basis for eventual analysis of assessment data. After performing the System Assessment, MCOTEA documents the assessment in a System Assessment Report (SAR).

Figure 3-3 provides a "menu" of possible ways for MCOTEA to be involved in System Assessments, along with the products that each type of involvement yields. Using

this table, MCOTEA works with the program sponsor to identify the exact nature of MCOTEA involvement in the System Assessment.

Quick Reaction Assessment

When a system must be fielded quickly an Urgent Operational Need Statement (UONS) or Urgent Universal Need Statement (UUNS) is typically issued for the system in development, or the system may be granted Rapid Deployment Capability (RDC) status by ASN (RDA). This urgency may necessitate modifying established MCOTEA OT&E processes in order to rapidly procure and deliver the urgently needed capability. In such cases, the program sponsor may request a QRA from the Director, MCOTEA. The QRA request should include the following:

- Purpose of the System Assessment and the specific system attributes the program sponsor wants assessed
- Time available for the System Assessment
- Concept of Employment
- Any available threat documentation
- Resources available for the System Assessment
- Forces that will deploy with the system before IOC

Execution of a QRA does not replace the scheduled operational testing as approved in the TEMP for ACAT Programs. Systems in RDC status, as approved by ASN (RDA), will normally undergo formal OT&E when they transition to program status.

AAPs and ACAT IV(M)

By definition, AAPs and ACAT IV(M) programs (the M stands for "monitor") do not require operational testing. They do, however, require a MCOTEA endorsement to obtain their designation. As part of the designation process, the PM requests from the Director, MCOTEA, a written endorsement of the proposed acquisition strategy.

Intermediate Assessments

Intermediate Assessments pertain to programs at the ACAT IV(T) (Test) level and above. They are governed by a SEP and are most commonly performed after DT Observation.

DT Observation

MCOTEA normally observes DT events to verify that the DT event was executed according to plan and to verify DT data results after receiving the DT report. Properly performed DT Observation enables MCOTEA to use DT data in

Figure 3-3. MCOTEA's Options for Involvement with System Assessments

Non-Programs of Record, AAPs, ACAT IV(M), and QRA						
	Concur with request for no OT; no further MCOTEA program involvement	MCOTEA only observes the testing to ensure a quality test is executed	MCOTEA-led event with no assessment	Provide assessment based solely on DT	Provide assessment based solely on MCOTEA-led test	Provide assessment based on both DT and a MCOTEA-led test
	Operational Task Analysis (OTA) not required				OTA required	
ACAT Designation Letter*	X	X	X	X	X	X
SAP				X	X	X
Observation Plan		X		X		X
Observation Report		X		X		X
Test Plan			X		X	X
Test Data Report			X		X	X
SAR				X	X	X

* For AAPs and ACAT IV(M) programs

overall system evaluation. In addition, MCOTEA's participation gives the PM insight into the system's developmental progress, materiel maturity, and readiness to enter a MCOTEA-led assessment or operational testing phase.

DT Observation events are specified in the TEMP. To prepare for attending the DT event, MCOTEA prepares a DT Observation Plan for internal use, based on the evaluation questions from the SEP pertinent to this event. MCOTEA may also participate in collaborative planning of the DT event, but only DT personnel execute the events under DT observation.

During developmental testing, system components are checked to ensure that they function as designed, and the system is checked to ensure that it meets the requirements derived from the ICD/CDD/CPD. MCOTEA generally uses the data gathered during DT to determine if the thresholds in the approved capabilities documentation have been demonstrated. In addition, aggregating DT data over time can be useful in determining a system's OS and OSur.

As with any assessment, OE/OS/OSur is not determined. After the DT Observation, MCOTEA writes an Observation Report and later after receiving the data in a DT Report, an IAR. The PM and MDA use the IAR to gauge a program's progress toward IOT and to become aware of any risks to program success.

Operational Assessment

MCOTEA may conduct an Operational Assessment (OA) to demonstrate selected system performance, with user support as required. An OA can range from a "paper assessment" to a physical operational test. The nature of the OA is described in the TEMP. An OA can be conducted at any time, but is normally done during the Engineering and Manufacturing Development phase of the acquisition cycle to evaluate selected Issues, KPPs, and other system attributes. An OA typically focuses on significant trends noted in developmental efforts, programmatic voids, areas of risk, testability of capabilities, and the ability of the program to support adequate operational testing. An OA does not determine OE, OS, or OSur. Any program on the DOD Oversight List must attain acceptable performance in an OA before entering the Production and Deployment phase (DOD 2008). An OA provides early information to the PM and/or decision maker about system progress in the following areas:

- Satisfying capabilities documentation
- Satisfaction of defined Attributes including KPPs and KSAs
- Readiness for LRIP
- Readiness for entry into IOT

Characteristics of an Operational Assessment include the following:

- May also be used to support program reviews or milestones
- May be conducted using technology demonstrators, prototypes, mock-ups, engineering development models, or simulations; production-representative articles not required
- May use typical users (Marines) as operators
- May be conducted under actual operational conditions
- Does not substitute for IOT&E needed to support full-rate production decisions

Early Operational Assessment

An Early Operational Assessment (EOA) is similar to an OA, but is conducted during the Technology Development phase of the acquisition cycle, before MS B, and is typically used as an input to determine whether a system should continue development and proceed to Engineering and Manufacturing Development.

Figure 3-4 summarizes assessment and test characteristics.

Figure 3-4. Assessment and Test Characteristics

Characteristics	Assessments				Tests
	System	Intermediate	Operational		IOT/FOT/MOT
	QRA, AAP, and ACAT IV(M)	DT Observation	EOA	OA	
May use technology demonstrations, prototypes, and mock-ups	X	X	X	X	
Production-representative models required					X
May use Marine Operators	X	X	X	X	
Must use Marine Operators					X
May use contractors to operate or maintain the system	X	X	X	X	
May be conducted under actual operational conditions	X	X	X	X	
Must be conducted under actual operational conditions					X
Does not substitute for IOT	X	X	X	X	
Uses representative forces (both friendly and opposing)					X
Employs realistic tactics and targets whenever possible					X
Determines OE/OS/OSUR					X

STEP 1 System Evaluation Plan

Task		Responsible	Provides Input	Task Output	Signature /Release Authority
Initiate Program		Exec	DH	Program Initiation Letter	Exec
Assign Team Members		DH	OTPO/OTAD	Team Assignment Letter	Exec
Develop and Write POA&M		OTPO	ORSA	POA&M	DH
Participate in Capabilities Development IPT		OTPO		Meeting Minutes	
Review ICD, CDD, CPD		OTPO	Military SME, ORSA	CRM	DH
Participate in T&E Working-level Integrated Product Team (WIPT)*		OTPO		Meeting Minutes	
Help develop IPT Charter				IPT Charter	DH
Develop SEP or SAP**					
	Write	ORSA	Military SME, OTPO, MS	Draft SEP	
	Review/Comment	OTPO	ORSA, DH, Exec	CRM	
	Adjudicate CRM & Prepare for CRB			CRB/signature-ready SEP	
	Prepare post-CRB Copy			Signature-ready SEP	
	Approve			Signed SEP	Exec
Develop M&S Accreditation Plan *** See Volume II M&S Chapter for details.					
	Write	ORSA	OTPO	Draft Accreditation Plan	
	Review/Comment	OTPO	ORSA, DH, Exec	CRM	
	Adjudicate CRM			CRB/signature-ready Accreditation Plan	
	Prepare post-CRB Copy			Signature-ready Accreditation Plan	
	Approve			Accreditation Plan	OTAD
Develop Failure Definition/Scoring Criteria Charter***					
	Write	ORSA	TM, ORSA, OTAD	Draft FD/SC Charter	
	Review/Comment		OTAD, DH, Exec	CRM	
	Adjudicate CRM			CRB signature-ready Charter	
	Prepare post-CRB Copy			Signature-ready Charter	
	Approve			Signed Charter	DH or Exec
Develop RAM Evaluation Plan***					
	Write	ORSA		Draft RAM Eval Plan	
	Review/Comment	OTPO	DH, OTAD, Exec	CRM	
	Adjudicate CRM	OTPO		CRB/signature-ready RAM Eval Plan	
	Prepare post-CRB Copy	OTPO		Approval-ready RAM Eval Plan	
	Approve	OTPO		Approved RAM Eval Plan	Exec

*WIPTs are continuous through step 6.

**SAP is written primarily for a Quick Reaction Assessment but is also used for ACAT IV(M)s, AAPs, and non-traditional programs of record.

***If required for program

Chapter 3-1. Step 1: System Evaluation Plan

Evaluation Purpose

MCOTEA's evaluations support stakeholders with information for pending decisions or validate decisions already made. Before beginning to develop an evaluation plan, the evaluator should understand the evaluation's exact purpose. A common purpose for a MCOTEA evaluation is to support the acquisition process through the determination of OE, OS, and OSur of materiel solutions.

MCOTEA's conclusions about OE, OS, and OSur are considered summative evaluations. The purpose of summative evaluation is to render a summary judgment on a system's performance (Scriven 1991). Summative evaluations determine whether the expectations for a system have been met. Their findings are intended for decision makers with major roles in system oversight. Such evaluations may influence significant decisions about the continuation of the system, allocation of resources, or restructuring. Therefore, summative evaluations must be based on information that is sufficiently credible under scientific standards to provide a confident basis for action and to withstand criticism aimed at discrediting the results (Rossi, Lipsey, Freeman 2004).

MCOTEA has the capability to evaluate nonmateriel solutions including Verification, Validation, and Accreditation (VV&A) of simulators and simulations; training methods; and TTPs. These nonstandard evaluations follow the same general process outlined in this chapter, even though terminology and evaluation questions may differ depending on the evaluation's purpose.

Benefits of Continuous Evaluation Reporting

- Reporting of information is timelier to the decision maker
- Evaluation products themselves are more focused on a smaller set of evaluation topics at greater depth
- Evaluation level can focus on the decision maker's needs at that phase of development
- System evaluation subsequent to operational testing can focus on mission performance rather than a combination of specification compliance and mission performance

Evaluation Paradigm: The Importance and Benefits of Continuous Evaluation

The evaluation of a system for OE/OS/OSur requires a wide range of data and information, more than can normally be derived from a single test event (Giadrosich 1995). The Defense Acquisition Guidebook recommends "an integrated DT/OT/LFT&E evaluation, using a phased approach that identifies key decision points and that generates timely and objective information for decision makers on the system's demonstrated capabilities to date." Furthermore, system evaluation plans should be prepared in recognition of the need for multiple assessments of the performance of a system under development. The information from the evaluations should be issued periodically throughout Integrated Test activities. This information provides a feedback loop to inform systems development and minimize the number of system faults that are discovered in late-stage operational testing (National Research Council 1998).

Much is learned about a system as it progresses through the developmental cycle. With a continuous evaluation approach the independent evaluator can assess the system's progress against standards appropriate for that phase of development. Early information about achievement of performance specifications is useful to the decision maker when the evaluation and information are provided with sufficient time to react and affect changes in design. The key point is that saving the evaluation of developmental test data for independent evaluation later in the developmental cycle when the operational testing occurs negates the point of the early information; information's usefulness diminishes as time passes. In short, to enable more timely use of information,

MCOTEA's independent assessments (and reporting) should occur as closely as possible to the test events generating the results. An increase in frequency of communication between the independent evaluator and the decision maker will increase the likelihood of positive changes in a system's design.

Review Program Documentation

The test team must develop a thorough knowledge of the system, the system's mission, the threat to the system and threat tactics, and the way its operators will employ the system to accomplish the mission. The team gains this knowledge by reviewing program documentation, including capabilities documents (ICD/CDD/CPD); the system's COE; the System Threat Assessment; the CONOPS; the Marine Corps Task List; unit TTPs; and any other relevant document that would help the team understand the missions associated with the system under test.

System Evaluation Plan (SEP)

The SEP is MCOTEA's plan for evaluating results from Intermediate and Operational Assessments and Tests. The SEP includes Background and Scope information as well as detailed evaluation Objectives and Methods. In addition, the SEP has two annexes, the FD/SC Charter and the RAM Evaluation Plan. These annexes are discussed in more detail later in this step (see the MCOTEA SharePoint templates library for a detailed template).

Background

The most important part of any plan is to ensure that all stakeholders clearly understand the problem statement. This is accomplished by providing a definition of the capabilities gap that the materiel solution is meant to address. In addition, this section contains the decision this plan is designed to support and concludes with a description of the system being evaluated. For purposes of the SEP, a system is defined as the Marine unit or crew and their equipment, which includes the materiel solution that will be used to accomplish missions. This is the case even if the exact composition of the materiel solution is not known when developing the SEP.

> ***SEP's Relationship to the TEMP***
>
> The SEP "feeds" the TEMP. In particular, the SEP provides input to the TEMP's Evaluation Framework. (Note that the SEP also contains a section called "Evaluation Framework," which is not the same as the TEMP's.) Although many of the TEMP's lower-level details are developed in approximately the same timeframe as the Test Concept (Step 2 in this Manual), the test team must consider the TEMP throughout the entire process. As with most steps in the MCOTEA process, the SEP and TEMP do not occur as discrete, serial events, but develop concurrently. For more details on preparing TEMP input, see step 2.

The description of the materiel solution and the system users will most likely come from the capabilities documents or urgent need statements. These documents provide descriptions of the materiel solutions to include the necessary KPPs, KSAs, and other Attributes for the system that are necessary to design and build a materiel solution. These documents also provide the quantities of systems to be fielded and the units who will receive them. Systems are composed of components, attributes, and relationships described as follows:

- Components are the operating parts of a system consisting of input, process, and output. Each system component may assume a variety of values to describe a system state as set by control action and one or more restrictions (Blanchard, Fabrycky 1990).
- Attributes are the properties or discernible manifestations of the

Systems Viewpoint

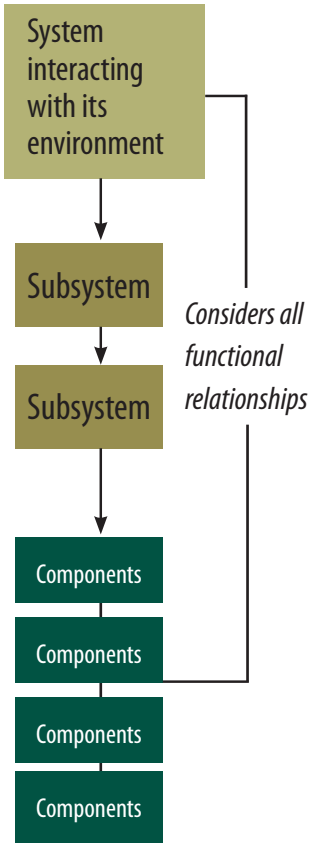


Figure 3-1-1. Hierarchical Depiction of a System

Uses of Operational Task Analysis

- Issues for Evaluation Framework
- Basis for mission essential functions
- Tasks necessary for training evaluation
- Defining the process flow for trials

components of a system (Blanchard, Fabrycky 1990). Attributes characterize the system; DOD further defines them as a testable or measurable characteristic that describes an aspect of a system or capability (DAU 2012).

- Relationships are the links between components and attributes.

A system is a set of interrelated components working toward some common objective. The objective or purpose of a system must be explicitly defined and understood so that system components provide the desired output for each given set of inputs.

Scope

This section contains the evaluation timeframe, a summary of operational missions, and the evaluation boundaries to include the system's position in the hierarchy of systems, and the system's functional relationships.

Every system is made up of components, and any component can be broken down into smaller components. If two hierarchical levels are involved in a given system, the lower is conveniently called a subsystem (Blanchard, Fabrycky 1990) (fig. 3-1-1).

In any particular situation it is important to define the system under consideration by specifying its limits or boundaries. Everything that remains outside the boundaries of the system is considered to be the environment. However, no system is completely isolated from its environment. Material, energy, and/or information must often pass through the boundaries as inputs to the system. In reverse, material, energy, and/or information that passes from the system to the environment is called output. That which enters the system in one form and leaves the system in another is usually called throughput (Blanchard, Fabrycky 1990).

The systems viewpoint looks at the system from the top down rather than from the bottom up. Attention is first directed to the system as a black box that interacts with its environment. Next, attention is focused on how the smaller black boxes (subsystems) combine to achieve the system objective.

The lowest level of concern is then with individual components. Focusing on systems, subsystems, and components in a hierarchy forces consideration of all pertinent functional relationships. Components and attributes are important, but only in that the purpose of the whole system is achieved through the functional relationships linking them (Blanchard, Fabrycky 1990).

Objectives

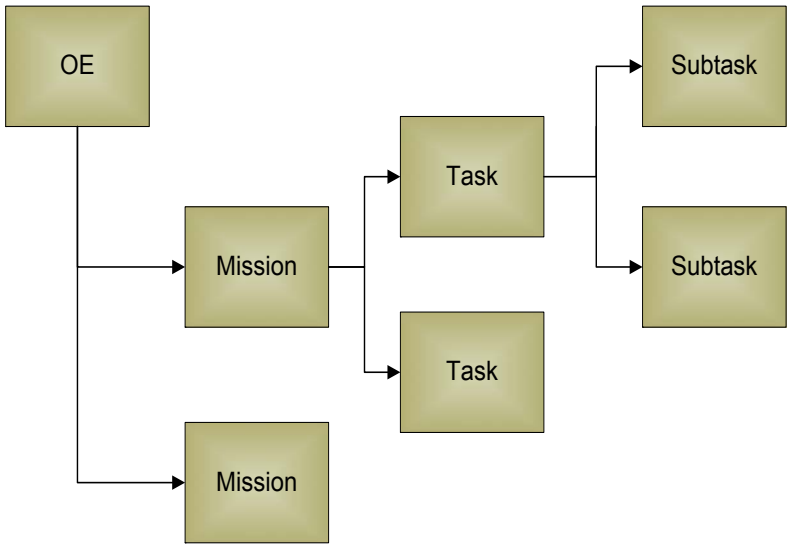
[Note: before beginning the Evaluation Framework, the test team should check the MCOTEA Lessons Learned database for helpful suggestions.]

At the top level of the hierarchy are the missions (COIs). Subordinate to the COIs are Tasks, followed by Subtasks, etc.

For OE/OS/OSur evaluations, each Task and Subtask represents an action to be accomplished by equipment, personnel, facilities, software, or any combination thereof. Each Task and Subtask also

represents a potential evaluation question. As seen in figure 3-1-2, the evaluation hierarchy flows from left to right. Added to that is a top-to-bottom addition of suitability and survivability characteristics as appropriate under each Task and Subtask. This hierarchy of COIs, Tasks, Subtasks, and their associated Issues forms the basis for the Evaluation Framework.

Figure 3-1-2. Evaluation Hierarchy



Missions form the basis for the COIs used to resolve OE/OS/OSur. Tasks, Subtasks, and suitability/survivability characteristics form the basis for the remainder of the evaluation questions (i.e., Issues) that support COIs. Answering the Issues associated with these Subtasks and Tasks at early stages of system development, if possible, provides assurances to the decision maker that the system is progressing as expected. Logically speaking, it is desirable to demonstrate the capability at a Subtask level before attempting the Task level.

Operational Task Analysis

MCOTEA uses Operational Task Analysis (OTA) as the analytic backbone of the Evaluation Framework. Task Analysis supports evaluations by breaking down complex evaluation problems into more manageable parts. OTA provides a disciplined method for developing the framework for evaluation questions below the level of OE/OS/OSur. OTA is top-down and mission-based. The methodology that follows can also be applied to evaluations of AAPs, ACAT IV(M)s, QRAs, and other non-programs of record performed by MCOTEA.

Identify Missions

The first step in identifying applicable missions is to start with the system's capabilities documentation supplemented by the Marine Corps Task List. An SME panel can be helpful in determining applicable missions for the system. The ultimate goal for identifying missions is to develop them into the COIs.

Identify Tasks

The next step in the top-down analysis is to identify the fundamental Tasks the system is expected to accomplish in each mission. These Tasks constitute the discrete actions that must occur to accomplish the mission

Other Sources for Task Identification

- Concept of Operations
- Concept of Employment
- Universal Naval Task List (USMC 2007) and/or the Universal Joint Task List (Joint Staff 2011)
- Mission Essential Task Lists of units that will employ the system under test or currently employ similar existing systems
- Mission documentation containing relevant TTPs
- Training manuals and battle books
- Subject Matter Expert panels
- DOD Architecture Framework (Operational View 6c)

(including suitability characteristics such as maintenance, transportation, and storage). These Tasks are founded in the capabilities the system is intended to address; therefore, the existing capabilities documentation is consulted initially. In fact, the capabilities documentation may state some Tasks explicitly. Since this step is accomplished early, the capabilities documentation can be supplemented with other authoritative sources. Determining the Tasks lays the foundation for the Evaluation Framework. The focus at this point should be on the Tasks that are required as opposed to how the Tasks will be accomplished. Determining how a Task is accomplished is the Materiel Developer's responsibility (when it comes to the materiel solution) using operational TTPs. At the end of this step, all Tasks by nature will be tied to at least one parent COI.

Identify Lower-Level Subtasks

At this point, the Tasks are subdivided into lower-level Subtasks. Like Tasks, these supporting Subtasks constitute the discrete actions that must occur to accomplish the Task. Some Subtasks may be associated with more than one Task; these should be listed with each appropriate Task. Subtasks are a means of identifying what operators must do to accomplish their missions, but at a lower level of indenture than Tasks. As with the Tasks, all Subtasks must be rephrased into a question (an Issue) to clarify the evaluation's intent. It may be necessary to go another level deeper into the Subtask hierarchy (the Sub-Subtask), but in general, the first level of Subtask should suffice. At the end of this step, all Subtasks will be tied by nature to at least one parent Task.

Figure 3-1-3 illustrates a completed OTA in block diagram format based on an example of a light armored anti-tank vehicle (hereafter called System ABC). Block diagrams efficiently document the decomposition of missions. OTA block diagrams are set up left to right so the top of the hierarchy is at the far left. At the top of the hierarchy is the system, defined as the Marine unit/crew and their equipment, which includes the materiel solution that will be used to accomplish missions. This is the case even if the exact composition of the materiel solution is not yet known. The remaining blocks are the Missions (COIs), Tasks, and Subtasks (Issues).

The OTA is a working document, and given its potential size, may be more efficiently used electronically rather than on paper. In any case, although the document is not a printed part of the SEP, it must be available for use and inspection in the official SEP files.

Once the OTA is completed, consideration should be given to the development of the Trial Flow, figure 3-1-4 (further defined in Step 2 of this Manual).

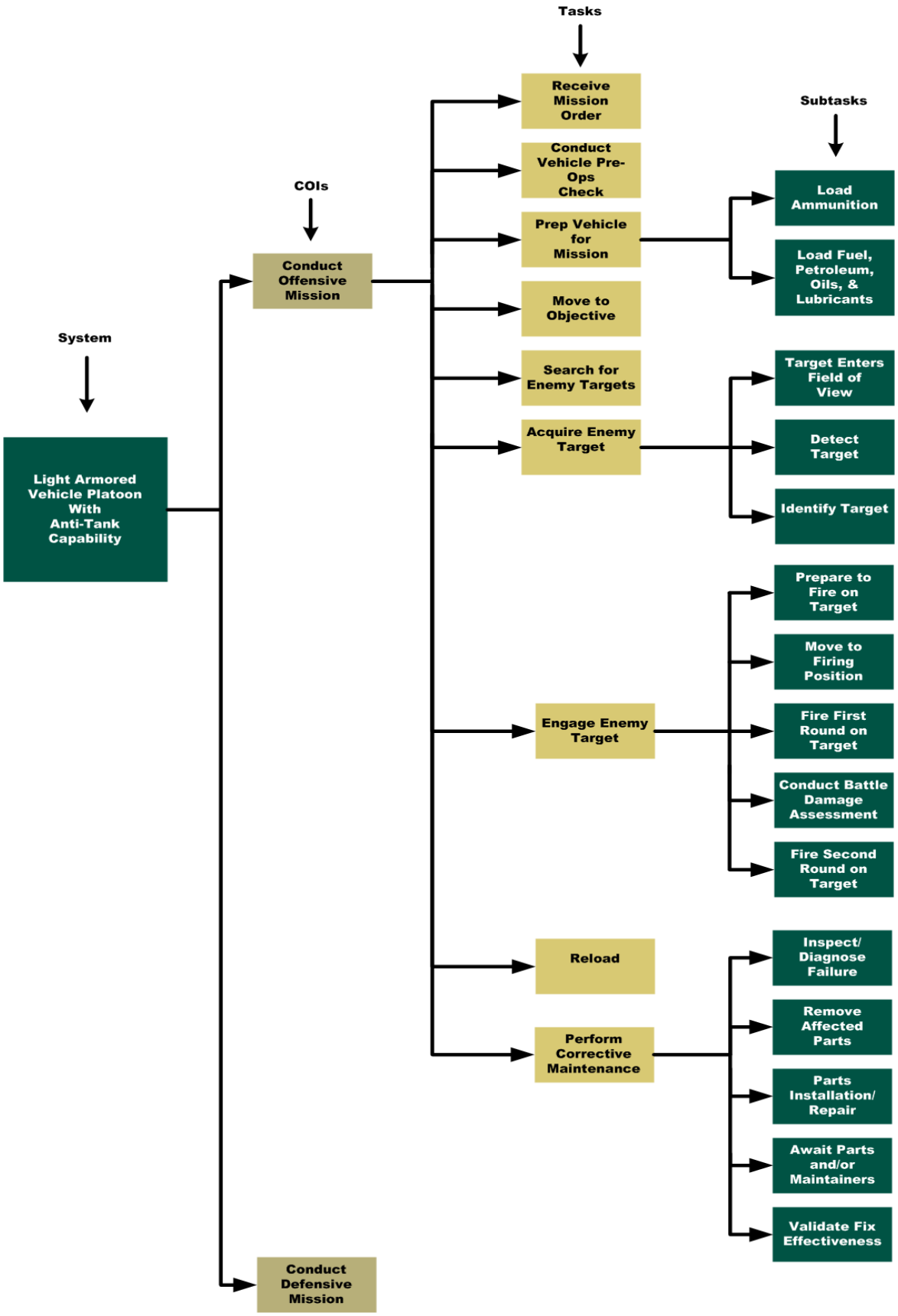


Figure 3-1-3. A completed Operational Task Analysis, which is always completed in a block diagram

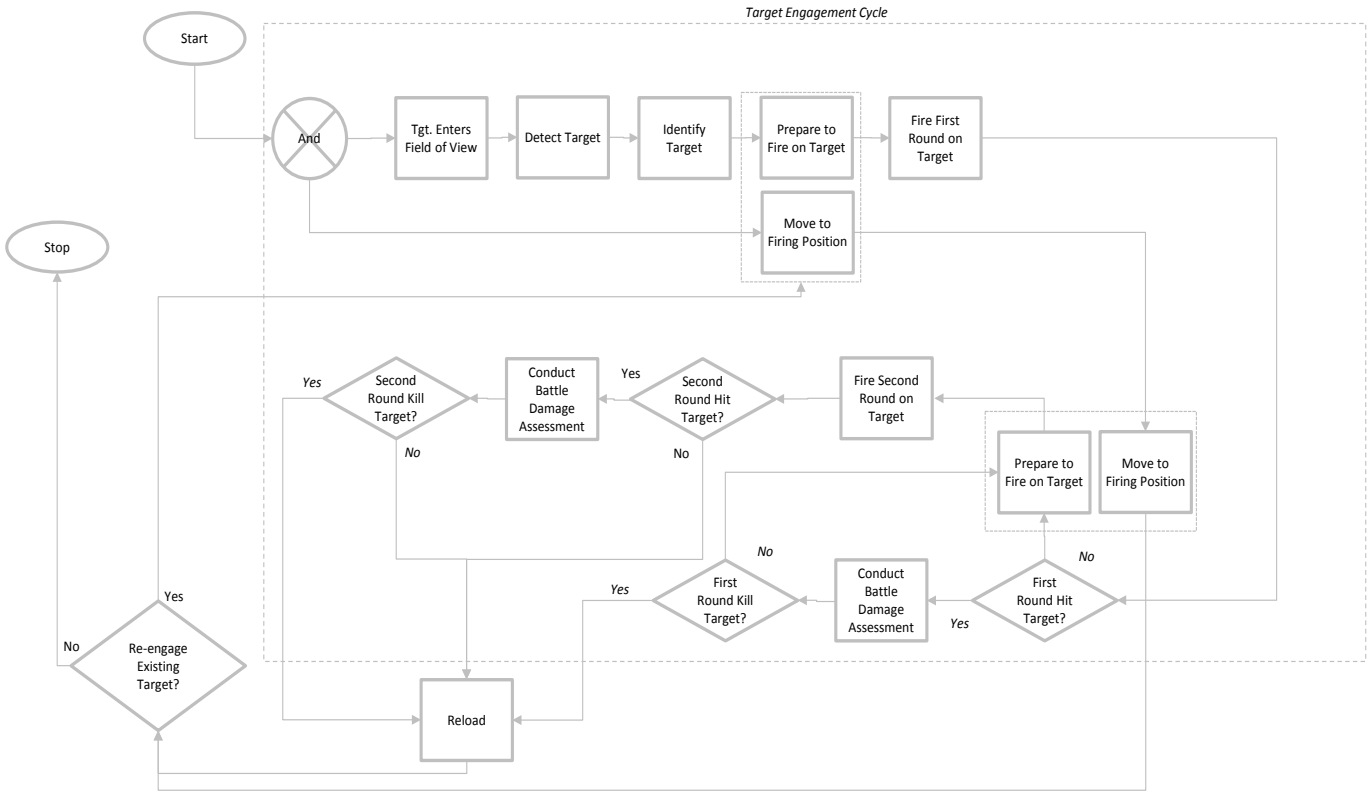


Figure 3-1-4. Beginning the Trial Flow Process for System ABC

Develop Evaluation Questions

The next step in any evaluation is to develop the evaluation questions. MCOTEA develops evaluation questions for acquisition programs that require OE/OS/OSur to be determined. In addition, MCOTEA performs a variety of nonstandard evaluations to support an array of decisions depending on stakeholder needs. The discussion that follows generally applies to any type of evaluation that MCOTEA might perform; however, the determination of OE/OS/OSur is only associated with IOT, FOT, or MOT.

Evaluation Questions

COIs and lower-level Issues are generically termed evaluation questions in this chapter. These represent operational questions that must be evaluated. COIs are mission-level questions, while Issues correspond to questions based on the Tasks and Subtasks associated with the system as well as Issues associated with aggregated suitability (e.g., Reliability, Availability, Maintainability, etc.) and survivability concerns. If a system is found to be Not OE, OS, or OSur, the Issues help to determine why.

Characteristics of Evaluation Questions

Evaluation questions should be operational in nature, observable, and testable (DAU 2012 and Clemen, Reilly 2001). Furthermore, evaluation questions must be answerable; they must involve performance dimensions that are sufficiently specific, concrete, practical, and measurable so that meaningful information can be obtained about their status (Rossi, Lipsey, Freeman 2004 and Clemen, Reilly 2001).

Formulating unanswerable evaluation questions without realizing it is easy to do. For example, "Does System ABC provide effective anti-armor support to the MAGTF?" is ambiguous: what does "effective" mean? How would an evaluator determine "effective fire support"? Evaluation questions may also include so

few observable indicators that little can be learned about them. For a question to be answerable it must be possible to identify some evidence (observables) that can realistically be obtained and that will be credible as the basis for the answer. Finally, the distinguishing feature of an evaluation question is its relationship to performance and its association, at least implicitly, with some criteria by which that performance can be judged (Rossi, Lipsey, Freeman 2004).

Top-Down & Mission-Based

OT&E follows a basic pattern of reasoning in its practice of evaluation. The Defense Acquisition Guidebook recommends that evaluators focus on the mission that a unit or crew will accomplish when equipped with a system and identify operational capabilities critical to mission accomplishment (DAU 2012). Doing so starts a "top-down" methodology leading to COIs, Issues, MOEs, critical LFT&E and other evaluation Issues, Measures of Performance (MOP), and data requirements.

OE/OS/OSur Interrelationship

OE, OS, and OSur are related hierarchically as seen in figure 3-1-5. OE is achieved through a combination of factors to include the performance of the system coupled with its suitability and survivability characteristics.

Examples of requirements for mission effectiveness can include the following:

- System is deployable to the mission theater (suitability)
- Operators know how to use the system properly (suitability)
- System performs as expected (performance)
- System does not adversely affect other mission equipment (suitability)
- System does not create a vulnerability to its operators or the operators of other systems (survivability)

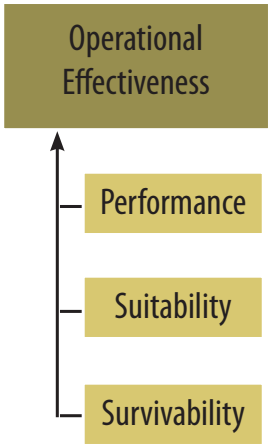


Figure 3-1-5.
OE Hierarchy

Operational Effectiveness

OE is an expression of the system's overall ability to accomplish its missions by typical users in the environment planned or expected for operational employment. Considerations include organization, doctrine, tactics, system performance, suitability, survivability, vulnerability, and threat.

MCOTEA is required to determine OE for systems that require IOT by law. OE is determined by measuring the effects or outcomes of the missions where a system under evaluation is being employed. The effect is unique for each system and depends on the missions in which the system is employed.

Operational Suitability

OS is the degree to which a system can be placed and sustained satisfactorily in field use (DAU 2011).

OS, like performance, forms the basis for the second tier of the evaluation questions below OE. MCOTEA is required to determine OS for systems that require IOT by law.

OS is determined by measuring the

Operational Effectiveness

Mission

COI-X. Can System ABC destroy a like-size armored enemy force during offensive operations?

Measure: Force Superiority

Threshold: ≥ 1.54

Task

I-X. Can the System ABC crews destroy enemy targets in less than or equal to 60 seconds after identification?

Measure: Target Engagement Cycle Time

Threshold: ≤60 seconds

Suitability Characteristic (subordinate to a Task)

I-X.X. Does the Tube-launched Optically-tracked, Wire command data link, guided missile (TOW) system onboard System ABC have a Reliability of at least 0.95?

Measure: Reliability

Threshold: ≥ 0.95

Subtask

I-X.X. Is the maximum effective range of the TOW system greater than or equal to 3,750 meters?

Measure: Max Effective Range

Threshold: ≤ 3,750 meters

suitability characteristics of the system and then determining what impact, if any, these characteristics have on the effects or outcomes of the missions. Suitability characteristics include the following:

- availability
- compatibility
- transportability
- interoperability
- reliability
- wartime usage rates
- maintainability
- safety
- human factors
- habitability
- manpower
- logistics supportability
- environmental effects
- documentation
- training requirements

Operational Survivability

OSur is the capability of a system and its crew to avoid or withstand a manmade hostile environment without suffering an abortive impairment of its ability to accomplish its designated mission considering the following (DAU 2012):

- electromagnetic environmental effects
- susceptibility
- vulnerability
- Information Assurance
- Chemical, Biological, Radiological, and Nuclear (CBRN) survivability

According to the OSD, "Typically, survivability testing for information and business systems will be based on information assurance" (OSD 2010). MCOTEA interprets this to mean the OSur component of the OE for information and business systems is based on the security, integrity, availability, confidentiality, authentication, and non-repudiation of the data that the system comprises.

Like performance and suitability, OSur forms the basis for the second tier of the evaluation questions below OE. MCOTEA is required by DOD instruction to determine OSur for systems that require IOT (DOD 2008). OSur is determined by measuring the survivability characteristics of the system, assuming realistic friendly and threat tactics, and then determining what effect, if any, these characteristics have on the effects or outcomes of the missions.

COIs for Joint, Multi-Service, and Oversight Programs.

Organizations external to MCOTEA (e.g., DOT&E, Other-Service OTAs) may recognize COIs that do not conform to the mission-oriented paradigm. These additional COIs should be added to the TEMP to support other organizations' evaluations, but not to MCOTEA's SEP. The SEP represents the USMC-oriented perspective on the system being evaluated. The Measures associated with the additional COIs may be added to the SEP as appropriate.

Critical Operational Issues

The system's operational activities (i.e., missions) form the basis for MCOTEA's COIs. The goal is to obtain an initial set of COIs early enough so they are available for use by the AoA. After additional review, the COIs will eventually be used in mission-based testing to help determine OE/OS/OSur.

COIs should be stated generally in most cases but can be written more specifically when a test is relatively simple. For example, a COI for a complex test of the Ground/Air Task Oriented Radar (G/ATOR) asks, "Can the operators using the G/ATOR system perform Air Surveillance and Control of Aircraft?"

For a relatively straightforward test such as for the anti-tank vehicle, a COI reads, "Can a Light Armored Reconnaissance platoon equipped with anti-tank variants destroy a like-sized armored enemy force during offensive operations?"

Issues

Evaluations are focused on answering questions. Issues are defined as any aspect of the system's capability, either operational, technical, or other that must be questioned before the system's overall military utility can be known (OSD 2008). Issues in the evaluations are categorized in two basic ways: Tasks/Subtasks and suitability/survivability.

Tasks and Subtasks

Tasks and Subtasks are means to identify what the operators need to do to accomplish their missions. All Tasks and Subtasks result in questions (i.e., Issues) to clarify the evaluation's intent. See the previous sections in this chapter on Tasks

and Subtasks to review the details of their characteristics.

Suitability and Survivability

Addressing suitability and survivability within the Evaluation Framework rather than in a separate dendrite helps illuminate and determine their overall impact to effectiveness at the mission level. Suitability and survivability are comprehensively examined by progressing through the hierarchy, beginning at the Subtask and moving to the Task level.

Not all Subtasks will result in an evaluation question. Some Subtasks, especially at lower levels of indenture, may not apply to the evaluation of the materiel solution. However, leaving them in the framework is useful to examining suitability and survivability.

For example, Subtasks required for mission accomplishment but that do not apply to the materiel solution can be used to identify equipment and actions pertaining to interoperability and compatibility. Using notional Anti-Tank System ABC as a simple example, information from the laser rangefinder must be exchanged with the fire control system for the weapon system to adjust to the appropriate elevation when firing on targets. Therefore, it is important to validate the interoperability of the two to ensure task accomplishment.

Another reason to include suitability and survivability in the Evaluation Framework involves their relationship with OE. Here is a simple example of this dependent relationship from the suitability perspective: if target kills are the desired effects for an anti-tank crew, but the turret malfunctions (Reliability), then the effect cannot be achieved.

From the survivability perspective, if the new turret has a highly reflective surface that readily reveals the vehicle's position to the enemy, and the vehicle is shot before accomplishing the mission, the desired effect cannot be achieved.

Figure 3-1-6 illustrates the incorporation of applicable suitability and survivability characteristics for a single Subtask, "acquire target." This process should be repeated for every Task and Subtask in the Evaluation Framework to identify potential suitability and survivability Issues that can affect Task or Subtask accomplishment.

Determine Level of M&S Support Required

An integral part of evaluation planning is to determine if M&S support is needed and how it will be used. At this point, the test team must have a general idea of the data that can be generated from testing. If additional data will be required for situations or environments that cannot be tested because of limited test asset availability, lack of time, test range limitations, cost, or safety considerations, M&S might be used to supply this data. Early in the SEP process, the test team must decide the candidate applications for M&S support (see M&S chapter, vol. II).

Construct Evaluation Standards and Measures

Any question to be evaluated needs two things: a standard for determining worth or value and a method of measurement. The process of identifying standards begins with mapping system Attributes to the Tasks and Subtasks in the OTA diagram. The process ends when each COI and Issue has a clearly defined, unambiguous standard for performance that can be observed, understood, and measured.

Developing Standards

The word "standard" is used generically to refer to thresholds or other defined ranges of acceptable performance. Thresholds are defined as a minimum acceptable operational value below which the utility of the system becomes questionable (CJCS 2012).

Attribute Variations

Attributes are defined as quantitative or qualitative characteristics of an element

"Orphaned" and Implied Attributes

During the Attribute tracing process, it is possible to find an Attribute that does not trace to any Task or Subtask. These "orphaned" Attributes may appear to have little to do with the OE/OS/OSur of the system, but the test team should ensure that orphaned Attributes are not indicating the need to identify additional Tasks or Subtasks.

The inverse is also possible; Tasks/Subtasks may not be associated with existing Attributes, which indicates the existence of implied Attributes; these will need to be identified.

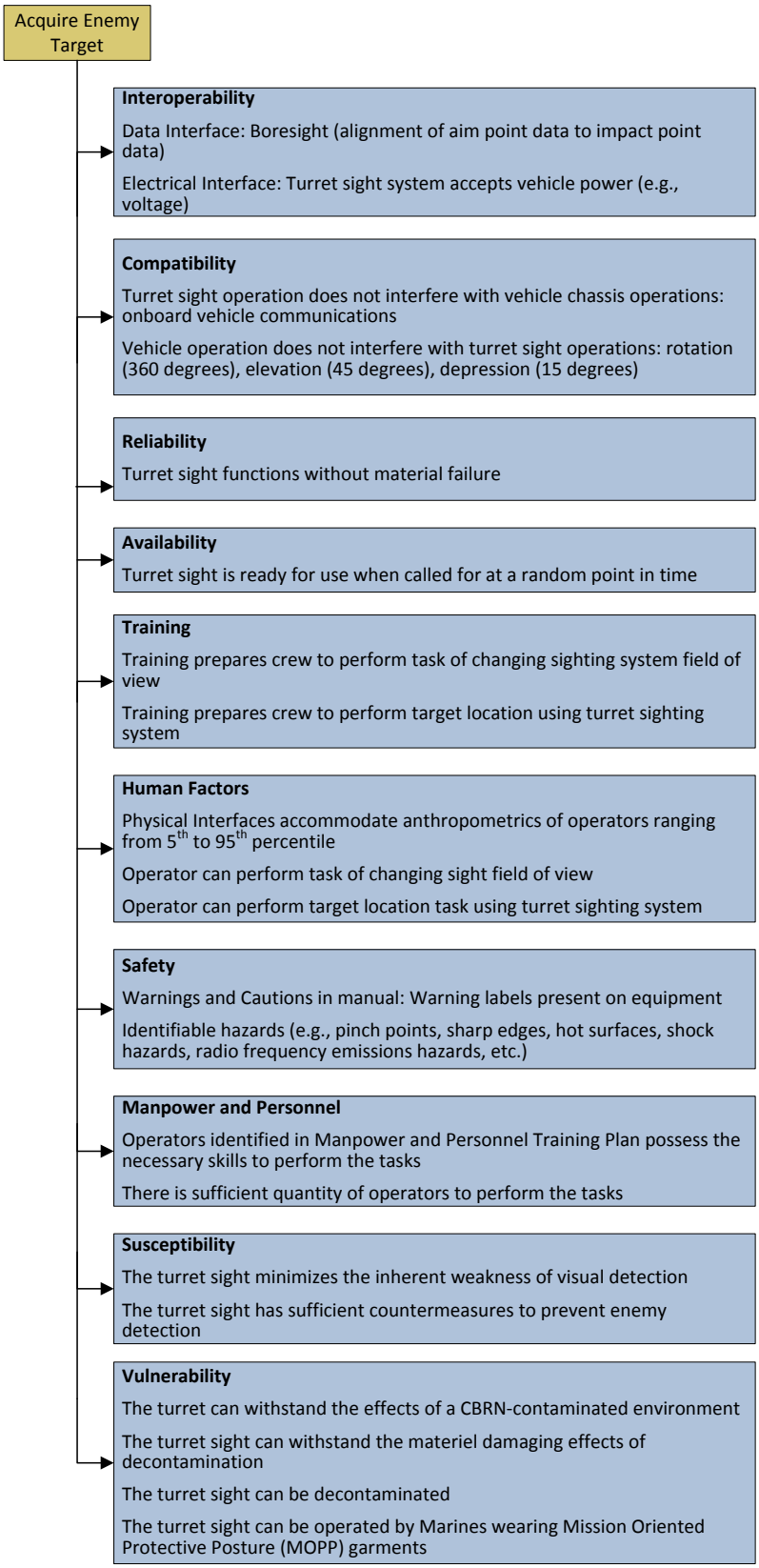


Figure 3-1-6.
Example of incorporating Suitability and Survivability Factors into Subtask

or its actions (CJCS 2012). The term Attribute is used here generically to refer to KPPs, KSAs, and other Attributes of the system outlined in capabilities documents. However, Attributes take many shapes and forms, and not all Attributes come from capabilities documents. Some Attributes are specified by law, regulation, or instructions. For example, DODINST 8500.2 provides Information Assurance Attributes. Figure 3-1-7 includes examples of Attributes from a single capabilities document, the Rapid Engagement Precision Rifle (REPR) CDD. The examples illustrate a variety of Attributes ranging from mandatory components to field use parameters.

Mapping Attributes to the Evaluation Framework

Attributes in the capabilities documentation should trace to Subtasks, Tasks, and COIs. The tracing process supports identification (and sometimes development) of standards for the COIs and Issues; in essence, the minimum acceptable outcome or effect.

The resulting Evaluation Framework links satisfaction of COIs to the capabilities identified in the JCIDS documents as the basis for accepting the system (CJCS 2012).

The tracing process is also useful for identifying the standards for Task/Subtask performance and the conditions under which Tasks/Subtasks are to be performed.

This process can help identify suitability and survivability standards as well.

At this point the capabilities documentation plays a prominent role. From the Materiel Developer's point of view the process of allocating requirements begins by assigning top-level system requirements to the various subsystems and lower-level elements of the materiel solution.

The evaluator views the allocation process differently. Since evaluation is concerned with task accomplishment, the Attribute mapping process occurs after the Missions, Tasks, and Subtasks have been determined with the Attributes mapped to the lowest level Subtasks/Tasks. When the Materiel Developer ultimately maps components and subcomponents to the Attributes in the functional analysis and MCOTEA traces these same Attributes to the Tasks and Subtasks, the link between the Capabilities Development, Materiel Development, and Operational Evaluation is complete.

Attributes Mapping Matrix

The Attributes Mapping Matrix is a working document that captures the work done to map Attributes to the Tasks and Subtasks. This matrix also accounts for any MCOTEA-derived implied Attribute and provides the references for developing standards. Given its potential size, the matrix is probably best used electronically rather than on paper. Like the OTA, however, the Attribute Mapping Matrix must be kept current and available in the official SEP files and filed in the T&E Records Management.

Establish Standards for Evaluation Questions

With Attributes mapped to the Evaluation Framework the evaluator can begin to establish standards.

Some standards may align directly with the accomplishment of the Issue or COI. For example, if the Issue at the Task level is to "engage enemy targets" and the Attribute mapped to it is Probability of Hit greater than or equal to 0.70, then the standard and Task are directly aligned. The evaluator should be aware that some Tasks and/or Subtasks may not have a standard that directly speaks to the accomplishment of the Task. In many cases the requirements speak to the critical technical parameters of the materiel solution rather than the capability itself. The evaluator must decide the nature of the evaluation to take place at

Figure 3-1-7. Types of Attributes

Note: Attributes for sample System ABC are not used here because the actual REPR system provides a ready-made and accurate set.

Attribute	Attribute Description, Threshold (T), and Objective (O)	Threshold Performance	Threshold Condition
Forward Assist	The REPR shall include a forward assist. (T = O)	N/A	N/A
Color	All external and visible REPR surfaces including magazines and suppressor shall have a dull finish that is paintable, consistent with current camouflage colors and patterns, and minimizes infrared signatures. (T)	N/A	N/A
Rail System	The REPR shall have a MIL-STD 1913 quad forward rail system that is integral to the upper receiver. The 12, 3, and 9 o'clock rails must be capable of maintaining sight zeros while conducting routine firing combined with combat movement and operational training drills. (T)	Maintain sight zeros (ambiguous)	While conducting routine firing combined with combat movement and operational training drills
Precision (KPP)	The REPR shall provide a precision of fire ≤ 1.0 minute of angle (MOA) at 800 meters when fired from an accuracy fixture in nominal conditions unsuppressed. (T)	Minute of Angle (MOA) ≤ 1.0	At 800 meters when fired from an accuracy fixture in nominal conditions unsuppressed
Hit Probability	A fully trained and current sniper firing the REPR shall achieve 8 out of 10 hits (80% probability) within 1.0 minutes of angle (MOA) at 800 meters firing 10 rounds in 10 minutes or less on an "NRA Bulls-eye" target under nominal conditions. Nominal conditions are defined as 70 degrees F ±10 degrees and unlimited visibility during daylight. (T)	8 out of 10 hits (80% probability) within 1.0 minutes of angle (MOA)	A fully trained and current sniper firing the REPR at 800 meters firing 10 rounds in 10 minutes or less on a "NRA Bulls-eye" target under nominal conditions. Nominal conditions are defined as 70 degrees F ± 10 degrees and unlimited visibility during daylight.
Trigger Pull	Pull weight shall not exceed 4 pounds. (T)	shall not exceed 4 pounds	N/A
Weight	Weight with scope, sling, bipod, suppressor, and magazine loaded with 20 rounds shall be 17 pounds or less. (T)	shall be 17 pounds or less	Weight with scope, sling, bipod, suppressor, and magazine loaded with 20 rounds
Multiple-Target Engagement	The REPR shall be capable of engaging 3 E-Type Silhouette targets (modified for MCMP Table II showing head, chest, and pelvic girdle scoring areas) placed 10 feet apart with one shot a piece in the head or chest scoring area at 500 meters in 15 seconds or less. (T)	15 seconds or less	The REPR shall be capable of engaging 3 E-Type Silhouette targets (modified for MCMP Table II showing head, chest, and pelvic girdle scoring areas) placed 10 feet apart with one shot a piece in the head or chest scoring area at 500 meters
Ergonomic Enhancements	The REPR shall have an adjustable stock and cheek-piece that shall accommodate shooter length of pull adjustments/optics alignment. The adjustable stock shall accommodate cheek weld, stock weld, and eye relief of the 5th-95th percentile of Marines. The stock must not interfere with the charging handle or cycle of operations of the weapon in any configuration. (T)	1. 5th-95th percentile of Marines 2. not interfere with the charging handle or cycle of operations	1. Cheek weld, stock weld, and eye relief 2. Any weapon configuration
System Ruggedness	The REPR shall perform reliably in High Temperature–160° F, Low Temperature–minus 25° F, Salt Fog, Sand and Dust, Icing/Freezing Rain, and after immersion in mud (T=O).	shall perform reliably (ambiguous)	High Temperature–160° F, Low Temperature– minus 25° F, Salt Fog, Sand and Dust, Icing/Freezing Rain, and after immersion in mud
Engagement Ranges	The REPR shall be capable of engaging targets between 300 and 800 meters.	shall be capable of engaging targets (ambiguous)	between 300 and 800 meters

every level of the operational task hierarchy (see sidebar).

At lower levels in the hierarchy, evaluation by proxy may be sufficient to mitigate risk. Evaluation by proxy does not directly measure the ultimate objective. For example, measuring the number of tanks killed could be a proxy for measuring success in battle. Evaluating the Task or Subtask directly may also be impractical, in which case evaluation by proxy is again acceptable.

The Subtask "identify target" from figure 3-1-3 provides a simple example of evaluation by proxy. If the Task is for the crew to acquire enemy targets, then evaluating the probability of target identification or time to identification could be an acceptable way to indicate the crew's ability to accomplish the Task. In this case the standard for the critical technical parameter becomes the standard for the evaluation question for this Subtask.

In more complicated circumstances, development of a standard for an evaluation question may be the result of piecing together multiple requirements from lower-tiered Subtasks to arrive at a COI or higher-tiered Task threshold. The technique for accomplishing this may be an analytic model, discrete system event model, numerical analysis, or stochastic model. For example, determining the standard for a COI where System ABC (i.e., Blue Forces) is required to destroy a like-sized armored enemy force (i.e., Red Forces) during offensive operations requires the use of analytic modeling. For the Blue Force platoon to triumph over the Red Force platoon, they must be more effective and/or have numerical superiority.

Figure 3-1-8 defines the engagement between the two forces in terms of their effective firing rates and their force sizes. Assuming the forces' sizes are equivalent, what remains is the effective firing rate, or the rate at which each force can attrite the opposing side. The effective firing rate can be solved for by taking the reciprocal of the time to kill a target on the opposing side, otherwise known as the Target Engagement Cycle (i.e., a_{blue}=1/TEC).

$$\phi_0 = \frac{M_{blue}}{M_{red}} \sqrt{\frac{a_{blue}}{a_{red}}}$$

Where

Φ_0 = the Force Superiority parameter

M_{blue} = the Blue Force size

M_{red} = the Red Force size

a_{blue} = the effective firing rate for the Blue Force = Reciprocal of the Target Engagement Cycle (TEC) time (i.e., $1/TEC_{blue}$)

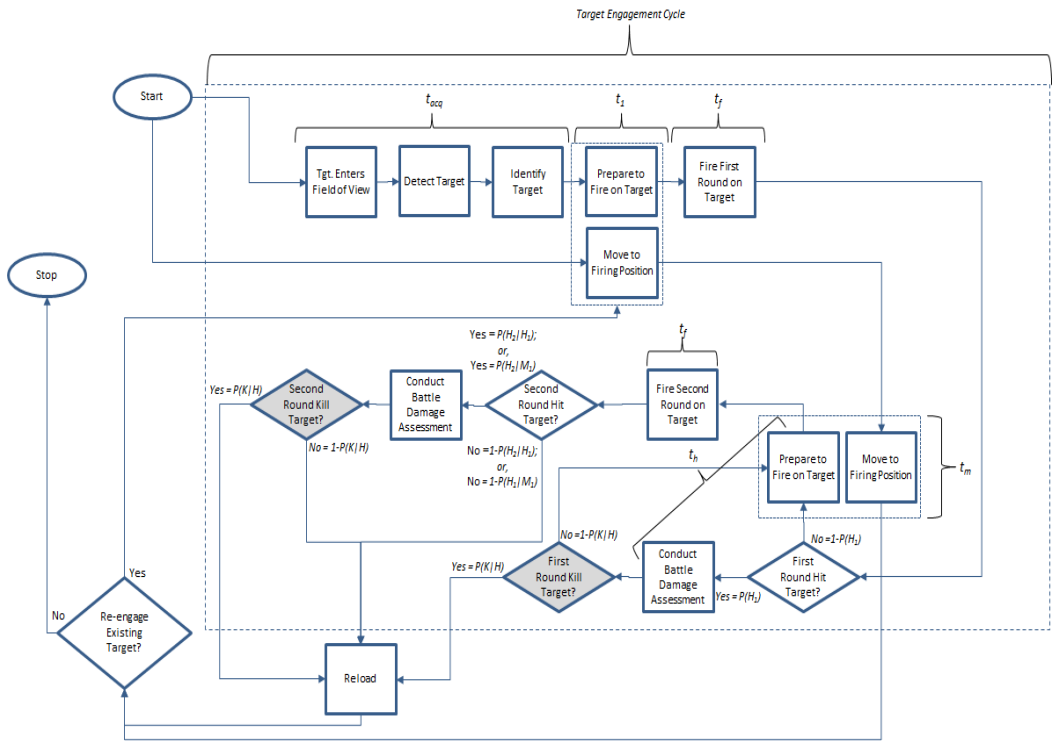
a_{red} = the effective firing rate for the Red Force = Reciprocal of the Target Engagement Cycle (TEC) time (i.e., $1/TEC_{red}$)

Figure 3-1-8. Example of Mathematical Expression Relating Mission- level Effects

Given this formula, TEC time can be further expressed by using the thresholds for probabilities of hit and kill; the time to acquire, fire, and assess damage; and missile flyout time (fig. 3-1-9).

$$E[TEC] = t_{acq} + t_1 - t_h + \frac{t_h + t_f}{P(K|H)} + \frac{t_m + t_f}{P(H_1|M_1)}\left[\frac{1 - P(H_2|H_1)}{P(K|H)} + P(H_2|H_1) - P(H_1)\right]$$

Figure 3-1-9. Example of Mathematical Expression Relating Measures to Process Flows



Standards for Performance and Conditions

The standards sought are for performance and for the conditions under which the performance must take place as noted in figure 3-1-10. The conditions encountered may affect the performance of a Task or Subtask.

Conditions can be the result of the physical environment (e.g., sea state, terrain, weather), the military environment (e.g., forces assigned, threat, command relationships), or the civil environment (e.g., political, cultural, economic factors (USMC 2007)). Operational conditions should be determined and associated with Tasks and Subtasks as appropriate. For example, the Attribute "hit probability" for a sniper rifle maps to the operational Task "engage targets" and forms the basis of the performance threshold. The Attribute "System Ruggedness" also maps to that Task, but serves as the basis for the threshold conditions for achieving hit probability.

The process of tracing Attributes has the unintended consequence of identifying gaps in the capabilities documents that must be filled for a successful evaluation.

Examples of MOPs

- Probability of detection
- Ammunition expenditure rate
- Rounds to adjust
- Casualties per dose
- Percent of tasks satisfied
- Time to adjust
- Range to detection
- Operator opinion (rating)
- Onload time
- Offload time
- Embarkation time
- Fuel consumption
- Radioactivity

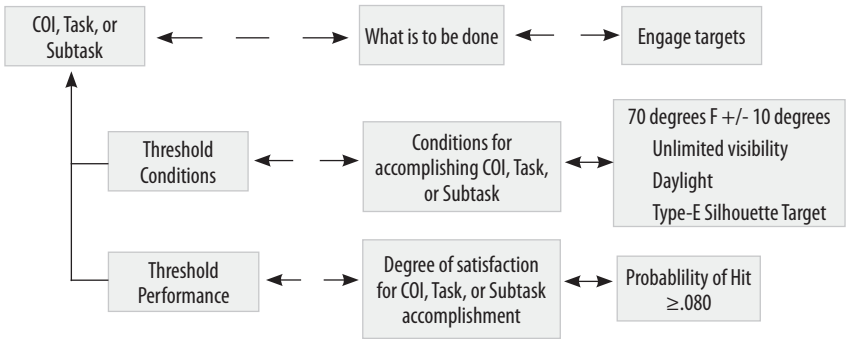


Figure 3-1-10 Relationship of Threshold Performance and Conditions to COI, Task, or Subtask

Using the previous example with hit probability and system ruggedness, the threshold for hit probability in operational conditions is not clear. The nominal conditions (70 degrees F ± 10 degrees) defined under hit probability do not agree with system ruggedness conditions (see figure 3-1-7).

The apparent disagreement leads to the following question: "What is the threshold probability of hit under other-than-nominal conditions?" In the process of deriving evaluation questions based on Tasks and Subtasks, the test team will find the need for standards that do not appear in the capabilities documentation. The team brings any questions to the attention of the capabilities officer early in the acquisition cycle, while the capabilities documentation remains in draft. The test team may use an SME panel, ideally including the DC, CD&I Action Officer for the program, to determine preliminary value for these standards. Any standards not clearly stated in the capabilities documentation will be identified as an evaluation assumption in the SEP.

Not all Attributes are measurable or have identified standards. If all attempts to establish standards for an Attribute are unsuccessful then MCOTEA may choose to characterize an Attribute rather than directly evaluate it. In these cases MCOTEA should identify the Attribute as a nonevaluatable requirement and document it as a limitation in the evaluation plan. Next, MCOTEA needs to identify a method to characterize performance relative to the requirement without levying a final conclusion about satisfaction.

Finalize Evaluation Questions

While the stage is now set for the evaluation questions, the work is not complete. Good evaluation questions will, when possible, convey the performance criterion or standard that is applicable as well as the Measure that is at issue (Rossi, Lipsey, Freeman 2004). Each evaluation question identified to this point should now be tailored to incorporate the standard and Measure.

Each question identifies what is of concern, how well it should be accomplished, and the dimension of measure. When an evaluation question lacks one of these critical elements, these shortcomings must be identified as early as possible and brought to the capabilities officer's attention though the Capabilities Development Integrated Process Team.

Developing Measures

Measures are needed to gather the data to satisfy the evaluation questions. The Measures dictate, at least in part, the data that needs to be gathered as part of the test event. The Measures will also be used later in the test design process to determine what factors (also called variables) will be varied and controlled in the testing process.

Types of Measures

The types of Measures relevant to system evaluations are Measures of Effectiveness (MOE), Measures of Performance (MOP), Measures of Suitability (MOS), and Measures of Survivability (MOSur). MOEs are needed to establish the system's military worth and value, while MOPs and MOSs are needed to design, build, and support the system. MOSurs are used to determine how well the system and its operators can survive to accomplish their mission in a combat environment.

Properties of Measures

The evaluator must consider three initial properties of MOEs, MOPs, MOSs, and MOSurs when selecting the best Measures for evaluation. These properties are reliability, validity, and sensitivity.

- Reliability is the extent to which the Measure produces the same result when used repeatedly to measure the same thing
- Validity is the extent to which the Measure succeeds at measuring what it is intended to measure

Fundamental Measures

- Power
- Area
- Flow
- Volume
- Torque
- Pressure
- Angles
- Frequency
- Temperature
- Velocity
- Distance
- Acceleration
- Mass
- Force
- Energy

- Sensitivity is the extent to which the values of the Measure change when a change or difference occurs in the thing being measured

An effective Measure conveys essential information without ambiguity or excess wording, both of which detract from a clear understanding of what data is required for test and evaluation. Examples of fundamental Measures that focus on essential information include the examples in the upper left sidebar.

Examples of MOSs

- Time between failures
- Time to repair
- Maintenance ratio
- Availability
- Time between maintenance actions
- Time to perform preventive maintenance
- Logistics Down Time
- Time between unscheduled maintenance actions

Measures of Effectiveness

An MOE is designed to correspond to the accomplishment of mission objectives and achievement of desired results. Generally, only a small number of generic MOEs are available to support system evaluations. Evaluation Measures are typically limited to:

- Percents of total events of a specific nature
- The time it takes for a specific event to occur
- The range at which specific events occur
- A qualitative assessment of specific events

Depending on the Issue (evaluation question), MOEs may be decomposed into MOPs, MOSs and MOSurs.

Measures of Performance

A MOP measures a system's performance expressed as speed, payload, range, time-on-station, frequency, or other distinctly quantifiable performance features. MOPs may have a greater number of observable phenomena to measure than are available for MOEs. Observable phenomena for MOPs include (but are not limited to) those mentioned for MOEs above plus the examples in the MOP sidebar.

Examples of MOSurs

- Probability of system detection by threat
- Probability of system hit given detection
- Probability of system damage given a hit
- Probability of casualties given a hit
- Probability of working countermeasures
- Reaction time to threat
- Probability of system defeating the threat
- Probability of information systems compromise

Measures of Suitability

An MOS measures an item's ability to be supported in its intended operational environment. An MOS typically relates to readiness or Operational Availability, and hence Reliability, Maintainability, and the item's support structure.

Measures of Survivability

An MOSur examines the degree to which using the system in combat places the system itself , the operators, or other systems/operators at risk. For information and business systems, survivability is interpreted as the ability of the system to maintain the confidentiality, availability, integrity, authentication, and nonrepudiation of the system's data.

Preferential Measures

MCOTEA has a preference for the types of Measures used in evaluations. In constructing Measures, the evaluator should consider a Measure's scale and its alignment with objectives.

Measure Scales

Measures are scaled as either natural or constructed. A natural scale Measure is one found in general use and having a common interpretation: "number of kills" is a natural scale Measure for lethality of a system. Natural scale Measures provide efficiency for the evaluator because they do not require scale definition. Their use may also be less controversial than constructed Measures because they are in general use. The difficulty is that natural scales may not fit the intended use, depending on what is being evaluated.

A constructed scale Measure is developed for a particular problem to measure the degree of attainment of an objective. Constructed scales are used in a variety of situations where natural scale Measures are not appropriate. Operator opinion (rating) Measures using Likert scales, for example, are constructed scale Measures.

Measure Alignment with Objectives

A direct Measure measures the degree of attainment of an objective, again using the example "number of kills." A proxy Measure reflects the degree of attainment of its associated objective, but it does not

directly measure the ultimate objective. For example, measuring the Gross National Product is a proxy for economic well being.

Measure Data Type

Measures can be continuous or discrete. A continuous Measure can take on an infinite number of values in an interval or collection of intervals; for example, the "distance from target" can be represented with an infinite number of values, depending on the precision of the instrument of measure. Continuous Measures provide more information and require fewer resources than noncontinuous discrete Measures. Discrete Measures may assume only a finite or countably infinite number of values; for example, the "number of fatalities" can only be an integer value (e.g., 1, 2, 3,…).

A discrete Measure with only two possible values is referred to as binary; for example, "Pass/Fail" is a binary Measure. When binary Measures are used, larger amounts of experimental resources are requried to evaluate a system process. Discrete (i.e., binary) Measures should be avoided whenever possible. Continuous Measures highly correlated to the binary response can be used in the analysis, resulting in large savings of experimental resources; for example, the "vibration of a device" (continuous) during processing can be highly related to whether the device will be "defective/nondefective" (binary).

Figure 3-1-11 depicts MCOTEA's preferences for types of Measures used, 1 being most preferred and 8 being least.

Figure 3-1-11. MCOTEA's Preference for Evaluation Measures

MCOTEA Measure Preference	Data Type				
	Continuous		Discrete		
Type of Scale	Direct	Proxy	Direct	Proxy	Alignment with Objective
Natural	1	3	5	7	
Constructed	2	4	6	8	

Several questions commonly arise in developing evaluation Measures:

- Should a natural scale, proxy, continuous Measure be used, or should a constructed scale, direct, discrete Measure be developed?
- Should an Issue (evaluation question) be subdivided into more detailed subconsiderations for which natural scales might exist, or should a scale be constructed to measure the evaluation consideration without subdividing it further?
- Should a natural scale that is precise but uses technical jargon be used, or should a constructed but possibly less precise scale be used that some stakeholders may understand more readily?
- How carefully should the scale definition for a constructed scale be specified?
- Can a continuous Measure be used to accurately portray the effectiveness of the system process or be used in conjunction with a highly correlated binary Measure?

Establishing Dominant Measures

A COI or other Issue (derived from a Task or Subtask) may have one or more MOE, MOP, MOS, and/or MOSur. When possible it is desirable to develop a dominant Measure for each evaluation question. A dominant Measure is a single Measure, which when evaluated, will consistently yield the same answer. When more than one Measure is needed for a COI or Issue, weights must be assigned to the relative importance of these competing MOEs for the decision maker's awareness unless both measures must be satisfied independently. Any COI or Issue with more than one evaluation Measure must also adhere to the principles of mutual exclusivity to avoid double counting. Said another way, if more than one evaluation Measure indicates the degree of attainment for a particular objective (that is, the evaluation Measures are redundant), then that objective will probably receive more weight than was intended when the weights are assigned to the various evaluation Measures either

explicitly or implicitly.

Methods

The ability of an evaluation result to withstand scrutiny rests in its foundation, the scientific method. An element of the scientific method is transparency of process, and an evaluation model with explicit methods provides that transparency. Furthermore, a transparent evaluation process can be repeated by others to confirm findings, and systems can be designed with the full understanding of the expectations that exist all the way up to the highest levels (OE) in a predictable manner. Predictability is important because it keeps evaluation expectations from becoming a moving target that is difficult and expensive to achieve.

Evaluation occurs in a continuum as the system is developed and test results become available. Early evaluations of the system (at the Issue level) consist of comparing the tested results for each Issue with its standard. At these early stages of evaluation when aggregation is not necessary, no need exists to construct an evaluation model. Generally speaking, evaluation models are necessary when some form of aggregation is required to collapse multiple components into a single evaluation answer, as with evaluations to determine OE/OS/OSur.

Properties of the Evaluation Model

The evaluation model is used to evaluate the system's test results to arrive at the evaluation conclusions, including OE/OS/OSur. The evaluation model is constructed to overcome inconsistency, a real and pervasive problem in human decision making. When asked to evaluate the same information twice, humans frequently give different answers (Kahneman, 2011). This problem can be especially prevalent in the OE, OS, and OSur determinations because they are summary judgments formulated using complex information, and humans can be highly inconsistent in making summary judgments under these circumstances. The model may employ a variety of techniques to aggregate and collapse the information across the dimensions of OE/OS/OSur in a manageable and understandable way. Most evaluations will employ some form of screening criteria, analytic model, and decision model to facilitate the system evaluation.

Screening Criteria

A screening criterion is a binding constraint on the system. The system must meet the screening criterion, the use of which can simplify the evaluation process (Kirkwood 1997). Screening criteria can reduce the number of Issues to only those essential for determining worth or value. A system that fails to meet minimum screening criteria should not proceed to later stages of the evaluation.

Aggregation Method

Care should be taken to aggregate only when necessary. Aggregation is necessary when multiple COIs exist in the hierarchy (i.e., a multimission system). Tasks and Subtasks can be evaluated and reported out individually as needed, in accordance with the TEMP, to support engineering and system progress reviews and to mitigate program risk. Although some Tasks and Subtasks may be evaluated individually to ensure that the system is ready for IOT, they may also be evaluated under operational test conditions with typical users and production-representative articles.

When a materiel solution begins to show performance shortfalls, tradeoff decisions must be made. These decisions are important, and aggregation and importance weighting are once again used to help resolve the issues.

Properties Necessary for Aggregation

When an evaluation contains more than

one COI, the need exists to enforce additional requirements, given the added complexity of the evaluation. One such complexity is the evaluator's ability to keep all of the COIs in mind at once, which is nearly impossible (Clemen, Reilly 2001). The accomplishment of one objective can also impede the progress of another (Clemen, Reilly 2001). The system under evaluation may have one or more competing objectives related to the COIs. For example, one mission for a system may require a high degree of off-road mobility, while another mission may require high levels of ballistic protection. Since increasing ballistic protection (i.e., adding the weight of armor) also reduces mobility, the two COIs have competing objectives. To address the complexities of multiple COIs, they should be collectively exhaustive, mutually exclusive, operable, and small in number (Parnell 2007). To complement the additional complexities of evaluating multiple COIs, screening criteria and weighting should also be used.

Collectively Exhaustive

An evaluation to support a decision is collectively exhaustive if it includes all aspects of a decision (Clemen, Reilly 2001, Kirkwood 1997, Parnell 2007). In other words, if the evaluation covers every mission required of the system as well as all relevant aspects of suitability and survivability, then the evaluation will be collectively exhaustive.

Mutual Exclusivity

Mutually exclusive COIs means that a given mission should be covered only once in the evaluation hierarchy (Parnell 2007, Kirkwood 1997). Overlap between COIs, especially when they are weighted, tends to overemphasize the importance of a particular dimension of the evaluation, sometimes referred to as "double counting" (Kirkwood 1997).

Evaluation Framework Operability

An example of an operable hierarchy is shown in figure 3-1-3, Operational Task Analysis. When using multiple COIs a tendency exists to continue to add evaluation considerations to achieve completeness, until the framework becomes so complex that any analysis using the framework will be difficult to conduct and interpret. In the quest for completeness, evaluators must balance the practical side, including cost and time to complete the analysis, within reasonable time limits (Kirkwood 1997). For this reason the COIs and MOEs should be few in number (DOD 2008, Kirkwood 1997, Parnell 2007). The COI and the accompanying Task/Subtask framework should be as small as possible without compromising needed detail. A smaller framework can be communicated more easily and requires fewer resources to estimate performance across the various evaluation Measures (Kirkwood 1997).

Keeping the evaluation framework as small as possible may seem to contradict the previous discussion on OTA. However, if the evaluations are accomplished over time, then each phase addresses relevant aspects of the framework rather than attempting to collapse information from the bottom to the top in a single evaluation. In this sense, taking the evaluations one layer at a time has the effect of making the evaluations smaller and more concentrated on the relevant characteristics of performance/suitability/survivability and keeps the evaluation focused on a single level of the system.

Evaluation Framework Weighting

Finally, multiple COIs vying for the evaluator's attention creates the need for weighting, which subjectively assigns relative importance to competing COIs according to the combat developer's priorities. In the earlier example of high

Quick Definitions for Evaluation Models

Collectively Exhaustive

The evaluation covers every mission required of the system as well as all the relevant aspects of suitability and survivability.

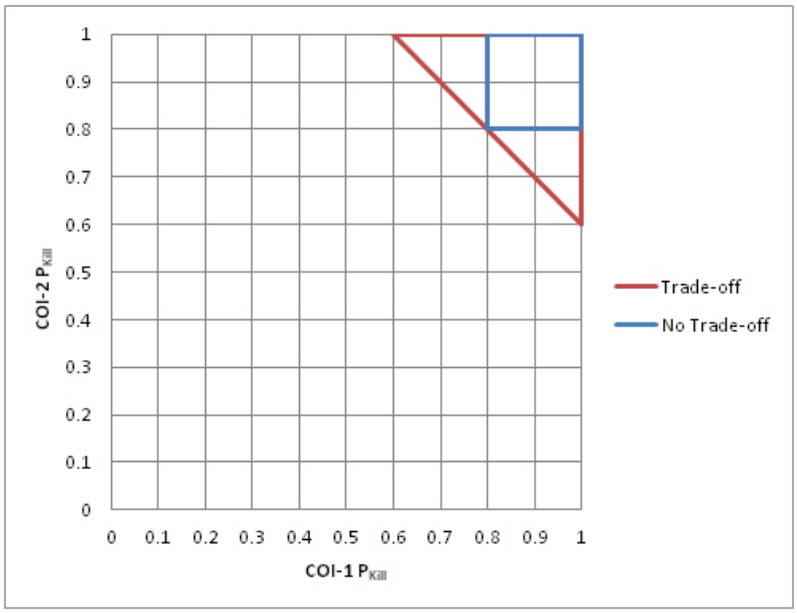
Mutual Exclusivity

The same objective should be covered only once in the evaluation hierarchy; no overlap should occur between the COIs.

mobility versus ballistic protection, the evaluator must know which requirement is more important and by how much. Weighting allows the evaluator to pay proper attention to the missions in terms of the combat developer's priorities.

Caution is in order when assigning weights because the assignment of weights implies that trade-offs can be made between COIs. It is inappropriate to automatically make this assumption without first validating that expectation with the Capabilities Developer. There are valid reasons for a system to be required to provide a minimum effect in multiple COIs without compromise. Simply put, when these situations arise it limits the range of feasible solutions that can be considered operationally effective. Figure 3-1-12 highlights the differences in the acceptable feasible region.

Figure 3-1-12. Example of Acceptable Feasible Regions for Trade-offs



The shapes within the figure represent operationally effective regions. The triangle represents the boundaries of the acceptable region where an OE solution exists for a system that allows trade-offs between different mission capabilities. The reduced region bounded by the square represents the boundary when no trade-offs between missions are acceptable. In other words, when no trade-offs between mission effects are acceptable, then the system under test must perform well in both missions rather than allow exceptional performance in one mission to offset substandard performance in another mission.

Building an Evaluation Model

Several key steps go into building an evaluation model to determine OE/OS/OSur. Most evaluations will employ screening criteria and analytic and decision models.

Identifying Screening Criteria

Screening criteria can be thought of as gates that force evaluations to occur in steps as the system matures or information becomes available. Not using screening criteria causes information to pool for a single, massive evaluation, which is cumbersome and inefficient.

Transportability is a common characteristic that may ultimately become a screening requirement. For example, certain systems are required to be transportable by CH-53E helicopters. If there is no other way to operationally deploy the system, this transportability requirement would be termed a screening requirement because inability to be transported by this platform would prevent mission accomplishment altogether.

In terms of evaluations, screening criteria can be considered binding constraints that can force a particular conclusion, such as "Not Operationally Suitable." In the example above, if the system cannot be transported by the CH-53E, then the system would automatically be evaluated "Not Operationally Suitable" regardless of performance in other suitability areas. The system should not proceed to operational evaluation to determine OE/OS/OSur until this requirement has been satisfied.

The failure to successfully satisfy screening criteria can be mitigated before the final evaluation of OE/OS/OSur in one of two ways. First, the system can be retested after appropriate fixes are in place to mitigate shortfalls. Second, the owner of the requirement can relieve the Materiel Developer of the requirement by modifying or abandoning it altogether.

Issues (both Task/Subtask and survivability/suitability) form the basis of screening criteria. Determining which Issues will become screening criteria and which decisions to apply them to is largely subjective, although mapping Attributes to Issues is valuable in determining screening criteria. An Issue mapped to a KPP or KSA may be a candidate for becoming a screening criterion. In addition, Issues (Tasks or Subtasks) that represent a critical path to mission success may also be selected as screening criteria.

Also subjective in selecting screening criteria is the timing of their use. Evaluations of system performance/maturity should occur over time. However, early screening criteria should be lower-level Issues, at the Subtask level, for example, in the Evaluation Framework. Issues identified as screening criteria may prevent systems from progressing past early Gate Reviews or Critical Design Reviews until their performance is satisfactory. Later, at the time of the OTRB, screening criteria may be used to ensure that the system is sufficiently mature for the rigors of operational test. Finally, some screening criteria, such as safety, may be used at any stage of the evaluation, because the effect on mission accomplishment may not be observable in an operational test.

Screening criteria that affect all COIs are considered global, while all others are considered local. Global screening criteria constrain the evaluation of OE/OS if not satisfied, while local screening criteria constrain the COI if not satisfied.

Thus, it is acutely important for the system evaluator to designate only those Issues critical to success and/or the decision maker as screening criteria; stated another way, not all requirements should be treated as screening criteria.

Local screening criteria influence one or more but not all COIs. Like global criteria, they are considered independent of affiliations that may exist with other screening criteria. Unlike global criteria, local screening criteria can be associated with more than one COI, in which case the criterion is considered independently for each COI and, in essence, is evaluated multiple times. Another difference from global criteria is that a failed local criterion affects only the applicable COI and not OE/OS.

Issues identified as screening criteria are noted in Annex A of the SEP in tabular format and are ultimately included in the TEMP. The logic for using screening criteria and their effect on the evaluation's outcome must be clearly identified, especially when combined with an analytic model for evaluation.

Constructing the Analytic Model

A model is a simplified representation of some aspect of the real world. Models provide a concise description of the essential features of a complex situation. Formal analytic models enable the evaluator to consider several variables simultaneously. By temporarily setting

aside unimportant variables, models serve as powerful tools for studying interrelationships among important variables.

Analytic models serve as tools for developing expectations for mission accomplishment when evaluation plans are drafted, and for performing post-test sensitivity analysis. In turn, the degree of mission accomplishment (or effect) depends on performance, suitability, and survivability, meaning that an analytic model for the COIs should incorporate all three of these dimensions. Incorporating suitability and survivability parameters into the analytic model is critical to determining their relative impact on effectiveness.

Simpler models are better than complicated models. The urge to over-populate a model with an abundance of parameters should be resisted because many parameters may have little or no real effect on the decision.

Under ideal circumstances KPPs and KSAs would populate the analytic model, but given the potential lack of consistency and specificity within various capabilities documents, this may be difficult. Top-level parameters such as Operational Readiness (OS parameter) and Probability of Hit (OSur parameter) are likely candidates for including in the model. Finally, selecting parameters to include in the analytic model always depends on the system being evaluated, as discussed in the following sections.

Constructing an Example Decision Model

The remainder of this chapter presents an example decision model highlighting key elements. The example presented illustrates the use of a Decision Analysis technique known as Multi-objective Decision Analysis with Value Focused Thinking (MODA with VFT) that will be continued throughout subsequent chapters; however, the example presented is not the sole means to construct a decision model.

The conclusions for OE/OS/Osur can be a direct result from COIs to a common scale, the Mission Capability Level (MCL), which is a score used to assess how well Marine operators using the system under test can be expected to fulfill their intended mission in a realistic environment. Arriving at MCL confers four distinct advantages to evaluation:

- Provides a systematic methodology for arriving at OE, OS, and OSur conclusion
- Allows the aggregation of Measures using different units by converting the measurement results to the dimensionless MCL value function
- Provides a framework for aggregation when multiple COIs (missions) are an element of the evaluation
- Normalizes evaluation results to a common scale (between 0–100), allowing decision makers responsible for multiple programs to assess capabilities across their portfolio in consistent terms

The 0-100 scale for MCL is divided into three intervals, defined by scores of 50, 80, and 100 (fig. 3-1-13). The 100-level score represents the capability corresponding to the system meeting all the objective values of the parameters in the COI analytic models. The score of 80 corresponds to the threshold values, while 50 corresponds to the current capability fielded for this mission, if a current capability exists.

Mission Capability Level	Range	
Fully Mission Capable	80	100
Partially Mission Capable	50	<80
Not Mission Capable	0	<50

Figure 3-1-13. Definitions of Mission Capability Level

The three intervals are defined more specifically as follows:

- **Fully Mission Capable** represents the highest section of the interval where a system scores at least 80. A system categorized as Fully Mission Capable means the system, in the mission context, has achieved at least the equivalent of threshold performance. A system must be fully mission capable to be considered OE.

- **Partially Mission Capable** represents the middle section of the interval where a system scores at least 50 but less than 80. A Partially Mission Capable system is considered to be at least as good as the current capability, but still falls short of the threshold. This categorical description only applies if a current mission capability exists and can be quantified. When the current capability does not exist, a system may still score between 50 and 80; however, the system is considered Not Mission Capable in this range.
- **Not Mission Capable** represents the lowest section of the interval where a system scores less than 50. A Not Mission Capable system does not improve on current mission capabilities. Fielding a system that scores less than 50 may still be justified by other aspects of the system, such as lower cost or overcoming technological obsolescence. The range for Not Mission Capable is expanded from 0–<50 to 0–<80 when no current mission capability exists for the missions the system is designed to address.

With MCL defined, the next step in building the Evaluation Model is to construct the mathematical functions for each COI to be used in deriving MCL results from the COI Measures. The functions can be curvilinear or, more commonly, piecewise linear, discussed here.

Constructing the piecewise linear function is relatively straightforward if standards have been established for the COIs. Certain data points are needed to construct the functions, including the standards that represent the objective, threshold, current capability values,

and a value to establish a zero point. Figures 3-1-15 and 3-1-16 constitute an example of the parameter values and the COI analytic model needed to construct the piecewise linear function.

Using the parameters from figure 3-1-14 with the equation in figure 3-1-15, a piecewise linear function can be constructed using points calculated for current capability, threshold capability, and objective capability. These lines are then plotted on a graph where the x-axis represents the MOE results and the y-axis represents the MCL scale.

Figure 3-1-16 illustrates the piecewise linear function for the data in figure 3-1-14.

The current capability point will only be used for missions where a current capability exists. An increase in capability expressed as a new mission area will not have a current benchmark for comparison; therefore, no Partially Mission Capable category will exist between 50 and <80. Instead, only two segments, 0 to <80

Figure 3-1-14. Example of Parameter Values for MCL Graph

Parameter	Description	Lowest Possible	Current Capability	Threshold Capability	Objective Capability
$R(t_m)$	mission reliability	0.40	0.55	0.75	0.90
k	probability of failure detection	0.90	0.90	0.90	0.90
$M(t_r)$	probability of a repair given a specified turn around time	0.50	0.65	0.85	0.95
M_o	operational maintainability	0.25	0.50	0.85	0.90
t_a	time to acquire the target	20	20	20	20
t_1	time to fire the first round	180	120	60	20
t_f	time of flight of the projectile from weapon to target	22	22	22	22
P(H)	single shot probability of hit	0.70	0.70	0.70	0.90
P(K|H)	probability of kill given a hit	0.90	0.90	0.90	0.90
t_m	time to fire a round given the preceding round was a miss	180	5	5	5
t_h	time to fire a round given the preceding round was a hit	180	5	5	5
$P(H_2\|H_1)$	probability of second shot hit given preceding round hit	0.90	0.90	0.90	0.90
$P(H_2\|M_1)$	probability of second shot hit given preceding round missed	0.90	0.90	0.90	0.90
E[TEC]	expected time for the target engagement cycle	314	174	114	68
POR(ta)	operational readiness given a specified turnaround time	0.67	0.81	0.94	0.99
D	dependability	0.55	0.78	0.96	0.99
m_{blue}	expected blue force size	2	4	5	6
a_{Blue}	blue force effective firing rate	0.003	0.006	0.009	0.015
$-a_{Blue}m_{Blue}$	attrition rate of red forces	-0.006	-0.023	-0.044	-0.087805
Φ(x-axis)	superiority parameter	0.37	1.00	1.54	2.40
MCL (y-axis)	Mission Capability Level	0	50	80	100

Figure 3-1-15. Example of Analytic Model Used to Calculate MCL

The expected force size for the Blue and Red forces which results from the product of initial number of forces (n) times the force system effectiveness (SE).

$$M_{blue}, M_{red} = E[M_{blue}], E[M_{red}] = n \times SE$$

The attrition rates for the Blue and Red forces, or reciprocal of the expected Target Engagement Cycle.

$$a_{blue}, a_{red} = 1/E[TEC_{blue}], 1/E[TEC_{red}]$$

The time to complete the Target Engagement Cycle from target acquisition to target kills for Blue and Red forces.

$$E[TEC] = t_{acq} + t_1 - t_h + \frac{t_h + t_f}{P(K|H)} + \frac{t_m + t_f}{P(H_1|M_1)}\left[\frac{1 - P(H_2|H_1)}{P(K|H)} + P(H_2|H_1) - P(H_1)\right]$$

System Effectiveness is the product of a system (vehicles) Operational Readiness and Dependability for a given mission.

$$SE = P_{or}(t_a) \times D$$

Operational Readiness is the likelihood of a system (vehicle) of a given force being ready for a mission when needed within a specified turnaround time.

$$P_{or}(t_a) = R(t_m) + \left[(1 - R(t_m))kM(t_r)\right]$$

Dependability is the likelihood of a vehicle of a given force completing a mission of a specified duration.

$$D = R(t_m) + (1 - R(t_m))M_o$$

Fire Superiority (phi) is a measure comparing the effectiveness of Blue forces to Red forces.

$$\phi_0 = \frac{M_{blue}}{M_{red}}\sqrt{\frac{a_{blue}}{a_{red}}}$$

Where

Φ_0 = the superiority parameter

M_{blue} = the Blue Force size

M_{red} = the Red Force size

a_{blue} = the effective firing rate for the Blue Force = Reciprocal of the Target Engagement Cycle (TEC) time (i.e., $1/TEC_{blue}$)

a_{red} = the effective firing rate for the Red Force = Reciprocal of the Target Engagement Cycle (TEC) time (i.e., $1/TEC_{red}$)

(representing Not Mission Capable) and 80 to 100 (representing Fully Mission Capable) will appear. Figure 3-1-17 illustrates the modification to the piecewise linear function when no existing mission capability exists.

Weighting Elements of the Decision Model

An evaluation with multiple COIs that must be aggregated into a single overall answer, such as OE, needs a weighting methodology to balance the multiple competing objectives. For proper weighting to occur, COIs must be mutually exclusive. Because a system's individual missions do not usually depend on each other, setting up mutual exclusivity should not be difficult. For example, the outcome of a humanitarian mission is not generally influenced by the outcome of a separate and distinct attack mission. However, the evaluator should be mindful of this process to prevent inadvertent overweighting.

The most effective time to establish the weights, given their subjective nature, is when the COIs are established, preferably around MS A in the acquisition process. The weighting should reflect the needs of the Warfighter and the intent of the Combat Developer. Early establishment of COI weighting is especially important from the Materiel Developer's standpoint because the developer will want to optimize system performance for the most

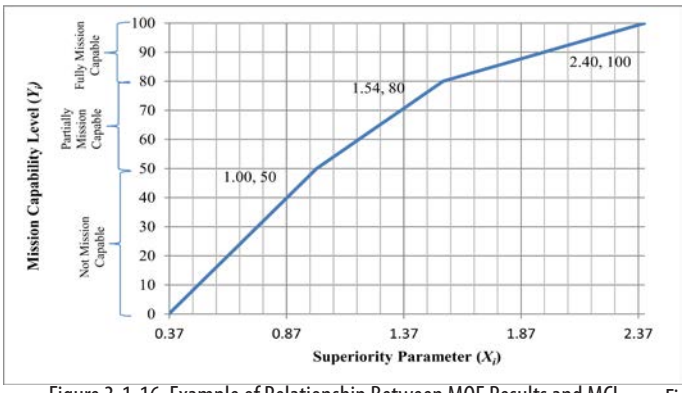


Figure 3-1-16. Example of Relationship Between MOE Results and MCL

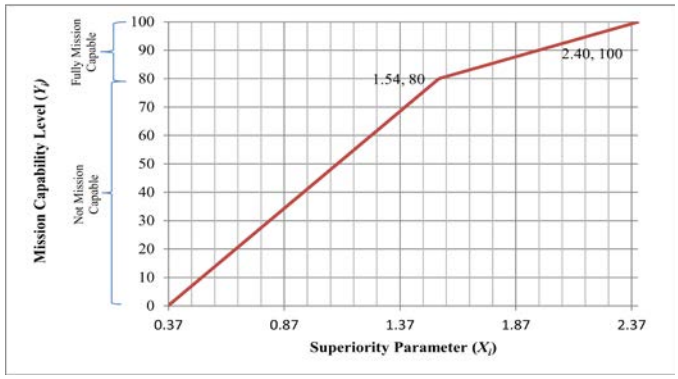


Figure 3-1-17. Example of Relationship Between MOE Results and MCL with No Current Capability for the Mission under Evaluation

important mission type when tradeoffs must be made. Establishing weights for the COIs later in the acquisition process could lead to an inappropriately optimized system or costly re-engineering for the most critical missions.

Finally, if a weight is determined late in the acquisition cycle, it may be more reflective of the developer's capabilities than the Warfighter's needs. Figure 3-1-18 is a sample decision model that incorporates weights for the COIs.

Combining Screening Criteria, Analytic Model, and Decision Model

Once the three elements of the evaluation model (the screening criteria, the analytic model, and the decision model) have been developed, the last step is to assemble the components, which establishes the roadmap for arriving at conclusions as the evaluation progresses over time.

Additional SEP Elements

Issues and Screening Criteria

MCOTEA uses a tabular format to capture issues that cover areas that may not be directly measurable in a mission profile, and might otherwise go unexamined in the course of the evaluation if not considered before operational testing. These Issues are questions and their associated Measures, developed to resolve system thresholds (identified during capability document review/parsing) and Task and Subtask performance (identified during the Operational Task Analysis). Some of these issues are further identified as Screening Criteria. Screening Criteria are used as binding constraints on the evaluation of the system. A system must meet its Screening Criteria to be OE, OS, and OSur. A template and additional guidance for this annex is available on the MCOTEA SharePoint site.

FD/SC Charter

Preparation

An OT's primary objective is to provide accurate and comprehensive information to the MDA regarding a system's operational

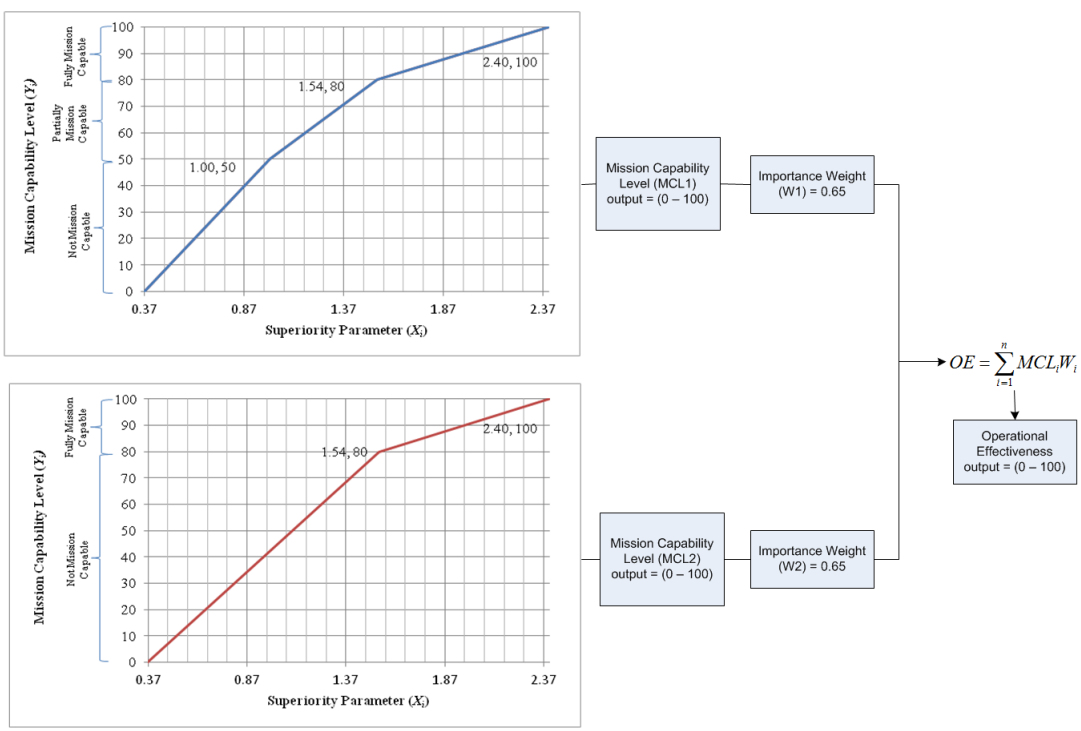


Figure 3-1-18. Example of Weighting Multiple COIs to Obtain the OE Result

performance. Some of the data collected during test is used to calculate system Reliability and Availability. This data is generally collected in Test Incident Reports, which document various types of failures during test.

Before data collection occurs, DC, CD&I, MCSC, and MCOTEA determine the basic categories of failures and the basic definitions of what constitutes those failures.

The failure categories are formalized in the FD/SC Charter, which establishes, up front, the guidelines used to classify the cause and effect of test incidents. The outcome of scoring these incidents is used to determine the Reliability estimates for that system at that point in time.

A single FD/SC Charter should be developed before testing begins and used for all contractor and government testing to score test incidents during DT and OT.

At a minimum the FD/SC Charter should contain the following information:

- FD/SC conference membership and responsibilities
- Rules of conduct for the FD/SC conferences
- System description, including components that are government-furnished and contractor-furnished equipment
- Mission Essential Functions (mef)
- Classification, chargeability, hazard severity, and hazard probability guidelines

Mission Essential Functions

Mission Essential Functions are the focal point of every charter and the element unique to every system. Mefs should flow directly from the OTA. They may also be derived from capabilities documents or developed during the charter process with concurrence from DC, CD&I. The test team reviews the system's operational mission profiles and mission scenarios and develops a short list of functions or tasks the system must be able to perform to accomplish its mission. For example, a resupply truck must be able to move to accomplish its mission of resupply, and a radio must be able to transmit digital or voice signals to accomplish its communication mission.

Mission Essential Functions are the basis of the FD/SC Scoring Conference. The failure of a system to perform any of its mefs during test results in an Operational Mission Failure (OMF), which adversely affects the system's Reliability and Availability ratings. Therefore, the development of mefs is critical to every test.

Tailoring the FD/SC Charter

Using the mefs identified in their analysis, the test team tailors the FD/SC Charter to reflect the system's unique elements. Issues such as identifying government-furnished equipment and contractor-furnished equipment are critical to the TIR scoring process.

RAM Evaluation Plan

Reliability, Availability, and Maintainability, collectively known as RAM, is a critical component of OS inextricably linked to OE. The RAM Evaluation Plan provides amplyifying information to assist in the evalution of RAM data. The plan clearly describes the system's time dendritic, definitions, and other pertinent information needed to explain measurements, calculations, or evaluation techniques. The RAM Evaluation Plan is developed concurrently with the SEP. A template is available on the MCOTEA SharePoint site (see the Volume II RAM Chapter for more details on RAM evaluation).

This page intentionally left blank for two-sided printing.

STEP 2 Test Concept and TEMP Input Development

	Task	Responsible	Provides Input	Task Output	Signature/ Release Authority
Develop Test Concept (Note: DT Observation will likely begin in this step. See Step 3 for details.)					
	Write	MS	Military SME, OTPO, TM, ORSA, DM, OTAD, DH	Draft Test Concept Brief	
	Review/Comment	OTPO	TM, MS,OTAD, DH, Exec	CRM	
	Adjudicate CRM			Test Concept ready for inclu-sion in TEMP	CRB if program is DOT&E oversight
Participate in Test and Evaluation Integrated Product Team		OTPO		Meeting Minutes	
	Write OT inputs to TEMP	OTPO		Draft OT Inputs	
	Review/Comment		TM, ORSA, MS, CRB	CRM	
	Adjudicate CRM			CRB signature-ready TEMP Input	
	Prepare post-CRB Copy			Approval-ready TEMP Input	
	Approve			TEMP Input	Dir
Develop Integrated Data Requirements List		TM	OTPO, ORSA, MS, DM,OTAD, DH	Draft IDRL, CRM, IDRL ready for inclusion in Test Plan	

Chapter 3-2. Step 2: Test Concept, TEMP Input, and Charter Development

MCOTEA develops a Test Concept for each test event that examines an evaluation question. Test Concept development directly supports TEMP section 3-1, Test and Evaluation Strategy, by ensuring that the Test Concept addresses the COIs (IOT) and Task/Subtask-level Issues (DT and Early Operational Assessment (EOA)/ Operational Assessment(OA)).

Test concept development is a MCOTEA working-level, slide-based effort (not a formal test process product). When required, such as for programs on DOT&E oversight, this format can easily transition to a brief.

Building on the SEP, the MCOTEA test team develops Test Concepts by employing Design of Experiments (DOE) to ensure that a rigorous methodology supports the development and analysis of test results.

Using the SEP as the basis and moving ahead with new details, the OTPO/TM/ OA/MS address the following topics:

1. System Definition

2. COIs and Measures

3. Trial Process Flow

4. Factors (table format; shows Constant, Nuisance, Testable, and Limitation Factors)

5. Design Type (including Reliability estimates, design, sample size, power analysis)

6. Analytic Method (one-sentence statement of method type)

7. Time Estimates

8. Key Resources

- Test Range
- Test Articles
- Funding for test
- Operating forces personnel
- Threats
- Targets
- Modeling and Simulation
- Instrumentation
- Specific requirements for hardware (tanks, trucks, C4 equipment, etc.)

9. Test Limitations

Technical Information about Test Concept Development

This section contains technical information that the test team will need to design test concepts.

Identify the Test Objective

To identify the test objective, the test team brings forward information from the SEP to ensure that each evaluation question and each Attribute with a threshold in the capabilities documentation is assigned a test event. Included with each evaluation question from the SEP are the Measures, standards, and conditions associated with the question, which serve as the impetus for the data collection methodology and for the identification of testable factors.

Define the System

The test team must provide a system definition for the part of the system that applies to each test event. The system definition defines the boundaries (what constitutes the "system" for the test, to include operators) for the system under test as clearly as possible. In early test concept development this may prove challenging, especially if the materiel solution has not yet been chosen.

Descriptions may vary since some test events exercise only subcomponents of the system and are operated by non-user representative operators, whereas other test events may exercise the complete system with Marines.

Define the Trial

A trial is one observation of the test. A collection of trials is the sample that will be used to answer the relevant evaluation questions. A trial is based on process flow (step-by-step execution of tasks and subtasks that are to occur during a trial) and the conditions under which the trial takes place. Establishing the process flow for an individual test starts with the applicable portion of the OTA developed for the SEP. The process flow is used to elicit cause-effect relationships and allows the test team to estimate resources. Operational experience is invaluable in determining process flow; familiarity with unit TTPs can be very useful in defining a trial. Figure 3-2-1 depicts a process flow for a Light Armored Reconnaissance platoon mission.

By understanding what needs to transpire during the trials, resource needs such as equipment, simulators, targets, and models and simulations can be identified. In the simple example of figure 3-2-1, it is possible to identify some major end items that will be needed for the test.

Identify Cause-Effect Relationships

When trials (as defined in step 1) are intended to be identical, they must be conducted in precisely the same way every time. Precise replication gives the tester a reasonable chance of isolating cause-effect relationships when differences in effects during the trial are noted. The OTPO/TM must ensure that the test team maintains the discipline needed to replicate identical trials throughout the test.

MSs and OTPO/TM/ORSAs, along with SMEs, identify cause-effect relationships. The MS guides the effort, which begins as a brainstorming exercise, and documents the results. To help guide the process the MS can build Ishikawa, or fishbone, diagrams as seen in figure 3-2-2.

Another useful technique involves replacing the six Ms of the fishbone diagram with diagonal lines representing each process step from the process flow diagram. The brainstorming effort may generate the same set of factors, but using

$$E[TEC] = t_{acq} + t_1 - t_h + \frac{t_h + t_f}{P(K|H)} + \frac{t_m + t_f}{P(H_1|M_1)}\left[\frac{1 - P(H_2|H_1)}{P(K|H)} + P(H_2|H_1) - P(H_1)\right]$$

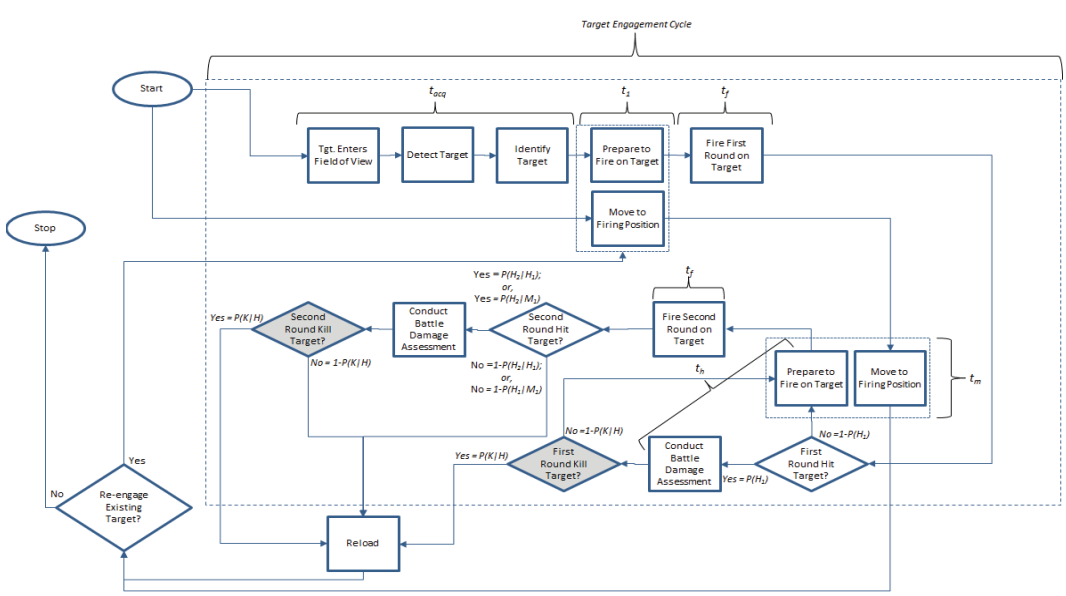


Figure 3-2-1. Sample Trial Flow Diagram

process steps has the potential to drive the brainstorming effort in a direction that might otherwise be overlooked when developing factors.

Factors and Cause and Effect Relationships

To begin identifying cause-effect relationships, the MS and OTPO/TM determine factors, also called independent variables. Factors are things the tester believes might affect the outcome (dependent variable) of the trial. See sidebar on the page 3-2-6 for factor definitions.

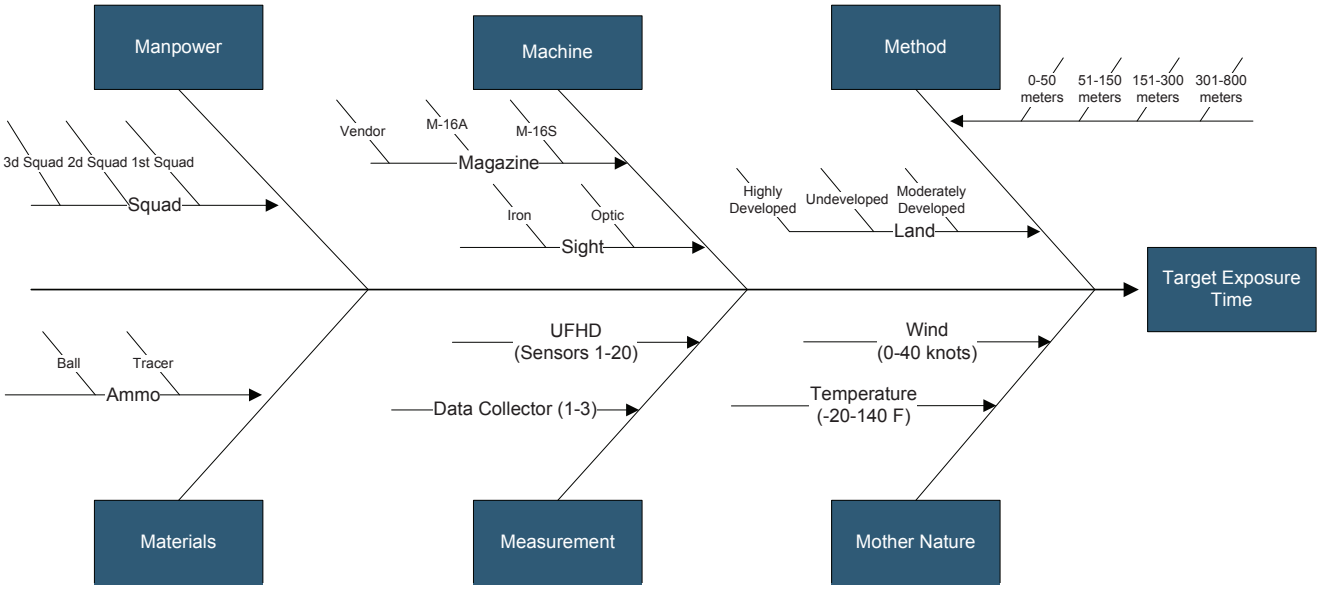
A simple example is a test designed to measure the time to get to work (outcome). Incidental to determining this time, the tester also wants to know what effect route selection (factor) has on the time to get to work. Factors can be broken down into levels. The route selection can be broken down into route-1 and route-2 (levels). In this case, the test's objective is now twofold: 1) determine the time to get to work, and 2) determine if route selection has a cause-effect relationship on the time.

Assume in this hypothetical example that although the speed of travel was supposed to be held constant during the trials, it was not. In fact, the speed traveled on one route was noticeably faster than the other route. Because the speed of travel was not adequately controlled, the tester cannot determine if route selection actually has a cause-effect relationship.

Link Inputs, Process Flows, and Outputs to Cause-Effect Relationships

After all factors believed to affect output

Figure 3-2-2. Example of a Fishbone Diagram



The six Ms, depicted by the diagonal "bones" of the figure, contain factors on each horizontal line. Not all factors can be represented as discrete values. For example, temperature is a factor that assumes a range of values. Ultimately temperature may be set to discrete blocks, but leaving the factor as continuous during the brainstorming effort is sufficient.

Six-M Definitions

Manpower - The causes attributed to the people working on the process; training would be placed here, for example.

Machine - The causes due to the machines or the equipment used in the process.

Method - The way the operation is conducted to cause the effect; target distance or type of weapon mount used, for example.

Materials - The potential causes due to the materials used, such as the difference between two ammo types.

Measurement - The causes related to how the process is measured, such as stopwatch or a ruler.

Mother Nature - The causes related to surroundings, such as external temperature or humidity.

have been identified, the next step is to select the method of control, which helps to categorize factors as nuisance, constant, testable, and limitation.

Figure 3-2-3 illustrates the linkage of the factors to a trial process and output measures. This diagram and the fishbone diagram are useful in leading a discussion of factors when brainstorming, but are not conducive to formal documentation; a Test Factors table is more appropriate in size and shape.

Select Sample Size, Design Type, and Analysis Method

Generally speaking, sample sizes and the number of trials are determined based on the balancing of resource constraints, confidence level, and the ability to detect a desired effect. Sample sizes are also determined by the design type selected. A wide variety of design types and analysis strategies is available. Basic designs include full factorial, partial factorial, and central composite. Analysis techniques include Analysis of Variance (ANOVA) and regression. Based on the design selected, resource estimation can begin in earnest to complete the test concept process in support of informing and building a TEMP.

The level of detail needed to explain the selection of sample size and the advantages and disadvantages of design types and analysis methods is beyond the scope of this publication. Analysts should research these topics independently, using a text on DOEs.

Test Limitations

The test team will attempt to evaluate all missions and capabilities of the system; however, in some cases areas will exist where the appropriate level of testing is not possible. The SEP provides a strategy for completely evaluating a system in an unconstrained test environment, but the SEP must also report any known, up front limitations affecting the evaluation strategy. A Test Limitation is a shortfall in OT depth or breadth that may affect the resolution of a test Issue. For example, some conditions simply cannot be tested in peacetime, e.g., open-air nuclear detonations cannot be used

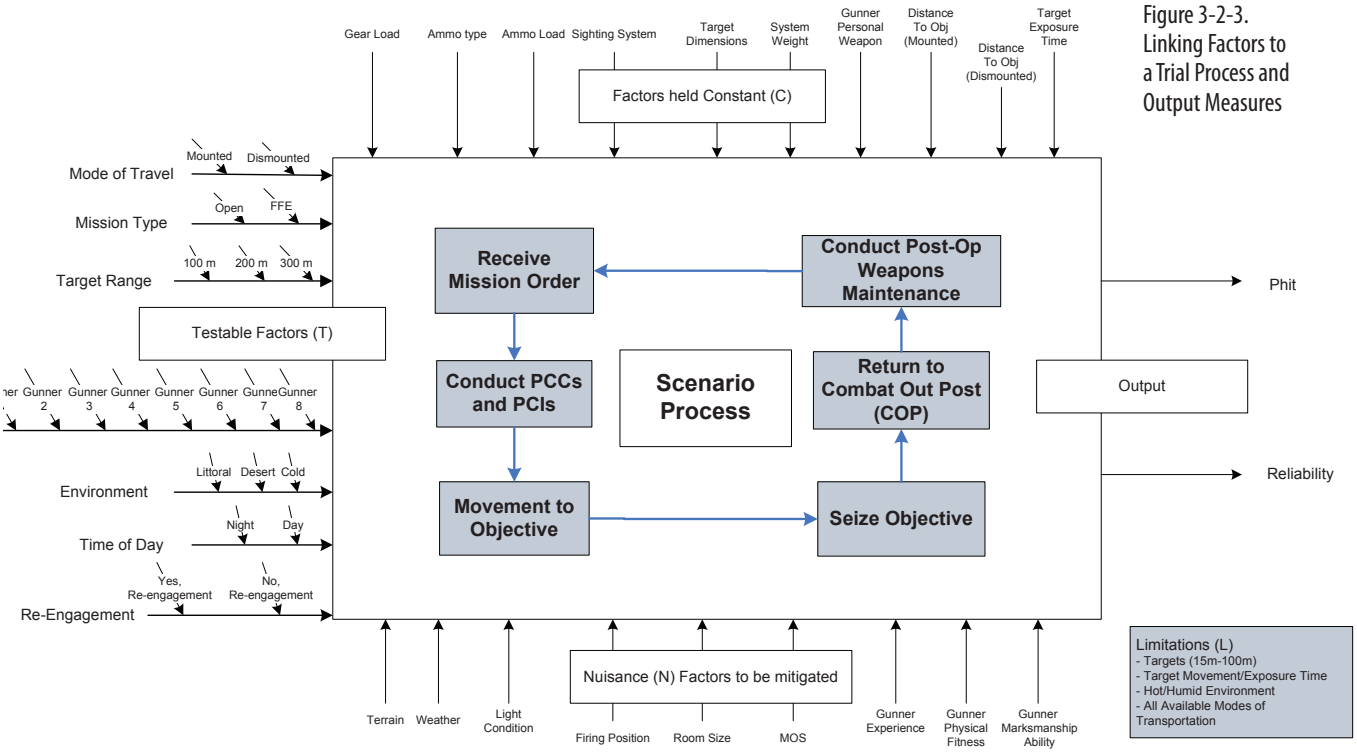


Figure 3-2-3. Linking Factors to a Trial Process and Output Measures

Nuisance Factors

Nuisance factors are uncontrolled sources of variation that represent noise in the testing process. The influence of nuisance factors can be mitigated by averaging out their effects, discussed further in the section on detailed test planning.

Constant Factors

Constant factors are controllable but not of interest to the tester because they do not contribute to answering the evaluation questions. These factors are, as the name implies, held constant for the duration of the test trials. Doing so helps ensure that knowledge gained from the test event is usable.

Testable Factors

Testable factors are used to help answer the evaluation questions. The threshold conditions from the SEP are usually the factors identified as testable conditions. Factors selected as testable are systematically varied as inputs to the test to determine their relative cause-effect relationship to the output.

Limitation Factors

Factors identified as limitations are believed to affect the outcome, but for resource, logistical, or other reasons cannot be dealt with effectively in the test design process. Limitation factors are not of interest to the tester, cannot be held constant, and cannot be mitigated through the test design process. Limitations are documented in the test concept process (TEMP, para 3.6.3 and in MCOTEA's Test Plan) to alert decision makers of the risk involved in accepting the limitations.

to operationally test Electromagnetic Pulse hardening. Limitations are also created by cost, schedule, or facilities; such limitations may not be acceptable to the Director, MCOTEA or the MDA, and, therefore, must be clearly and prominently described in the TEMP.

A Test Limitation highlights an area where the performance of the system under test may not be completely known at the completion of the evaluation. If the Test Limitation implies "inadequate" OT, the Director will request that the MDA either accept the increased decision risk associated with the limitation or increase OT resources to eliminate the limitation.

The test team must identify all potential test limitations including threat realism, resource availability, limited operational (military, physical, and civil) environments, limited support environment, maturity of tested system, inadequate M&S support, safety, etc., that may affect the resolution of operational Issues.

SECNAVINST 5000.2 states "When significant test limitations are identified, [Director, MCOTEA shall] advise the MDA of risk associated in the procurement decision." Accepted Test Limitations must be identified in the TEMP and addressed in the appropriate test event plan and report. After the test has been executed, any unanticipated test limitations that were encountered must be identified when test results are reported.

Integrated Data Requirements List

Throughout step 2, the test team builds an Integrated Data Requirements List (IDRL) as part of a dendritic analysis process to decompose Measures to data elements and data collection requirements. The IDRL is created in two spreadsheets (Data Reduction and Data Requirements) and is delivered electronically as part of the Test Concept package. After the Test Concept is approved, the IDRL is finalized as an electronic annex in Test Plans.

The IDRL identifies and displays all required data elements to be collected during test execution. It ties the data elements to the Issues, Measures, any pretest acceptance criteria, data collection instrumentation requirements, sample sizes, data structures, formats, and sources, data collection media,collection method (e.g., automated, manual form, etc.), and other information that supports data collection requirements.

MCOTEA uses Annex F of the *Analyst's Handbook for Testing in a Joint Environment* to develop a comprehensive IDRL. This handbook is available electronically (Joint Test and Evaluation Methodology 2009).

Preparing TEMP Input

The SEP provides the groundwork for MCOTEA's participation in developing the TEMP. The OTPO and TM are responsible for developing MCOTEA's contributions to the TEMP. MCOTEA uses the Defense Acquisition Guide TEMP template as a guide when providing input or reviewing the TEMP

TEMP Background and Structure

The TEMP is the contract between the developer, user, and operational tester that documents the plans, schedule, and required resources of the T&E program. The MCSC/PEO Land Systems Program Manager is responsible for producing the TEMP with the support of the T&E Working-level Integrated Product Team (T&E WIPT) (DAU 2012 and USMC 2010). MCOTEA is a member of the T&E WIPT and the Director, MCOTEA is an approving official for the TEMP.

The TEMP is a dynamic document published in support of MS B and must be reviewed and updated as required after major program changes and at each program milestone. The TEMP must be consistent with the acquisition strategy, the approved ICD, CDD, or CPD as well

as the System Threat Assessment, the Information Support Plan, the MCOTEA SEP, and other relevant documents. Figures 3-2-4 and 3-2-5 provide general information.

- Part I of the TEMP discusses Purpose, Mission Description, and System Description and is the PM's responsibility, although as a member of the T&E WIPT, MCOTEA may provide general input.
- Part II, also the PM's responsibility, requires the T&E WIPT to ensure that the OT&E schedule will support the system evaluation. Schedule is critical in supporting the engineering and decision-making processes. The PM is primarily responsible for overlaying test and reporting timelines on the acquisition timeline. MCOTEA's responsibility is to participate in the development of the POA&M in the TEMP and to inform the PM when the requirements of testing (and reporting) cannot realistically meet the expectations of the acquisition schedule.
- Part III, to which MCOTEA makes a significant contribution, contains the T&E Strategy; Evaluation Framework; Operational Test Objectives; DOE Factors, Levels, and Response Variables; and Operational Test Limitations (risk), among other topics. This section includes information about all planned OT, including any EOAs, OAs, and the IOT. Any use of Models or Simulations is also discussed, including the specific M&S; the Verification, Validation, and Accreditation (VV&A) plans; and how the M&S will be used to supplement data taken during testing.
- Part IV is a Resource Summary, which requires initial test planning information from MCOTEA.

Preparing TEMP Part III Inputs

MCOTEA uses information from the SEP to "feed" the TEMP, In the TEMP, the "Evaluation Framework" is a top-level view of the overall evaluation approach. MCOTEA uses the term to mean a complete hierarchy of COIs, Tasks, Subtasks, and their associated Issues. Although the SEP is critical to evaluation at MCOTEA, it is a separate document whose purpose is different from the TEMP; therefore, portions of the SEP cannot be dropped directly into the TEMP. Further analysis and development now occur to build the level of detail that the TEMP requires, beginning with allocation of COIs, Issues, and Attributes with thresholds to test events.

TEMP Part III Contents

3.1 T&E Strategy
3.2 Evaluation Framework
3.3 Developmental Evaluation Approach
3.4 Live Fire Evaluation Approach
3.5 Certification for IOT&E
3.6 Operational Evaluation Approach—Includes Limitations
3.7 Other Certifications
3.8 Reliability Growth
3.9 Future Test and Evaluation

Figure 3-2-4. The TEMP Outline

TEMP Outline

A detailed discussion and outline of the TEMP can be found in the Defense Acquisition Guidebook.

Multi-Service/Joint Programs

A single TEMP is used for multi-Service and Joint programs. Component-unique testing requirements can be documented in a separate annex to the TEMP.

Family of Systems

A "Capstone TEMP" may be developed for a collection of individual but interrelated systems.

Updates

The TEMP must be updated when a program baseline is breached, significant changes occur in the program, and as a part of each acquisition program milestone review.

No TEMP

A TEMP is no longer required once the system enters full-rate production and has no unresolved Major Deficiencies, or when no further OT&E or LFT&E is required. T&E requirements for any upgrades to a system are cause for new TEMP development.

Oversight

The TEMP for an ACAT I program, or any other program designated on the DOT&E oversight list, must receive OSD approval.

Figure 3-2-5. General TEMP Information

Allocating COIs, Issues, and Attributes with Thresholds to Test Events

The Evaluation Framework from the MCOTEA SEP identifies the evaluation questions that must be answered during the test program. In Part III of the TEMP, these evaluation questions are allocated to a testing source to ensure that an event will be available to generate the data needed to answer the question.

A test event may satisfy the information needs of more than one evaluation question; conversely, a single evaluation question may require more than one test event to be answered satisfactorily. The type of information needed dictates the type of test event required. Test events are of two basic types: developmental or operational.

Part III of the TEMP allocates evaluation questions and the Attributes with thresholds in the capability documentation to test events. In allocating evaluation questions to test events, the test team can begin with the expectation that test events from DT through IOT will typically address these Issues and thresholds as seen in figure 3-2-6.

Figure 3-2-6 . Focus of Test Events

DT (MCSC)	EOA, OA (MCOTEA)	IOT (MCOTEA)
Issues for OS, OSur	Issues at the Subtask level	OE, OS, OSur determination
Issues at Subtask level and below	Issues at the Task level	Missions at the COI level
Attributes with Thresholds	Issues for OS, OSur	Issues at the Task level
		Remaining Attributes w/ thresholds & Issues at Subtask and Task level depending on success of previous tests

Although the goal is to obtain relevant DT data on all KPPs and thresholds before IOT, this may not be possible. A few thresholds may need to be examined during IOT for the first time, or re-examined if significant changes to the system have been made after DT data was collected on them. In any case, as a result of the integrated test and evaluation strategy, the goal is for DT data to address all of the thresholds in the approved capabilities documentation. This will allow the operational test team to concentrate on the Mission-Based Testing approach to address satisfaction of COIs, assess system impact to combat operations, provide additional information on the system's operational capabilities, and determine the OE/OS/OSur of the system.

Summary of TEMP Part III Development

Using the SEP as a basis, the test team allocates COIs and Issues to test events and develops test concepts. The results of this work are applied to various paragraphs of TEMP Part III in conjunction with MCOTEA's participation in the T&E WIPT.

Preparing TEMP Part IV Inputs

The PM drafts Part IV to include all resources required for all types of T&E. Thus, MCOTEA's test team must ensure that all projected resources for OT&E are included in this section. The T&E WIPT should ensure that the number of required LRIP items is explicitly stated in Part IV (fig. 3-2-7, next page).

Identifying Resource Requirements

An experienced operational expert and a test management professional are essential in identifying required resources for a successful test. The preceding discussions about test concept focused on what is to be done; the test team now focuses on how, where, and what is needed to accomplish the testing.

Identify Test Locations

The T&E WIPT's job is to identify ranges, laboratories, and/or facilities needed for the test, based on the following top-level factors:

- What is to be accomplished (test concept)
- Conditions necessary for testing (factors)
- Timeframe required to support the appropriate decisions (test schedule)
- Test range capability for gathering required data
- Cost of available sites (which are most cost-effective?)

Other considerations include proximity to the personnel or using unit, support facilities, billeting, and infrastructure requirements such as maintenance bays, wash racks, secure stowage, and transportation available at testing locations.

Estimate Time Requirements

Estimating the time requirements for test events is based on sample size, time to complete a single trial, time for trial reset, training and pretest setup time, on-site daily transportation and setup times, and posttest teardown time.

Identify M & S, Cyber, and LFT&E Needs and Required Resources

The T&E W IPT develops a list of models and simulations that can be acquired or developed, VV&A'd, and used to support system evaluation. The T&E WIPT also addresses Cyber and LFT&E needs, *all of which must be adequately resourced*. See Vol. II for a full discussion of M&S and LFT&E.

Identify Key Resources

The number of test articles depends on what the test is to accomplish and the timeframe for doing so. Since test articles are a resource that must be procured, the test team must pay careful attention to the quantity and configuration of the test articles. The threshold conditions for testing will dictate the need for targets, threats, communications architectures, support equipment, and instrumentation.

Estimate Costs

MCOTEA begins the detailed cost estimation process for the TEMP by conducting a comprehensive analysis of the specific resources needed to support the evaluation strategy identified in the SEP. This analysis covers all aspects of the OT&E program. The process begins with the identification of key test resources.

Test team skill sets and levels of effort are identified and the sourcing of those skill sets are coordinated within MCOTEA. Detailed coordination between Test Divisions (including Cyber and Live Fire) and OTAD is conducted in order to ensure a common understanding of requirements and the intended sourcing of team members. For skill set requirements that will not be met using MCOTEA personnel due to limited availability, the costs of Other Government Agency (OGA) or contractor support will be estimated and included in the overall costs identified in Part IV of the TEMP.

Once the resource list is complete, MCOTEA captures the cost of testing by putting cost estimates to the resources identified.

Guidance for developing Part IV of the TEMP is found in the Defense Acquisition Guidebook (DAU 2012). Potential data sources for cost estimate development are the Program Office, MCOTEA BusMgt/S-3/S-4, MCOTEA Records Management document library, and the MCOTEA Lessons Learned from other OT events.

In addition, tools are available within MCOTEA to assist in cost estimation. Most of these tools are spreadsheet-based and come with a representative collection of resources that are found in the majority of the OT situations. All can be easily modified by the test team to reflect a tailored list of resource categories appropriate for a specific test. MCOTEA develops the OT funding summary by fiscal year, aligned to major events or phases.

Inputs and the Operational Test Plan

Test Concept and TEMP input development supports detailed test planning for Operational Assessments and Tests. SEP, Test Concept, and TEMP input are pulled forward into MCOTEA's Test Plan for further development before test execution.

TEMP Part IV

Topics in Part IV of particular interest to operational test planning include the following:

4.1.1 Test Articles (number and timing)

4.1.2 Test Sites and Instrumentation

4.1.3 Test Support Equipment

4.1.4 Threat Representation

4.1.5 Test Targets and Expendables

4.1.6 Operational Force Test Support

4.1.7 Modeling & Simulation and Testbeds

4.1.9 Special Requirements (special databases, geodesy, physical requirements, etc.)

4.3 Manpower/Personnel Training (this may be key to successful OT)

4.4 T&E Funding Requirements (Funding Schedule by FY, including pre/post OT support requirements)

Figure 3-2-7. TEMP Part IV

STEP 3 Test Planning

Task		Responsible	Provides Input	Task Output	Signature/ Release Authority
Prepare for Intermediate Assessment, if performing					
	Request Developmental Test Plans (DTP)	OTPO	TM	DTP Request Letter	DH
	Review/Comment on DTPs	TM	OTPO, MS	Correspondence (General Admin & Management)	
	Write DT Observation Plans		OTPO	Draft DT Observation Plan	
	Approve DT Observation Plan			Signed Observation Plan	
Participate in T&E Integrated Product Team		OTPO		Meeting Minutes	
Construct Feasibility of Support (FoS) Message		OTPO	Military SME, TM	FoS Message	
Prepare for Operational Assessment, if performing*					
	Write Early Operational Assessment Test Plan	TM	Military SME, OTPO, MS, DM	Early Operational Assessment Test Plan	Exec
	Write Operational Assessment Test Plan			Operational Assessment Test Plan	Exec
Write System Assessment Test Plan if performing QRA or for non-typical program*		TM	Military SME, OTPO, MS, DM	System Assessment Test Plan	Exec
Write Event Design Plan, if required (LF Div)*		TM	Military SME, OTPO, MS, DM	Event Design Plan (Live Fire Division)	Exec
Develop Operational Test Plan					
	Write	TM	Military SME, OTPO, MS, DM	Draft Test Plan	
	Review/Comment	OTPO	Military SME, TM, ORSA, MS, OTAD, DH, Exec	CRM	
	Adjudicate CRM & Prepare for CRB			CRB signature-ready Test Plan	
	Prepare post-CRB Copy			Signature-ready Test Plan	
	Approve			Signed Test Plan	Exec
Write Multi-Service Operational Test Plan*		TM	Military SME, OTPO, MS, DM	Multi-Service Operational Test Plan	Exec
Write FOT Plan*		TM	Military SME, OTPO, MS, DM	FOT Plan	Exec
Conduct Data Collection V & V Event		DM	OTPO, TM, ORSA, MS, DH, OTAD	CRM	
Coordinate Test Range, Logistics, Instrumentation, Transportation, & Personnel Resources		TM	OTPO	Trip Report	
Develop OTRB Briefing		OTPO	Military SME, TM	OTRB Briefing	DH
				Meeting Minutes	
Prepare for OTRR		OTPO	Military SME	OTRR Briefing	DH
				Pre-OTRR Memo	

*Follow Review and Approval Process outlined for Operational Test Plan

Chapter 3-3. Step 3: Test

While the MCOTEA test team prepares TEMP Part IV input, other operational test planning can and should occur simultaneously. For example, the test team can make initial test site visits, and the TM and OA can develop test trials.

When the basic information required by TEMP Part IV is complete, the SEP and the TEMP are aligned and will only require updating as program elements, such as cost and schedule, continue to change. The test team may now turn its full attention to developing the details of test plans, both for integrated testing and operational tests.

> "Integrated Testing is the collaborative planning and collaborative execution of test phases and events to provide shared data in support of independent analysis, evaluation, and reporting by all stakeholders, particularly the developmental (both contractor and government) and operational test and evaluation communities," (Secretary of Defense 2008).

The purpose of MCOTEA's test planning, execution, and reporting activities is to prepare for and conduct individual test events in support of the overall system test and evaluation plan in the TEMP and SEP. MCOTEA's involvement with testing will vary depending on the scope and size of the overall testing program outlined in the TEMP. MCOTEA observes and assesses developmental test events and conducts operational test events, supplemented by modeling and simulation as appropriate, to gather data in support of the system evaluation.

Integrated Testing

Integrated Testing, which can occur sequentially or simultaneously with MCOTEA's Intermediate or Operational Assessments, takes place before IOT. Integrated Testing can provide MCOTEA with quality test results that save duplication of effort. MCOTEA's cadre of test professionals collaboratively plans and carefully observes DT events (when appropriate) and assesses the quality of DT results.

MCOTEA does not need to observe all developmental test events set forth in the TEMP. For example, some DT events are used purely for engineering purposes to experiment with and perfect designs, manufacturing techniques, and/or operating and training procedures. MCOTEA's participation is purely optional in such cases because there is no intent to use data from such events for system evaluation. However, MCOTEA may request or be invited to attend, on a not-to-interfere basis, purely to gain a better understanding of system operations, which will aid in detailed OT planning.

When MCOTEA does intend to use data from a particular test event for system evaluation, the OTPO/TM will ensure that a knowledgeable and independent observer witnesses the event. The Divisions within MCOTEA are responsible for observing and reporting on test event execution and for analyzing the resultant test report for accuracy.

MCOTEA follows the processes described in the USMC Integrated T&E Handbook (USMC 2010). This handbook should be consulted to answer any questions dealing with Integrated Testing.

Writing DT Observation Plans

Before MCOTEA can write a DT Observation Plan, the cognizant MCOTEA Division must obtain the DT Plan from the DT organization at least 15 days before the test event (USMC 2010). Initial review should identify those parts of the plan that

- Support the independent evaluation by providing the needed data
- Need modification to support MCOTEA's evaluation
- Are to be observed by the operational tester

The test team reviews the plan and provides comments to the PM, following the overall observation process (figure 3-3-1) regarding any inadequacies, inconsistencies, or vague instructions that rise to the level of affecting the data MCOTEA requires from the test event in accordance with the TEMP. MCOTEA may also make suggestions so the DT event will produce data that MCOTEA can use. If the test team identifies no problems with the test plan and has no suggestions from the MCOTEA perspective, no requirement exists for feedback. MCOTEA does not approve/disapprove developmental test plans. However, if the operational testers identify problems in the test plan that would invalidate test data previously planned in the TEMP for MCOTEA's use, MCOTEA is obliged to inform the PM. This mandatory notification is done by standard naval letter from the head of the Division to the Program Manager and documents that MCOTEA will be unable to accept test event findings for use in the independent system evaluation unless the problems are corrected.

If no test plan exists, MCOTEA may still consider sending an operational tester to observe the event for system familiarization purposes. In no case will MCOTEA use data from an event without a plan for system evaluation purposes. Having no plan strongly indicates that results will be highly suspect. Without a plan, findings are not reproducible and cannot be independently validated, a basic quality of good scientific practice.

When MCOTEA chooses to attend an event that lacks a plan, the Division still writes an Observation Plan and a Report of its own, but only for internal purposes. While attending the event, MCOTEA observers must not in any way indicate that test event results will be used in system evaluation.

Observing Test Events

When the test event begins, the responsibility of the Division is to observe test event execution for adherence to the DT Plan. Under no circumstances should MCOTEA personnel interfere with the conduct of the event. MCOTEA's function is to observe test conduct, any deviations from the DT Plan (no matter how minor), changes to the system or its setup, or other observations that would affect the character and validity of the test event's data.

A MCOTEA observer with subject matter or operational expertise may feel the need to comment on system performance during DT observation or in the subsequent Observation Report. However, MCOTEA's focus during DT observation is on execution of the test event, not system performance. The MCOTEA observer attends the DT event as a test professional, not a system SME. MCOTEA's purpose in focusing on test event execution is to build a valid data set over the life of the test program. Doing so ultimately contributes to making the final system OTA Evaluation Report completely defensible.

MCOTEA personnel may cite system performance observations as causal factors when documenting deviations from the test plan. For example, the observer may need to note that "on the 5th trial the test was stopped because the system did not appear to be functioning. The remainder of five planned trials was abandoned while the system was inspected." The intent here is to provide the rationale for test event deviation, not to inject opinion about system performance adequacy.

If MCOTEA is unable to attend a DT event, MCOTEA may use the data results under the following circumstances:

- MCOTEA has a copy of the test plan
- A government representative (can be a contractor representing the government)

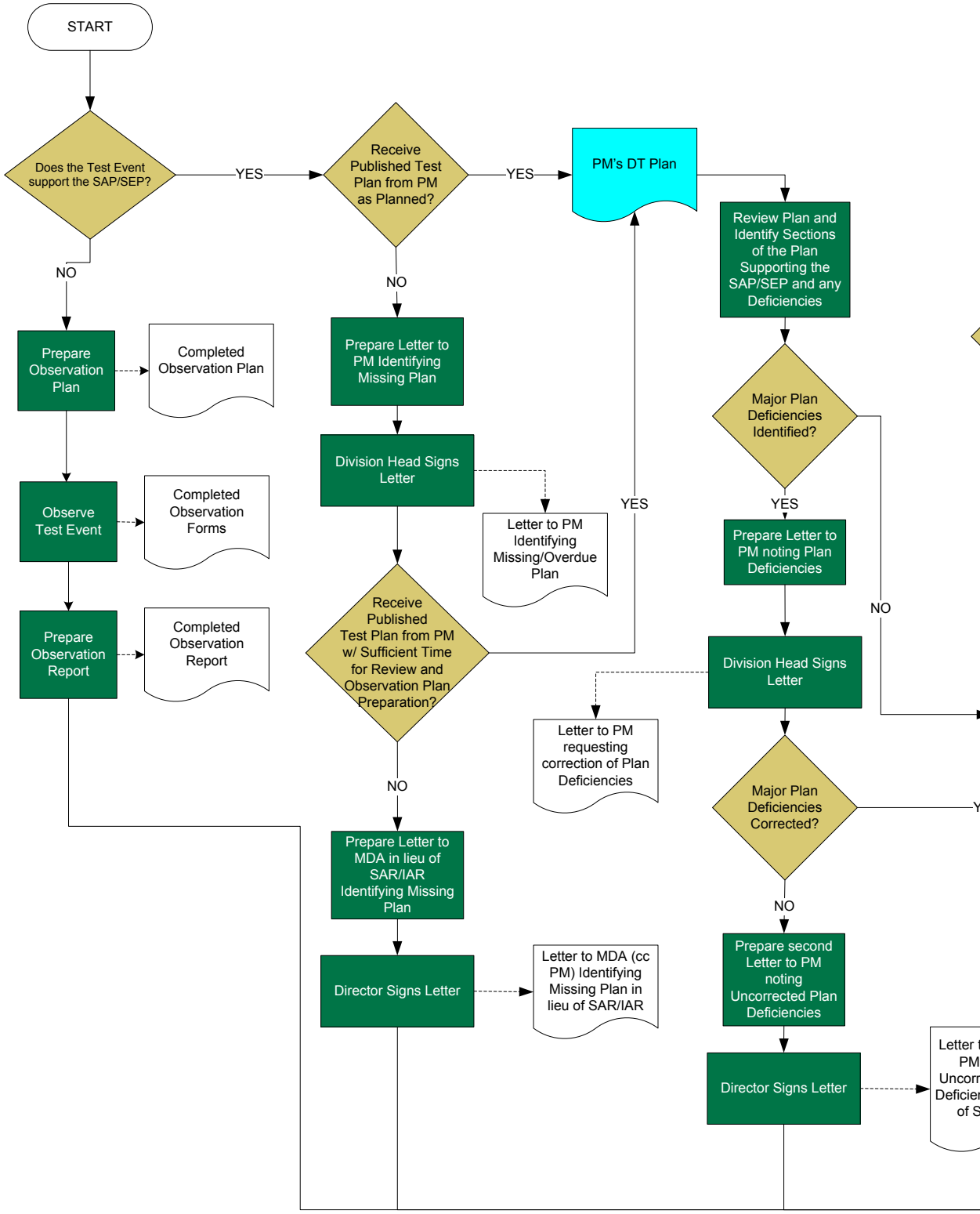
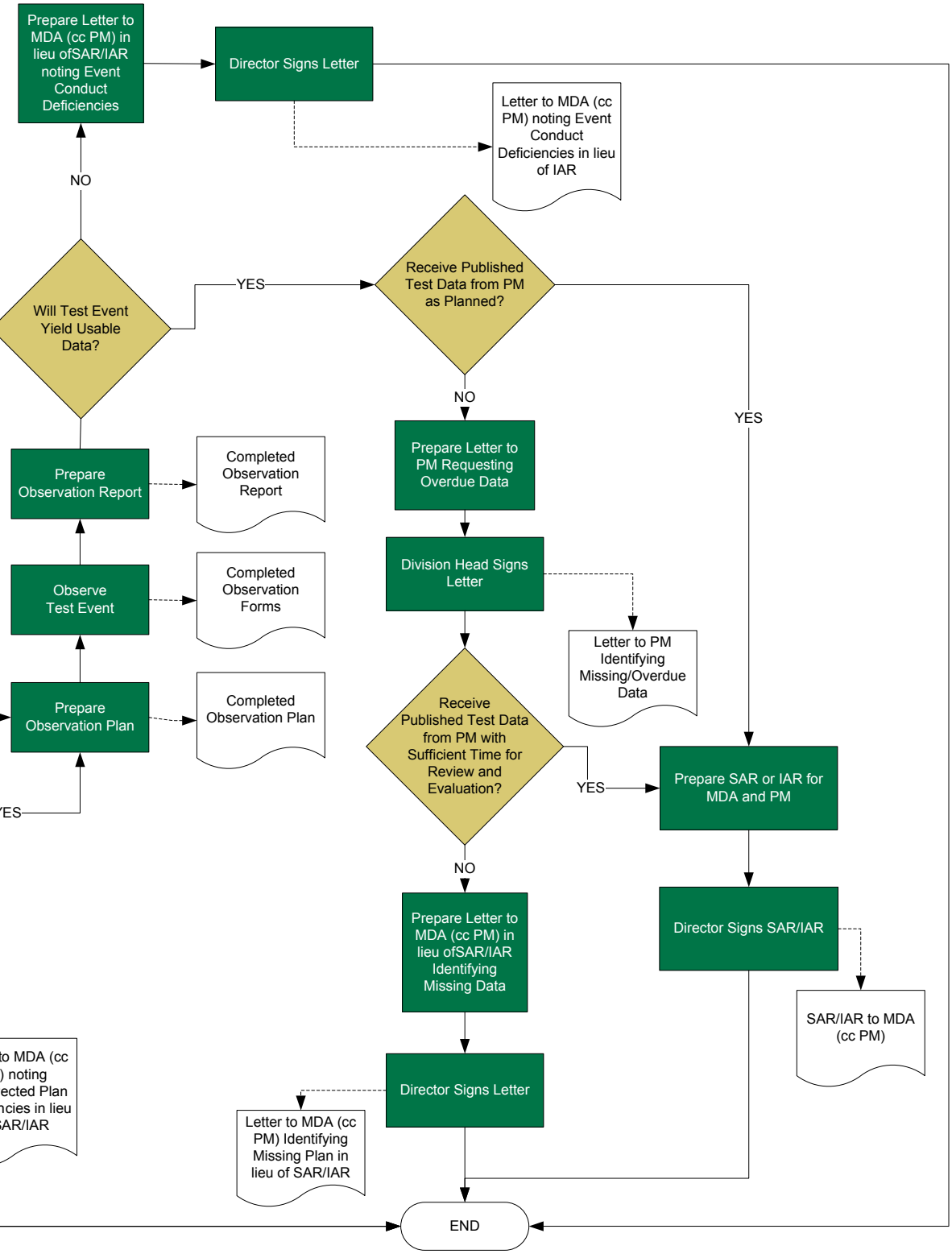


Fig. 3-3-1
MCOTEA's DT
Observation Process


Prepare Letter to MDA (cc PM) in lieu ofSAR/IAR noting Event Conduct Deficiencies
Director Signs Letter
Letter to MDA (cc PM) noting Event Conduct Deficiencies in lieu of IAR
NO
Will Test Event Yield Usable Data?
YES
Receive Published Test Data from PM as Planned?
NO
YES
Prepare Observation Report
Completed Observation Report
Prepare Letter to PM Requesting Overdue Data
Observe Test Event
Completed Observation Forms
Division Head Signs Letter
Letter to PM Identifying Missing/Overdue Data
Prepare Observation Plan
Completed Observation Plan
Receive Published Test Data from PM with Sufficient Time for Review and Evaluation?
YES
Prepare SAR or IAR for MDA and PM
YES
NO
Prepare Letter to MDA (cc PM) in lieu ofSAR/IAR Identifying Missing Data
Director Signs SAR/IAR
SAR/IAR to MDA (cc PM)
to MDA (cc
) noting
ected Plan
ncies in lieu
SAR/IAR
Director Signs Letter
Letter to MDA (cc PM) Identifying Missing Plan in lieu of SAR/IAR
END

familiar with the system being tested witnesses the test

- If the government representative is a contractor, this person cannot be employed by, or subcontracted to, the system developer
- The government representative records detailed observations of the test
- The government representative notes all deviations from the test plan
- The government representative notes all relevant caveats associated with data elements
- The government representative is available to answer MCOTEA's questions after the test
- MCOTEA has access to all recorded test data, the configuration of the test asset, and the actual test conditions under which each element of test data was obtained
- MCOTEA receives copies of all reports with data generated by the DT team

Whether the test is witnessed by MCOTEA personnel or not, MCOTEA may still use the DT data to determine the extent to which thresholds are met and may also use the DT data to help identify risks and determine OS and OSur. This data may also require regression testing, depending on the circumstances. In any case, MCOTEA will use DT data to indicate a system's progress towards overall readiness for OT.

Considerations About MCOTEA's Attendance at Early DT Events

MCOTEA may attend a Technology Demonstration or other early event (before a CDD, etc.) without intending to evaluate any results. The chief purpose in attending early events is to gather information about the system to aid in developing future integrated testing.

When MCOTEA is not evaluating results from an event, the DT Observation Report should state this and the report should be kept on file for reference.

DT Observation Reporting

MCOTEA expects to write two reports following a DT event, an Observation Report and an Intermediate Assessment Report.

The OTPO/Test Manager write the DT Observation Report after returning from the DT event. The process assumes that the OTPO/TM have not yet received the expected DT Report, meaning that only test execution, not system performance, can be discussed at this point. Observers must refrain from commenting on system performance in Observation Reports because many preliminary conclusions levied at test sites are often later found to be erroneous. Without data results in hand, conclusions about system performance remain opinion, not fact. More investigation into causality is required than can usually be provided on the test site.

The Division Head may send a copy of the Observation Report to the PM within 10 working days of event completion if there are significant recommendations or insights that need to be brought to the attention of the PM.

DT Report Review Process and IAR Preparation

After the DT event is complete, the DT team typically writes a Test Report with data, which MCOTEA should routinely receive for evaluation purposes as agreed to in the TEMP (USMC 2010). The MCOTEA OTPO/TM should follow up with inquiries if the DT data is not received within the agreed-upon time (see Chapter 4 Standard Documentation Timeline for details). If MCOTEA does not receive the DT data, or if it is not delivered in time to allow independent analysis, evaluation, and reporting before the Gate Review, then MCOTEA presents this information to the MDA in lieu of an Intermediate Assessment Report.

Early Operational Test Planning Activities

Operational testing is defined as field testing, under realistic conditions, of any item (or key component) of weapons, equipment, or munitions for the purpose of determining the effectiveness and suitability of the weapons, equipment, or munitions for use in combat by typical military users (DAU 2011) (see Guidelines, fig. 3-3-2). The principal tests that examine the Task and Subtask level are the Early Operational Assessment and Operational Assessment. The principal tests that

Operational Test Guidelines
Typical users shall operate and maintain the system or item under conditions simulating combat stress and peacetime conditions.
The independent OTAs shall use production or production-representative articles for the dedicated phase of OT that supports the full-rate production decision.
The use of modeling and simulation shall be considered during test planning. As a condition for proceeding beyond LRIP, IOT&E shall not be based exclusively on modeling, simulation, or an analysis of system requirements, engineering proposals, design specifications, or program documents. The extent of modeling and simulation usage in conjunction with OT&E shall be explained in the TEMP.
All hardware and software alterations that materially change system performance (OE, OS, and OSur) shall be adequately tested and evaluated. This includes system upgrades as well as changes made to correct deficiencies identified during T&E.
OT&E shall be conducted before full-rate production to evaluate OE, OS, and OSur as required by 10 USC 2399 for ACAT I and II programs. (SECNAVINST 5000.2 requires OT&E for all DON ACAT programs except ACAT IV(M) and AAP)
OTAs shall participate early in program development to provide operational insights to the combat developers, Program Office, and acquisition decision makers.
OT&E shall be structured to take maximum advantage of training and exercise activities to increase the realism and scope of OT and to reduce testing costs.
The use of system contractors in the OT&E conducted to support a decision to proceed beyond LRIP is restricted by 10 USC 2399. (Developing contractors may participate only to the extent that is planned for them to be involved in the operation, maintenance, and other support of the system being tested when it is deployed in combat.)
A contractor that has participated (or is participating) in the development, production, or testing of a system for a DOD component (or for another contractor of the DOD) may not be involved (in any way) in the establishment of criteria for data collection, performance assessment, or evaluation activities for the OT&E. These limitations do not apply to a support contractor that participates in such development, production, or testing, solely in testing for the Federal Government.

Fig. 3-3-2. Operational Test Guidelines

examine the mission level and answer COIs are IOT, FOT, and MOT.

Human Research Protection Program

MCOTEA is responsible to ensure that all Operational Test and Evaluation is assessed to determine any Human Research implications. The Director, MCOTEA, is designated as the Command's Institutional Official authorized to approve human subject research studies conducted under the jurisdiction of MCOTEA (Surgeon General 2012). Although the majority of Operational Test will most likely not fall under the Human Research definition, it is imperative that MCOTEA review all test plans to determine applicability and coordinate with the Institutional Review Board if necessary. OTPOs will be notified if there program is considered Human Research and will execute their programs in accordance with the MCOTEA Policy and Procedures for Research with Human Subjects (MCOTEA Apr 2012).

Creating the Feasibility of Support Message

The detailed effort associated with setting up an operational test begins when the Feasibility of Support (FOS) naval message is published, which should occur between 3 and 6 months prior to NET. Written by the MCOTEA Division, the FOS outlines the general test plan, personnel requirements, equipment requirements, facility requirements, logistical support, and any shortfalls in support needed for the test.

When creating the FOS, the test team should list all conceivable personnel requirements. (Reducing the numbers later is easier than increasing them.) The test team should clarify requirements using follow-up calls and e-mails with the

appropriate COMMARFOR and Marine Expeditionary Force (MEF) points of contact. If they do not respond within a reasonable time, the test team should consider going through the Plans, Policies, and Operations (PP&O) POC to establish contact. Often, higher headquarters must send a FOS message to the Division/ MLG/MAW G3 to determine if units are available to support test requirements. Ultimately, the MCOTEA test team must establish a POC with the supporting unit to facilitate official contact and receive a Direct Liaison Authorization (DIRLAUTH). Note: When preparing the FOS message, the test team should coordinate training requirements with the PM. Although the PM schedules and conducts operational test training, MCOTEA is ultimately responsible, meaning that the test team must ensure that assets, timeline, and facilities/training areas are included.

Test Planning Timeline

A notional operational test planning timeline of 1 year would dedicate 4–6 months to TEMP development. With the release of the FOS message 3-6 months before NET, the test team continues to develop and finalize test plan details, leading to the OTRR 30 days before NET. These time spans are rules of thumb (except for the 30-day pretest OTRR requirement) and will vary widely according to program.

Test Planning Process

Test Concept and TEMP input development supports detailed test planning for Operational Assessments and Tests. SEP, Test Concept, and TEMP input are pulled forward into MCOTEA's Test Plan for further development before test execution.

Check Lessons Learned Database

The test team should begin any test plan by reviewing the MCCLL database (www.mccll.usmc.mil) for operational tests similar to the current system. More information about reviewing and logging Lessons Learned.

Update Cost Estimates

TEMP Part IV contains the Integrated Test resources and estimates of funding requirements, although resource planning does not end with publication of the TEMP. Resource planning remains a critical activity as the test team periodically re-evaluates test schedules and mission trials, modifying resource requirements as required. Resource changes precipitate cost adjustments.

Writing the Test Plan

Detailed test planning takes the test concepts developed to support the TEMP and turns them into action plans for test execution. The test team must write the plan in enough detail to allow anyone with appropriate knowledge and skills to execute the test, more than once if necessary. The concept of repeatability is essential to good testing, and repeatability can only occur if the plan was sufficiently detailed in the first place.

Initial Sections of the Test Plan

As explained in detail in chapter 4, MCOTEA abides by a single template for test plans. Test plans contain a number of standard tables (as found in the template) but few unique graphics apart from the tables. Graphics that do appear will generally be maps or Trial Conduct diagrams unique to each program.

Refining the Schedule

At this stage in test planning, the test team adjusts the initial schedule from the Test Concept to ensure that the trials can be executed with logistical efficiency while satisfying the need to collect high quality data. In refining the test schedule, the TM and the OA may approach the same situation from two widely different viewpoints. The TM's focus is on executing trials logically and efficiently, whereas the OA may want sufficient randomization and blocking to mitigate confounding effects. Neither viewpoint is entirely right or wrong,

and the answer will most likely involve compromise.

When deciding between the two positions, the test team must always side with the opinion that will produce the highest quality data in the allotted time. The rationale is simple. An efficiently executed test event with insufficient analytic controls will most likely result in information that does not adequately explore the factors of interest to the evaluator. The test team's purpose is not simply to execute the test; the foremost purpose of the test is to gather relevant data.

To mitigate the effect of changes to the schedule, which are common and to be expected, the test team should create schedules using generic test days rather than calendar days; for example, Pilot Test Day-1 (PT-1) or Record Test Day-5 (RT-5). In identifying time as well, the test team should use generic labeling; for example, trial 4 may have a start time of Test Day Start + 8 hours, indicating that Trial 4 will begin 8 hours after the start of the test day. Specific mention of time should occur only when environmental conditions such as light levels (e.g., daylight, twilight, or nighttime) are required.

Operational Test Readiness Board/Review Process

The purpose of the OTRB/OTRR process is to ensure that the test team and system under test are ready to proceed to test. The OTRR occurs at least two times before IOT, MOT, or FOT. The pre-OTRR occurs at least 91 days before NET. Just after the pre-OTRR, MCOTEA holds the OTRB (90 days before NET). These reviews are explained below. The second and primary OTRR occurs 30 days before NET. Note that the materiel developer needs to issue the Pre-OTRR Memorandum no later than 91 days before OT. See the USMC Integrated Test and Evaluation Handbook (2010) for a detailed explanation of the OTRB/OTRR process.

Conducting the Pre-OTRR

At least 91 days before NET but always before the OTRB, the pre-OTRR is a vital opportunity for the Program Office and MCOTEA to examine the system's readiness to proceed to test. Candid discussions of system readiness are essential for two reasons. First, waiting until the last minute to cancel an operational test creates a burden on the operating forces by impeding their ability to plan and train for their normal duties. Second, proceeding to an operational test when the system is clearly not ready is a waste of valuable test resources.

The pre-OTRR is chaired by the acquisition lead or designated representative. Other attendees include the MCOTEA Division Head, the MCSC Product Group Director, the OTPO, the PM, and the DC, CD&I Action Officer. After the pre-OTRR, the Division Head reports the level of the system readiness for test to the Director, MCOTEA. In addition, the acquisition lead issues a

pre-OTRR memorandum documenting the expected state of system readiness for IOT. MCOTEA uses this memorandum as the basis for scheduling test support from the Operating Forces.

Conducting the OTRB

Approximately 90 days before NET but after pre-OTRR, the MCOTEA Division Head and PMO/PdM chair an OTRB (fig. 3-3-3) to determine the readiness of support packages, instrumentation, test planning, and test participants to support the OT. The OTRB identifies and mitigates any problems that may affect the start or proper execution of the OT. The OTRB is pre-briefed to the MCOTEA Director for concurence prior to release. Other attendees at the pre-brief include the DD, SA, and supporting staff (i.e. S-3, S-4, etc). In addition to the OT requirement, OTRBs are required for all MCOTEA-led events (i.e., QRA, EOA, OA, etc.).

Conducting the OTRR

The OTRR is conducted 30 days before NET. Its purpose is to determine if everything is ready for the operational test and entrance criteria are satisfied. The OTRR is chaired by the acquisition lead or designated representative. Participants

Example of Schedule Compromise

The ideal analytical setup for a tower-based sensor test called for two towers to be rotated between events. From the TM's viewpoint, however, the time and expense of moving the towers negated any analytic benefit. The test team decided to hold the location of the two towers constant and rotate the operators manning the towers. The OA was satisfied that doing so would average out the operators' influence between towers, thus reducing any confounding effects. However, if the initial test concept is modified, the test team must ensure that the OA's ability to answer evaluation questions has not been impeded.

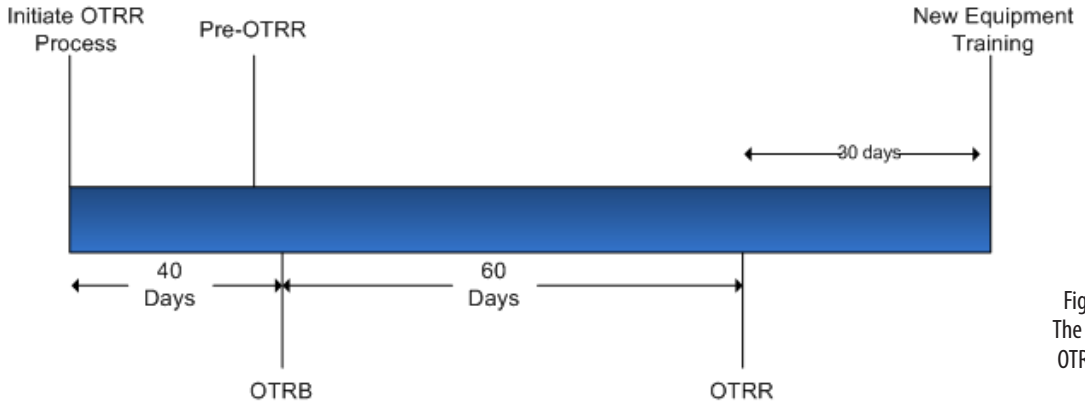


Figure 3-3-3
The OTRR and OTRB Process

include representatives from ASN(RDA) and DOT&E (for ACAT I and II programs); MCSC/PEO; DC, CD&I; and MCOTEA.

For OTRR, Commander or Executive Commander, MCSC certifies that the system is safe and ready for operational testing, unless otherwise directed by ASN(RD&A) for programs on the OSD T&E Oversight List.

The acquisition lead selects the OTRR agenda issues based on SECNAVINST 5000.2, a review of integrated testing results, and related program documentation, including certification of equipment to be safe and ready for OT&E. Agenda items may be nominated by any OTRR attendee.

Coordination of Personnel

Marine Officer in Charge and Test Unit

When it's time to schedule operating forces to support the test, the test team's top priority is to identify the MOIC and the test unit. Depending on the current operational tempo, this task can be difficult. The OTPO/TM must establish a working relationship as soon as (but not before) DIRLAUTH is received and the test unit is assigned. From this point on, the test team should include the MOIC in all site visits, scheduling meetings, test plan discussions, etc., for the MOIC to gain a better understanding of that billet's responsibilities, including working relationships and chain of command. The MOIC need not be from the supporting unit, but excellent leadership skills are important.

The MOIC is responsible for helping to execute the test plan and report test deviations to the OTPO, among other duties including the following:

- Helping to coordinate necessary resources required to support tests
- Supervising the Marines conducting the events described in Trial Conduct and ensuring that Marines collect data specified in Data Requirements
- Ensuring that the Marines collect the data in accordance with the Test Plan
- Maintaining a daily log that includes significant events and incidents that affect test conduct, test events completed, and personal observations of the test conduct and system functionality
- Tracking the daily review, editing, and compilation of all data collection forms and electronic data collection
- Reviewing TIRs for accuracy and completeness and providing preliminary scoring of TIRs for scoring conference members

Data Collectors

Data collectors can generally be recruited from one of four sources: government civilians, active duty Marines or soldiers (for Joint tests), contractors, or, in rare cases, Reservists, Sailors, and Airmen.

Other Government Agency civil service support can also be used to provide data collector support. Organizations such as the Naval Surface Warfare Center have a depth of T&E experience and with appropriate advance notice can provide resources for short to mid-duration data collection efforts.

Contractors

Generally, contractors are the easiest to schedule. They are usually experienced personnel who require very little training and can easily adapt to unexpected test-related situations. They are arranged for through a supporting contractor.

Active Duty Marines

Active duty Marines are able to fill dual roles: they can be trained as data collectors and they can function as alternates if the test requires Marines in a specific MOS or if someone must leave the test early due to an emergency.

Details about test planning logistics appear in Annex A at the end of this chapter.

Note: Early test planning may also include the need for verification, validation, and accreditation of modeling and simulation to be used for data collection. See vol. II of this manual for details.

Reservists

Reservists can sometimes be scheduled for duty during the test. If a Reserve Unit resides in the area, the test team can contact the I & I (Inspector and Instructor) to learn if any Marines need active duty time. Paying Reservists can be problematic; the test team should check with the Reserve Support Center if the I&I has no available funding. Discretionary funds are sometimes available to support active duty time for Reservists.

Detailed Test Planning

With the basics in place, the test team can now fill in the core of the Test Plan, which contains Sample Size and Test Design, Data Requirements, Data Collection Methods, Data Reduction, Data Analysis, and Trial Conduct for OTPO/TM/MS/MOIC use.

To begin this section of the plan, the test team brings forward information from the SEP, TEMP, and IDRL. Even if some of the following information is not covered in the TEMP or the SEP, the test team must address each item for each test:

Each COI/Issue is supported by the test and its respective Measures.

Figure 3-3-4 Operational Test Plan, con't.

Sample Size and Test Design				
	Illumination			
	Day		Night	
Target Distance	Armor	FFP	Armor	FFP
Near	4	4	4	4
Mid	4	4	4	4
Far	4	4	4	4
Power = 1- β at 20% alpha level to detect a signal/noise (δ) ratios of:				
δ=0.25	δ=0.33		δ=0.75	
Sample Size	Sample Size	Power	Sample Size	Power
228	228	0.885	228	1.000

Data Requirements	Data Collection Method
Trial Number	Trial Number, Target Distance, Target Type, Illumination, Blue Force Pre-Mission Force Size, and Red Force Pre-Mission Force Size will be predetermined by the trial sequence and prerecorded in the test trials database in the Offensive Trials Table.
Target Distance (Near, Mid, Far)	
Target Type (Armor, FFP)	
Illumination (Day, Night)	
Blue Force Pre-Mission Force Size	Trial Number, Intial Engagement Blue Force Size, Initial Engagement Red Force Size, Ending Engagement Blue Force Size, Ending Engagement Red Force Size, Mission Orders Issued, Mission Start, and Mission Complete will be recorded on Form X.
Blue Force Readiness Losses	
Blue Force Dependability Losses	
Initial Engagement Blue Force Size	
Red Force Pre-Mission Force Size	See MOP-1.1 for (Blue and Red) Target Engagement Cycle Time
Red Force Readiness Losses	See MOS-1 for (Blue and Red) Force Readiness Losses
Red Force Dependability Losses	
Initial Engagement Red Force Size	See MOS-2 for (Blue and Red) Force Dependability Losses
Ending Engagement Blue Force Size	
Ending Engagement Red Force Size	
Blue Target Engagement Cycle Time	
Red Target Engagement Cycle Time	
Mission Order Issued (Time)	
Mission Start (Time)	
Mission Complete (Time)	

Each COI is set up separately in the plan with its Measures.

Insert Sample Size and Test Design

Using information developed in the TEMP, the test team inserts sample size and test design information beneath the COI as seen in the template sample (figure 3-3-4).

The sample size and test design table in the template identifies the trials that allow sufficient spreading across the nuisance factors and testable factors.

Finalize Data Requirements

Using information developed in the IDRL, continue to refine the data requirements throughout the Test Planning step. The data requirements (sometimes called data elements) are the individual pieces of information needed to satisfy the Measures. In addition, there are data requirements to conduct appropriate analyses on the measured results, such as establishing cause-effect relationships. A principal job of the TM and MS is to develop data requirements to satisfy the evaluation questions.

Data Requirements for Measures

An example of a data requirement for a Measure is "time to set up" for each trial for which the elapsed setup time must be recorded. However, the OA should take care to consider the widest possible uses for the data. Test data has a temporal quality that should not be overlooked. Elapsed time for the Measure technically satisfies the Measure, but valuable data would be lost if "when" the trial took place was not collected. To capture both the elapsed time and the data's temporal quality, the following data requirement would be needed to satisfy the MOP:

- Time Start (hh:mm:ss dd/mm/yyyy)
- Time Stop (hh:mm:ss dd/mm/yyyy)

Given this data, the MS can use data reduction methods (Time Stop – Time Start = Elapsed Time) to reduce the data to its usable form and retain temporal quality by knowing when during the test period the trial took place. This is especially important in operational tests where time of day or task sequence is important in understanding what transpired during a mission.

Determine Data Requirements for Analysis

For analysis methods to be completed, additional data elements are required. Data from each trial must be collected to support each factor (constants, nuisance, and testable).

Using the MOP "probability of hit" as an example, the data requirements in figure 3-3-5 would be needed to satisfy some of the factors that need to be analyzed.

Factor Type	Data Element
Nuisance	Instrumentation Error
Nuisance	Test Day
Constant	LAV-AT Crew Size
Constant	System Location
Testable	Range to Target

Figure 3-3-5
Data Requirement Examples

Develop Data Collection Methods

Data collection methods fall into two categories, automated and manual, each with advantages and disadvantages. The best data collection methods for operational tests do not interfere with the accomplishment of tasks or missions, or do so to the slightest possible extent.

Automated Data Collection

Automated data collection involves some form of instrumentation that is set to monitor and record what is occurring on the test site. Instrumentation can be installed directly onto or into the system under test or on the test range. Automated data collection methods are useful when space requirements limit access to personnel outside of the crew or operators.

Automated data collection often can record information that would not be available using manual collection, and with speed and accuracy that manual efforts cannot duplicate.

A disadvantage of automated data collection is that it typically requires personnel with specialized skills to set up and operate. When ranges are instrumented with automated data collection methods, considerable preparation time may be required to set up the range before a trial can begin. Additionally, not all ranges are suitable for automated data collection, e.g., the need for external power sources may limit automated collection utility in a free-flowing operational event. When automated data collection methods are installed onboard or incorporated into the system under test, particular attention must be paid to ensure that the device does not interfere with system operations.

Manual Data Collection

The primary focus of a manual data collector is to observe and record. Data collection must be limited to collecting the necessary data elements, not scoring, tabulating, or calculating results, which are data reduction functions performed by the OA. Manual data collection can employ paper or electronic forms and has the advantage of being highly adaptable to changes.

However, manual data collection has many disadvantages, chief among them the possibility of distractions and documentation errors; space requirements; and training requirements. Other disadvantages include the following:

- Documenting data elements on a test requires attention to detail and the ability to ignore activities not of principal concern. A manual data collector can also misinterpret or incorrectly document an observation. Both problems result in poor quality test data.
- Manual data collection can be particularly challenging in tests where limited space is available for personnel outside of the crew or operating personnel. Manual data collection, in most cases, assumes that the data collector must be in place to observe and document the data elements, which requires sufficient space for that person. In tests of vehicle systems this challenge is particularly difficult to overcome. The initial thought might be to use a crewmember as a data collector or replace a crewmember with a data collector, but neither option is viable because both have the potential for producing a poor quality test result. Using a crewmember as a data collector means tasking that person with a job outside his particular skill set that he may not be equipped to handle. Tasking an operator with collecting data may also overburden the operator who retains crew responsibilities. Replacing an operator with a data collector has similar implications. A data collector cannot perform the duties of the replaced operator, and in any case the data collector must not be involved in system operation; the data collector must remain a passive, non-interfering observer of events.
- One of the greatest challenges for a manual data collector is collecting data on a multifunctional system. A data collector may be charged with recording failure information on one system function while simultaneously recording events of other system functions. Many people do not multitask well. The nature of data collection work is sequential tasking, where attention may move back and forth between different tasks, but not focusing on more than one at a time.
- Training data collectors requires time and effort to ensure that they understand their roles and responsibilities. While setup time may be reduced by using manual data collection methods, personnel requirements may increase, including the additional logistical and training burden. Most data collection efforts are unique to each operational test, meaning that data collectors must be trained for each test event. When employing electronic data collection devices such as PDAs, additional familiarization time may be required. Finally, despite being sufficiently trained, manual data collectors often make errors of omission, transposition, accuracy, or judgment.

Data Reduction

1. Select all records from the Offensive Trials Table with the following fields: Trial Number, Target Distance, Target Type, Illumination, Blue Force Pre-Mission Force Size, and Red Force

2. Sort all records in assending order based on Trial Number

3. For each trial in the Offensive Trials Table input the Initial Engagement Blue Force Size, Initial Engagement Red Force Size, Ending Engagement Blue Force Size, Ending Engagement Red Force Size, Mission Orders Issued (time), Mission Start (time), and Mission Complete (time) from Form X.

4. For each trial in the Offensive Trials Table input the Blue Force Target Engagement Cycle Time and Red Force Target Engagement Cycle Time from MOP-1.1.

5. For each trial input the Blue Force Readiness Losses and Red Force Readiness Losses from MOS-1.

6. For each trial input the Blue Force Dependability Losses and Red Force Dependability Losses from MOS-2.

7. For each trial compute the reciprocal of both the Blue and Red Force Target Engagement Cycle Times to arrive at the Effective Firing Rate using the formula:

$$1/TEC_i = a_i$$

Where:

TEC = Target Engagement Cycle Time

a = Effective Firing Rate

i = *{Blue, Red}*

8. Compute the Force Superiority Parameter value for each trial using the forumla:

$$\frac{M_{Blue}}{M_{Red}}\sqrt{\frac{a_{Blue}}{a_{Red}}} = \phi_0$$

Where:

MBlue = Inital Blue Force Size

MRed = Initial Red Force Size

aBlue = Blue Force Effective Firing Rate

aRed = Red Force Effective Firing Rate

Φ0 = Force Superiority Parameter

Data Analysis Method

Linear Regression. The general model form used, reference (*x*):

$y=\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_kx_k+\varepsilon$

Where:

y = response (dependent) variable, i.e., Fore superiority

β_0 = defines the intercept of the plane

β_j = regression coefficient (expected change in y per unit)

x_j = regressor (independent) variables from Test design

$j = 1,\cdots, k$

k = the number of regressor variables (i.e., Target Type, Target Distance, and Illumination)

ε = error term

Overall Model Hypothesis and Test Statistic

$H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0$

$H_1: \beta_j \neq 0$ for at least one j

Accept H_0 if $F_0 \leq F_{a,k,n-k-1}$

Reject H_0 if $F_0 > F_{a,k,n-k-1}$

Where

$$F_0 = \frac{SS_R / k}{SS_E / (n-k-1)}$$ = test statistic

SS_R = Sum of Squares for the regression

SS_E = Sum of Squares for the error

α = 0.10 significance level

n = total sample size

Individual Regression Coefficient Hypothesis and Test

$H_0: \beta_j = 0$

$H_1: \beta_j \neq 0$

Accept H_0 if $|t_0| \leq t_{a/2,n-k-1}$

Reject H_0 if $|t_0| > t_{a/2,n-k-1}$

Where

α = 0.10 significance level

$$t_0 = \frac{s_j}{\sqrt{\hat{\sigma}^2 c_{jj}}}$$ = test statistic

$\hat{\beta}_j$ = least square of the estimator of the jth regression coefficient

$\sqrt{\hat{\sigma}^2 C_{jj}}$ = standard error of the jth regression coefficient

Note: the Trial Sequence table is placed under the Data Reduction/Data Analysis Method table in an actual Test Plan, not next to it. The table runs for as many pages as necessary.

Trial Sequence

Trial Number	Target Distance	Target Type	Illumination
1	Night	Armor	Mid
2	Night	FFP	Near
3	Day	FFP	Mid
4	Day	Armor	Far
5 & PT-1	Night	FFP	Near
6	Night	Armor	Far
7 & PT-2	Day	FFP	Far
8	Day	FFP	Near
9	Night	FFP	Far
10 & PT-3	Night	Armor	Near
11	Day	FFP	Mid
12	Day	FFP	Mid
13	Night	Armor	Near
14	Night	FFP	Mid
15 & PT-4	Night	Armor	Far
16	Night	FFP	Mid
17	Day	Armor	Mid
18	Day	FFP	Far
19	Night	Armor	Near
20	Night	Armor	Mid
21	Day	Armor	Near
22	Day	Armor	Near
23	Night	FFP	Far
24	Night	Armor	Far
25	Day	Armor	Mid
26	Day	Armor	Far
27	Night	FFP	Far
28	Night	FFP	Far
29	Day	FFP	Far
30	Day	FFP	Near
31	Night	Armor	Mid
32	Night	Armor	Near
33	Day	Armor	Far
34	Day	Armor	Mid
35	Day	FFP	Far
36	Day	FFP	Near
37	Night	FFP	Near
38	Night	Armor	Far
39	Day	FFP	Near
40	Day	FFP	Mid
41	Night	Armor	Mid
42	Night	FFP	Mid
43	Day	Armor	Near
44	Day	Armor	Near
45	Night	FFP	Far
46	Night	FFP	Mid
47	Day	Armor	Near
48	Day	Armor	Mid

Fig. 3-3-6 Operational Test Plan, con't.

When manual data collection is the preferred method, the test team should consider who is best suited to the task. Using Marines as data collectors has some advantage in that they are familiar with the military operating environment, but data collection is not their purpose in the Marine Corps. In addition, using Marines as data collectors increases the burden on the Operating Forces.

Using civilian data collectors can lessen the burden on the Operating Forces. Civilian data collectors can also be obtained earlier in the test planning process to improve training and awareness of what is to be collected and the methods for doing so. However, a civilian data collector may be inexperienced in the harsh military environment and may be ill-suited for dealing with it.

Develop Data Reduction Methods

Data from the test must be reduced to a form useful to the OA, and the form will vary from test to test (fig. 3-3-6). The formal definition of data reduction is the transformation of information, usually empirically or experimentally derived, into corrected, ordered, and simplified form. The term generally refers to operations on either numerical or alphabetical information digitally represented, or to operations which yield digital information from empirical observations or instrument readings.

Data reduction methods should be documented for each Measure in a Test Plan and tailored to the data collection methods.

Adding Details to Test Trials

The Test Concept and TEMP provide the basic information required to produce test event trials, which are formed around the missions Marines will execute using the system under test and, therefore, may be multiple in number. The test team should anticipate that IOT will cover every mission associated with the system. Other types of tests, e.g., EOA or OA, may investigate only one mission or a single capability area requiring partial execution of multiple trials.

With more information available since TEMP development, the test team should re-examine the cause-effect relationship of factors; the six-Ms (Materiel, Methods, Manpower, Machine, Measurement, and Mother Nature) are more certain now. Figure 3-3-7 provides a detailed example of Trial Conduct development. The test team writes the instructions required to exercise systematic variation of the factors from trial to trial.

Test Plan Annexes

The main body of the Test Plan is followed by three annexes: Detailed Schedule, Data Collection Forms, and Safety Plan.

The Detailed Schedule displays the Measures and trials for each day's execution as well as logistical support needs.

Data Collection Forms are printed to ensure traceability with data requirements as the Test Plan is reviewed for approval.

The Safety Plan is unique to each test site and provides detailed instructions for different types of emergencies. Templates for MCOTEA's most frequently used test sites are available for convenience.

Figure 3-3-7. Detailed Example of a Trial

Example of a Complete Trial Conduct

a. **Trial Objective.** The objective of the trial is for the LAR platoon equipped with Anti-Tank variants of ABC Systems to accomplish the specified mission objective as noted in the Mission Order. The mission objective is to destroy like-size enemy forces.

b. **Equipment/Personnel**

(1) The Blue Force (friendly) equipment consists of four ABC Systems combat-loaded IAW reference (x) with a total of 12 crewmen (ref. y) and either 2 ABC-ATMs or 2 ABC-ATs combat-loaded IAW reference (x) and 6 crewmen (ref. y).

(2) The Red Force (enemy) equipment consists of four ABC-25s combat-loaded IAW reference (x) with a total of 12 crewmen (ref. y) and 2 M1-A2 Abrams Tanks combat-loaded IAW reference (x) with a total of 8 crewmen (ref. z). The four ABC-25s will simulate BMP-1s, BMP-2s, and BTR-70s, while the two M1-A2 Abrams Tanks will simulate T-72 Tanks.

(3) The White Force (mobile observers/controllers) equipment consists of two HMMWVs, each equipped with one AN PRC-150 radio, one Motorola range safety radio, 20 gallons of potable water, one Electronic Tablet (Data Collection Device), Driver, Test Manager, and a Field Data Collector.

(4) The White Force (stationary controllers) equipment will consist of six AN PRC-150 radios, two Motorola range safety radios, one force-on-force master display and computer, one force-on-force instrumentation base station, two force-on-force instrumentation technicians, two OE-254 mast antennas, three Electronic Tablets (Data Collection Devices), one OTPO, one MOIC, one DM, and two DCs.

(5) All Blue and Red Force vehicles will be equipped with force-on-force instrumentation for collecting time, space, position, shot, and hit information.

(6) All DCs will be equipped with ruggedized handheld data collection devices.

(7) 2,270 gallons JP-8 fuel (500 gal. x 2 M1A2 Tanks; 122 gal x 10 ABCs; and 25 gal x 2 HMMWVs

(8) Eight TOW 2B empty missile casings

(9) 2,000 rounds of 25mm blank fire ammunition

(10) Four Electronic Tablet Computers for SPOT, Situation, and SALUTE [Size, Activities, Location, Unit Identification, Time, Equipment] Reports

c. **Roles and Responsibilities**

(1) The role of the Blue Force is to seize, secure, defeat, clear, defend, or delay as specified in the Mission Order for each trial by executing fire and maneuver.

(2) The role of the Red Force is to resist the attempts of the Blue Force to accomplish their assigned tactical task by executing fire and maneuver.

(3) The MOIC will serve as the master trial controller (White Force) while simultaneously serving as higher headquarters for each force. The MOIC will issue Blue and Red Force Mission Orders and receive Blue and Red Force SALUTE reports and Situation Reports.

(4) Data Collectors will serve as recorders for test information gathered during a trial.

(5) Instrumentation Operators will prepare the instrumentation package on each vehicle and the White Force for operation and monitor the status of instrumentation during trial conduct. If the instrumentation malfunctions, the Instrumentation Operators are responsible for restoring functionality.

d. **Pretrial Preparation**

(1) Based on the trial being executed from the Trial Sequence table, the testable factors for the trial will be set as initial conditions (i.e., Target Distance, Crew, Target Type, Illumination).

Figure 3-3-7 is an example based on a Light Armored Vehicle system and illustrates the level of detail needed for enemy actions to be carried out as part of the test.

A complete Trial Conduct is developed by stating the trial's Objective, Equipment/Personnel, Roles/Responsibilities, Pretrial Preparation, Trial Conduct, Special Situations, and Posttrial conduct.

(2) The force-on-force instrumentation technicians will perform a pretrial inspection of the onboard instrumentation to ensure it is fully functional and replaced as needed before beginning the trial. During the instrumentation check, the data link for each vehicle (Blue and Red) will be verified on the White Force master computer and display.

(3) The Red Force vehicles will be marked with an up arrow symbol (i.e., ↑) on the front, sides, and rear of the vehicles with 2" white reflective tape to indicate their status as enemy vehicles. The markings will be approx. 2' in height by 1' in width on an exposed vertical (or nearly vertical) surface of the vehicle.

(4) The White Force (stationary controllers) will occupy the test site command center, perform radio communications checks and services, and contact range control to request permission to use the range or transit the training area.

(5) The HMMWVs will undergo pre-operations checks and services as required. Upon completion, but not less than 1 hour before mission start time, the White Force mobile observers/controllers will request permission from range control using the Motorola radios to transit the range, and upon receiving permission will depart the test site command center and move to their designated observation points. Designated observation points are identified in figure-x [INSERT DIAGRAM or MAP]. Upon arrival, the White Force mobile observers/controllers will contact the White Force stationary controllers using AN PRC-150 radios to indicate that they are ready for the trial to begin. The White Force mobile observers/controllers will also contact range control using the Motorola radios to indicate that they have reached their designated location.

Note the use of diagrams and maps as needed.

(6) The Red Force will receive its Mission Order, perform vehicle pre-operations checks and services, and move from the test site command center to its assembly area. The Red Force-designated assembly area is identified in figure-x for each trial.

(7) The OTPO will review the trial Go/No-Go criteria and determine if all Go criteria have been satisfied.

e. **Trial Conduct**

(1) *Initiate Trial.* The trial begins upon receipt of the mission order in five-paragraph format. The Blue Force commander will brief the mission and Blue Force crews will begin to prep all vehicles for the mission by loading and stowing fuel, ammunition, weapons, and personal gear and will conduct pre-Combat Checks and pre-Combat Inspections. The Blue Force will also conduct vehicle pre-operations checks and services as required by the technical manuals. If an Essential Maintenance Action is discovered, that vehicle's crew and supporting maintainers will follow the process outlined in MOS-x.x (Operational Readiness).

(2) *Mission Departure.* At the designated mission departure time specified in the mission order (table x), all Blue Force vehicles not currently undergoing Essential Maintenance Actions will depart the assembly area and begin movement along the designated route outlined in figure x. The Red Forces will also depart their respective assembly area. During the course of transit the Blue and Red Forces will search for their opposition forces and update their respective headquarters elements (i.e., White Force Stationary Controllers) with Spot Reports and situation reports (SITREP) via the vehicle onboard communications equipment every 30 minutes (see Annex B, Data Collection Forms). The Data Collectors located in the White Force (stationary controller) will record the reports in the database contained in the Electronic Tablets.

(3) *Record SALUTE.* At any point during the mission when the Blue or Red Forces detect what they believe are oppositional forces they will contact higher headquarters via vehicle onboard communications equipment using a SALUTE report format (see Annex B, Data Collection Forms) noting the enemy size, activity, location, unit, time, equipment, and any remarks. The Data Collectors in the White Force (stationary controller) will record the SALUTE report in the database contained in the Electronic Tablets.

(4) *Identify the Target.* Upon detection, the detecting force (Blue or Red) will take the necessary actions (e.g., maneuver or halt as tactically required) to identify the target (i.e., target identification) as friendly, neutral, or enemy. Once identity has been established they will

contact higher headquarters via vehicle onboard communications equipment using a SALUTE report format (see Annex B, Data Collection Forms) noting the enemy size, activity, location, unit, time, equipment, and any remarks. The Data Collectors in the White Force (stationary controller) will record the SALUTE report in the database contained in the Electronic Tablets. In the event that target detection and identification occur simultaneously, only a single SALUTE report will be reported to higher headquarters.

(5) *Target Engagement Cycle.* If the identity of the target is determined to be an enemy target the force will begin the target engagement cycle. The cycle begins by simultaneously preparing to fire on the target and maneuvering to a position as tactically required to support successful engagement. Once in firing position and ready to fire the firer will fire the first round at the target. The time and location of shot will be recorded by the force-on-force instrumentation. The force-on-force instrumentation on the targeted vehicle will record if a hit occurs and provide a visual signature (i.e., external light will illuminate) to notify the operator that the target was hit. No illumination from the light means the shot was a miss. Next, the firer will conduct a battle damage assessment to determine if the hit was a kill. A flashing means the target was hit, but not killed. If the target was not killed the firer will simultaneously prepare to fire again on the target and maneuver as needed to a position as tactically required to support successful engagement. Once in firing position and ready to fire the firer will fire the second round at the target. The time and location of shot will be recorded by the force-on-force instrumentation. The force-on-force instrumentation on the targeted vehicle will record if a hit occurred and provide a visual signature (i.e., external light will illuminate) to notify the operator that the target was hit. Next, the firer will conduct a battle damage assessment to determine if the target was hit and if the hit was a kill as previously described.

(6) *Reevaluate.* After the second shot, or after the first shot if the engagement is discontinued, the firer will reload. If targets still exist at the end of the first engagement they will determine if re-engagement is required per the mission order. If re-engagement is required the firer will repeat the steps described in paragraph 5. If re-engagement is not required and the mission objectives have not been achieved the firer will resume actions as described in paragraph 2. If the mission objectives have been achieved the mission will be ended and the mission will be scored by the White Force using the mission score card form located in Annex B, Data Collection Forms.

f. **Special Situations**

(1) Loss of a Mission Essential Function (mef). If at any point during the mission a vehicle experiences a failure the crew will determine if there is a loss to a mission essential function. If a determination is made that a mission essential function has been lost, that vehicle's crew will follow the process outlined in MOS-x.x (Mission Reliability). Regardless of the type of failure, the crew will issue an out of sequence SITREP to higher headquarters. The Data Collectors in the White Force (stationary controller) will record the SITREPs in the database contained in the Electronic Tablets and transcribe the SITREP onto a Test Incident Response form contained in Annex B, Data Collection Forms.

(2) Hazard to Personnel/Equipment. In the event of an emergent or ongoing threat to the health and/or safety of any test participant/equipment a radio message to all test participants will be relayed via the tactical frequency ordering the immediate cessation of activities. The OTPO will follow the protocols noted in Annex C of this plan and testing will not resume until the all-clear is given by the OTPO.

g. **Posttrial Conduct**

After the mission has ended the post-trial activities will begin to include a summarization of combat and non-combat losses incurred during the mission. The vehicle crews (Blue, Red, and White) will perform post-operations checks and services in accordance with the published maintenance procedures (ref. a). The vehicle crews will download and account for all unexpended ammunition and store in the field ammunition storage point. All Blue Force crews will complete post-mission surveys in accordance with MOS-x.x.

Using the process flow brought forward from the Test Concept, the test team begins to add detail to the trials through written instructions. The instructions include the actions of the operators and the functions of the system as well as test conditions (physical, military, and civil).

The test team writes detailed instructions for a trial to ensure the proper placement and timing of everything and everyone needed for trial success, relying heavily on operational experience and familiarity with unit TTPs.

Annex A: Test Logistics Support

Requirements for Other-Service Assets

Current operational tempo determines availability of other-Service assets, so they may be difficult to procure.

Navy Assets

If a test requires Navy assets, particularly amphibious ships or landing craft, the test team must obtain a Test and Evaluation Identification Number (TEIN) as described in SECNAVINST 5000.2. A specific format exists, and it must be sent via the appropriate Requirements office (DC, CD&I) for endorsement. With the TEIN in hand, the test team completes the Fleet Support Request (FSR) form. Since the East Coast and West Coast Fleet Commands differ in scheduling lead times, the test team must contact the appropriate scheduling coordinators at least 6 months in advance to be included on the scheduling conference notification lists. (This can be accomplished by contacting the current OPNAV N912C Project Officer, who answers to the OPNAV N091 scheduler.) The FSR usually needs to be submitted 6 months ahead, but the actual scheduling conference may occur within 3 months to 1 month of the test date. If another Service is the lead OTA, and the Marine Corps is the only party with an amphibious mission, the test team may have to schedule amphibious operations testing independent of the lead OTA.

Army Air Assets

Air asset requirements present unique challenges. If a test needs Army Air, such as a CH-47, to demonstrate internal or external lift capabilities, the test team should consider the Army Reserves or the Air National Guard. The test team's POC with the Army's OTA may be able to provide contact names and telephone numbers. If MCOTEA is the lead OTA and Army assets are required, they must be requested through the Operational Test Command via the Outline Test Plan (OTP), drafted by the Army POC. (Note: Modifying the OTP, once published, is a difficult and slow process.)

Marine Air Assets

The test team can usually coordinate the use of both fixed-wing and rotary-wing aircraft through the Marine Aircraft Liaison Officer (ALO) for the respective MEF. The MEF normally assigns the duties to a specific squadron and issues the DIRLAUTH for detailed planning. MV-22 Osprey support, however, proceeds differently. If Wing assets are stood up and available, planning will proceed through the ALO at MEF. If a test will use planes from VMX-22, the test team should coordinate with the squadron itself, since they work directly for DC AIR, not the MEF. The MEF G-3 will also be required to issue an authorization for FMF Marines to fly in VMX-22 aircraft, since the Marines belong to MEF and not the aircrafts' command. Scheduling the Wing may require flexibility, so the test team should provide alternate dates/times in the Test Plan. (Note: The test team should also consider that the qualifications and certifications of both the pilots and the ship affect whether the schedule requires shipboard landings. The pilots cannot take off from/land on the ship if their qualifications are not current.)

Air Force Assets

Air Force lift assets can often be arranged through the ALO at the MEF level. The Marine ALO can provide contact names and numbers and may be willing to perform the necessary coordination.

Equipment Requirements and Test Site Coordination

Although a published FOS has identified the Marine Corps forces (MARFOR) (LANT or PAC) that will support the OT, it has not necessarily defined the test location. Depending on the nature of the test, Camps Lejeune and Pendleton may suit a portion of the data requirements, but they seldom provide the extremes needed. Twentynine Palms is popular for desert/hot weather testing, but alternate locations may be less crowded. Several Army Reserve and National Guard sites offer adequate facilities for temperate and cold weather testing, such as Camp Ripley in Minnesota and Fort Pickett in Virginia. Time of year is a factor: Reserve and Guard units book their facilities from May through September for their 2-week training evolutions. Other government and civilian agencies are also potential candidates. Nevada Automotive Test Center in Nevada is an excellent motor vehicle test site.

Other options include Yuma Proving Ground, Aberdeen Test Center, and Naval Surface Warfare Center, Dahlgren. The S-3 has access to the capabilities of various test and training ranges around the country and should be consulted before deciding on the optimal test venue for each test. When using government labs, the test team must obtain a Rough Order Magnitude (ROM) cost proposal because these labs can be expensive. For all test sites, the test team must generate a ROM for site cleanup after the test.

Although the test team may begin informal identification and coordination with the test site, formal coordination is accomplished by the S-3. The S-3 will notify the test team of the test site POC after formal coordination is complete. At that time, the test team will assume responsibility for coordinating with the test site.

After selecting the test sites, the test team should communicate with a POC via telephone/e-mail to ascertain documentation requirements and to schedule a site visit. Although some details can be resolved over the phone, face-to-face contact ensures clear communication. Traveling with a representative from the PM is advantageous because scheduling training facilities and assets at once can save time and money. MOIC attendance on these visits is strongly encouraged.

During the site visit, the test team should attempt to establish POCs for billeting, messing, ranges/training areas, ammunition support (if needed), and network connectivity and should identify any special waivers, certifications, or area-peculiar requirements (e.g., OIC/RSO) certifications, port-a-johns in the field, dunnage collection schedules/costs, frequencies and radios, waivers for privately owned vehicles in the training areas, etc. If the program involves classified documentation or equipment, advance coordination for delivery and storage is mandatory. If the test team coordinates ammunition delivery procedures in advance, the process will be simplified as the test dates draw closer.

The test team should plan to visit the test site at least once more after the initial visit and before the test to finalize and confirm details previously arranged.

Identifying Required Facilities and Logistics Support

The S-4 helps the test team coordinate on-site logistics support for MCOTEA tests. Site visits enable the test team to identify and consolidate administrative and logistics support. Office space and equipment are most commonly needed. Sometimes one source can address phone, fax, and copier requirements as well, but the test team may benefit from shipping the items from MCOTEA to the test site. Maintenance spaces are another frequent issue, and if weapons, classified documents/

items, or serialized equipment are involved, armory or other secure storage facilities are key requirements. If training will be conducted immediately preceding the OT, the PM representative will be interested in scheduling classroom spaces, and the test team will need a place to conduct data collector training.

Data Collection Forms

One of the Data Manager's largest responsibilities is creating the Data Collection Forms that will be used during EOA/OA/OT execution and the final evaluation. The forms may be electronic (created and used in a portable data collection device) or paper-based and filled in manually. If data is collected and stored only in electornic form, a backup of the data must be created as soon as practical to protect against data loss due to electronic malfunction. Working from the Evaluation Framework, the Data Manager develops forms to collect each program's data requirements and to resolve all defined Measures.

Types of Forms

The DM creates forms to capture all of the requirements outlined in the Test Plan. Form structure is based on the types of Measures contained in the SEP. The relation of Measures to forms is illustrated below.

Quantitative Forms

Quantitative forms are used to collect numerical data, e.g., RAM and TIRs.

Qualitative Forms

Qualitative forms collect the ratings and comments of the operators and SMEs and are written as Survey Questions.

Verification Forms

Verification forms collect data for the purpose of proving that items exist or are included with the system to be tested. Additional forms may be created to characterize the operational test environment. While each form may be adapted to the particular event, certain reference information must appear on every form: e.g., the item being tested, operator ID, date, and time. From there the forms are designed to capture requisite information: for example, Test Incidents, RAM, Maintenance, Demographics, and Operations Log. Other forms that could be developed include Inventory Control, Weather Log, Information Assurance, and Crew Assessment. While a few basic forms (Operations Log, TIR, and Weather) may be similar, most of the forms must be built to capture program-specific data to answer the Measures in the SEP.

The DM must ensure that the forms flow logically and are easy for a data collector in the field to follow. Each set of forms is program-specific and will vary greatly in design and depth of data collected.

Survey Questions

Survey questions are the primary method of collecting qualitative data; each qualitative Measure has questions assigned to it. The DM works with the test team and an SME (e.g., the Human Factors and Safety SME) to develop the questions. The basis for questions can derive from the SEP, the Request for Clarifications, and the OMS/MPs. Another option for creating a survey is to perform a structured interview, in effect an open forum that asks the operators to state their opinions about the system in a structured way.

MCOTEA prefers using more quantitative data sources, but surveys can be useful in finding issues for further analysis and in helping to identify risks.

Data Collector Training

Data Collector (DC) training is the opportunity to provide instruction to the collection team on the purpose of the test and their role in it. DC training is usually done at the test site after the arrival of all personnel. This should occur a couple of days before the Pilot Test.

Everyone should understand that the purpose of the test is NOT to make the system work, but to obtain unbiased data on its performance, given the crew training and operating conditions particular to the event. DCs should understand that they are to gather the data requested on the forms, but not attempt to analyze or interpret the data or interfere with operators using the system.

Data collector training focuses on training the DCs to accomplish their mission in the test. This includes going over each data form in detail, paper or electronic. The instructors are usually the test team members responsible for creating the forms. A substantial portion of the training should be dedicated to practical application. If automated data collection is employed, the instrumentation supporting the automation should be used as an integral part of this training. The team should discuss the forms with the DCs and solicit their recommendations on such items as terminology, so that changes can be made and validated with the DCs before the test begins. The instructors should make notes of all questions asked and the responses given by the instructors to aid in consistency throughout the test. The Data Collector Handbook should be covered in the training. DCs may then use this reference book throughout the test.

Environmental Considerations

Data collection efforts on an operational test must occur in day or night, rain or shine, wet or dry, cold or hot, etc. The operating environment will affect the choices of data collection methods. Things to consider when choosing data collection methods include the following:

- Visibility (natural light or availability of sources of artificial light). Data collection under low light or no light situations presents unique challenges. Depending on the method of collection, paper for example, a data collector would need an artificial source of light to collect data. Care should be taken to ensure that the artificial light source does not interfere with the operations of the system under test or its operators. When using electronic means to collect data, the same holds true, except that the electronic means are often sources of light.
- Precipitation (rain, freezing rain, sleet, snow, none). Depending on the environment, data collection methods need to be resistant to precipitation. Waterproof paper and ruggedized data collection devices are available to protect data collection efforts.
- Temperature (cold/hot). Cold and hot environments can make data collection difficult. Electronic devices can fail in extreme cold and heat. Likewise, clothing designed for inclement weather may make paper data collection difficult to accomplish.
- Data Collection Mobility. Another serious consideration is whether data must be collected on-the-move. Movement by foot or vehicle can make collecting data very difficult. It is difficult to write or tap touch screens effectively while on-the-move.

Data Collection Based on Data Requirements

What is being measured and the data requirements themselves often dictate how the data is to be collected. For example, if "elapsed time" was the data requirement, then the analyst may choose to instrument the trial with a stopwatch. However, if "Time Start" and "Time Stop" are the data requirements, then the analyst may choose to instrument the trial with a device that creates time stamps for events, such as a ruggedized PDA.

Building a Data Repository

Once all data requirements have been developed, the DM builds an electronic data repository, an electronic medium for storing the collected data. The preferred method is a database, although spreadsheets may be used for smaller tests. All test data, including the data collected on paper forms, must be placed into the data repository for appropriate analyses. The repository must

be able to support the analytic requirements of every Measure for the test; if data to support every Measure is not included in the repository, the repository is inadequate.

Maintaining Data Integrity and Security

The test team DM is responsible for maintaining data integrity (completeness, correctness, and noting caveats associated with data elements) and security (no unauthorized changes). Limiting access to the repository through password protection maintains data security as does limiting write privileges inside the repository.

Data Collection Verification and Validation

The test team conducts a Data Collection (DC) Verification and Validation (V&V) to ensure that the Test Plan is adequate to support the data collection effort. Data collection, including the collection equipment, should be verified and validated before use in the actual test. DC V&V is performed once data collection methods and the data repository have been constructed and prior to the Test Plan Consolidated Review Board (CRB). Accordingly, the test team plans and conducts a DC V&V exercise that confirms the adequacy of the Test Plan Detailed Schedule, data collection methodology, and ensures that the data collection equipment functions properly and reliably.

All systems that will support data collection for the operational test must be programmed and present at the DC V&V (automated data collection devices, survey computers, primary forms, instrumentation or appropriate data from instrumentation, etc.). The DC V&V should include as many members of the test team as permissible, but at a minimum the DM responsible for data collection during the test, the TM, and ORSA should be present. Following the DC V&V, if the test team does not discover any issues, the items should be ready to ship to test. If the test team discovers issues, they should repeat the DC V&V following corrections (the test team can tailor the DC V&V to focus on the issues they discover). The Operational Test and Analysis Division (OTAD) is the approving authority for the DC V&V and provides a detailed process checklist to support the preparation and execution of the event.

Test Site Visit

Even if the test team has visited the test site earlier in the planning, another visit should occur at least 2 weeks before test to confirm the following:

- Dining and sanitary facilities are ready
- Range regulations have not changed
- Corpsman is available, if needed
- All shipping/receiving details are arranged
- Coordination with key staff officers in the host organizations and the Base Public Affairs Office has occurred
- Other range users and stakeholders know how the test may affect them (range closures, etc.)

Equipment Check

To prevent delays once testing begins, the test team should arrange to have limited technical inspections (LTI) and operations checks for all major test support systems and equipment before the items are transported from the providing commands to the test site. This can be as simple as ensuring that a generator is working or a road wheel on a vehicle will last for the duration of the test. No equipment should arrive at the test site that may require major preventive maintenance during test. Specific equipment configuration requirements should also be confirmed.

Instrumentation

Rehearsals of instrumentation setup, operation, and teardown should be conducted at least 2 weeks before test. Validation and data reduction procedures

for video data should be rehearsed before the Pilot Test, allowing adequate time to adjust instrumentation schematics and collection plans, if necessary.

Transporting the Test Team and Test Equipment

The S-4 helps the test team coordinate transportation to the test site. If many test participants are involved and the site is not within motor transport range, air transportation becomes the most viable option. The ALO at MEF can assist here. Although C-130 transport (USMC or Air Force) is ideal, these aircraft are usually overbooked and unavailable. Air Force transport (C-5, C-17, etc.) is possible: the ALO may be able to coordinate with the Air Force counterpart to inquire into aircraft availability. Commercial charter transportation may be the best option. The test team should coordinate with the Traffic Management Office (TMO) and provide a detailed roster, but this requires travel orders per Fiscal's guidance. Local transportation at the embarkation and debarkation points must still be arranged, but the local base transportation office can provide buses (military or civilian) for that purpose. In-and-around transportation will depend on the size of the test contingent. For groups of less than 50, test participants can drive rental vans. A regular bus schedule can be arranged through the Regional Transportation Facility (RTF) for larger contingents.

Note: If the test team uses commercial (rental) vans, the OTPO must procure a release from the RTF stating that government vans are not available. Upon receipt, the local Base Comptroller will generate a contract so that the Marine test participants will not be charged.

Travel Orders

Travel Orders for test participants, should they be needed, should be coordinated with MCOTEA fiscal.

Transporting Test Equipment

Normally the PM is responsible for transporting the test equipment to and from the test sites. The test team should coordinate with the PM's representative to arrange for the equipment's timely arrival on location. After equipment arrival at the test site, the test team should conduct a joint LTI (with PM and MCOTEA representatives) to ensure that nothing was damaged in transit. If training is scheduled for immediately before OT, the PM will probably need to use maintenance facilities to prep the articles. MCOTEA test equipment destined for the test site, including any electronic data collection devices and laptops, is usually boxed in secure shipping containers and sent via TMO. The test team can obtain the requisite documentation from the S-1, including documentation for the return trip.

Site-Specific Restrictions

When arranging travel plans, the test team must consider site-specific restrictions. For example, winter travel to the Cold Regions Test Center (CRTC) in Alaska includes a flight to Fairbanks and a drive to CRTC. However, travelers must remain overnight in Fairbanks if they arrive after 1500 because authorities discourage traveling the 100 miles in the icy darkness. The test team must locate adequate billeting for any test participants arriving after 1500. In addition, special permission is required for Marine use of 15-passenger vans.

Marine Corps Equipment

Finally, the test team must consider the availability of routine Marine Corps equipment. If a host unit is assigned as test support, that unit normally provides required assets, i.e., MTVRs, HMMWVs, weapons (M2, MK19, etc.), radios, etc. MCOTEA covers repairs, fuel, etc., as test costs. If no host unit exists, the test team should inquire into the existence of an equipment allowance pool, such as the one at Twentynine Palms. A good LTI will

help keep repair costs lower at the end of the test. The FOS should have identified these assets, and discussions with higher headquarters during the planning process should have identified the source.

Test Funding

During the site coordination visit, the test team must visit the base/facility comptroller to identify a POC. At most bases the test funds will be sent via funding document to the comptroller, who will be the central paymaster for test expenses. However, the Base comptroller cannot cross accounts, meaning that Base can cover expenses that most functional areas generate except those related to the Marine Division. If, for example, host unit equipment (a Division asset) needs repair, those funds must be filtered through the Division Comptroller. MCOTEA needs clarification of the various expense channels as early as possible, so the test team must provide the contact information (POC, telephone, and fax numbers) to the MCOTEA Fiscal section.

STEP 4 Operational Test Execution

Task	Responsible	Provides Input	Task Output	Signature/ Release Authority
Observe M&S V&V Event, if accrediting M&S	TM	OTPO	Completed Observation Plan Forms	
Observe Developmental Test Event, if performing Intermediate Assessment	TM	OTPO	Situation Report	OTPO
			Completed Observation Plan Forms	
Conduct Movement to Test Site for QRA, OA, or OT	TM	OTPO, S-4	Checked-off Equipment List from Test Plan	
Perform Pretest Setup and Coordination	TM		Situation Report	OTPO
	DM	OTPO, TM, MS	Daily Data Report readiness	
New Equipment Training	TM	OTPO, DM, MS	Situation Report	OTPO
Data Collector Training	TM	OTPO, DM, MS	Situation Report	OTPO
Execute Pilot Test	TM	Military SME, MS	Situation Report	OTPO
	DM	OTPO, TM, MS	Daily Data Report	DM
		Military SME, TM, ORSA, MS	Scoring Conference Results	OTPO
Execute Record Test	TM	Military SME, MS	Situation Report	OTPO
	DM	OTPO, TM, MS	Daily Data Report	DM
Hold FD/SC Scoring Conference and Data Review	DM	Military SME, TM, ORSA, MS	Scoring Conference Results	OTPO
	OTPO	Military SME, TM, ORSA, MS, DH, Exec	CRM for Scoring Conference Results	
Conduct Data Review	DM	OTPO, TM, MS	Situation Report	OTPO

Chapter 3-4. Step 4: Operational Test Execution

Pretest Activities

New Equipment Training

New Equipment Training (NET), including maintenance training, is typically the first official event of the OT and should occur immediately before the Pilot Test. It is the only OT event that involves the PM and is, in fact, the PM's responsibility. Although directed to the operators and maintainers participating in the test, all test team members should attend NET. Any materials used in the NET and subsequent operational test must conform to the requirements of MCO P5215.17. Operational Test must not begin until operators and maintainers are properly trained on the functions of the system and can use it in an operational environment.

DC Training

Data Collector (DC) training is the opportunity to provide instruction to the collection team on the purpose of the test and their role in it. DC training is usually done at the test site after the arrival of all personnel. This should occur a couple of days before the Pilot Test.

Everyone should understand that the purpose of the test is NOT to make the system work, but to obtain unbiased data on its performance, given the crew training and operating conditions particular to the event. DCs should understand that they are to gather the data requested on the forms, but not attempt to analyze or interpret the data or interfere with operators using the system.

Data collector training focuses on training the DCs to accomplish their mission in the test. This includes going over each data form in detail, paper or electronic. The instructors are usually the test team members responsible for creating the forms. A substantial portion of the training should be dedicated to practical application. If automated data collection is employed, the instrumentation supporting the automation should be used as an integral part of this training. The team should discuss the forms with the DCs and solicit their recommendations on such items as terminology, so that changes can be made and validated with the DCs before the test begins. The instructors should make notes of all questions asked and the responses given by the instructors to aid in consistency throughout the test. The Data Collector Handbook should be covered in the training. DCs may then use this reference book throughout the test.

Test Execution

Test execution is a team effort. Test team personnel and the MOIC of the Operating Forces must continuously coordinate their activities to ensure that all test events are executed and that all necessary data is collected. If necessary, the test team must take the time to adjust the schedule to ensure that all test and data collection objectives are met during the test events, not afterwards. This constant coordination often results in long days for the test team, who will arrive first and depart last each day.

Pilot Test

After NET and immediately before the Record Test, the test team executes the Pilot Test. By way of review, the Pilot Test is both a rehearsal and a readiness check for the Record Test; it is an abbreviated version of the Record Test conducted to detect deficiencies in the planning, instrumentation, data collection, data management, and test control. In addition, it provides an opportunity to prove out the daily battle rhythm and logistics support plan.

Operator Learning Curve During Pilot Test

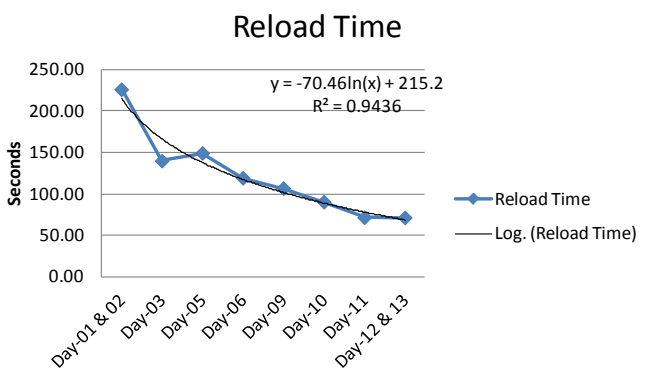
From an analyst's perspective, an operator should be proficient in using the system under test, not just trained. An operator who is merely trained will provide data, but the data will be grounded in a learning curve, diminishing the true measure of mission accomplishment.

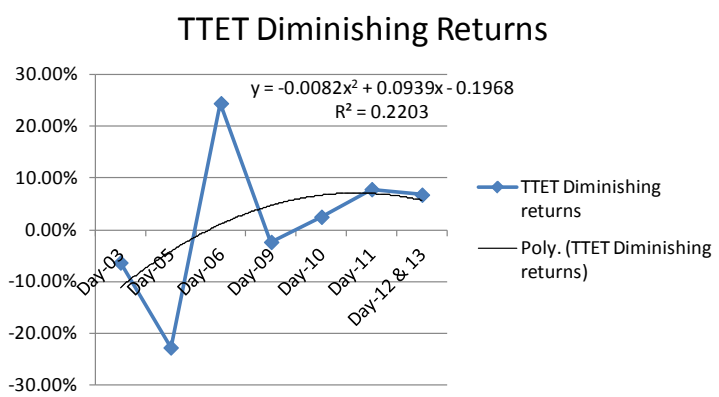
Conversely, a long Pilot Test allows an operator to increase his level of knowledge and expertise, burn through the learning curve, and emerge into proficient status. Data collected from that point forward is truer to mission accomplishment than data collected from lesser-trained operators. Tests relying heavily on instrumentation may require additional time after the Pilot Test to correct problems.

The process of plotting and analyzing learning curves should be defined in the Test Plan Data Analysis Method.

The example below depicts results from a Pilot Test for an automatic rifle. In figures 3-4-1 and 3-4-2, the solid black lines trace the learning curves for Reload Time and Total Target Exposure Time (TTET) Diminishing Returns, respectively. Once the data collected from the operators becomes less erratic and more level, the Pilot Test is considered complete. By the end of the Pilot Test, operators' reload time had improved significantly. The learning curve for Total Target Engagement Time, the test's most critical element, reflected continuous learning among all test team members as they operated more efficiently and adapted to new procedures. The curve leveled off early in the Record Test, leading to data that more accurately reflected mission accomplishment.

Problems revealed during the Pilot Test must be corrected before the Record Test. This may involve additional training, support, resources, or changes to the Test





Figures 3-4-1 and 3-4-2 Pilot Test Learning Curves

Plan. Enough workdays should have been scheduled between the end of the Pilot Test and the beginning of the Record Test to incorporate any changes.

Pilot Test Data Collection

During the Pilot Test, data should be collected and managed in the same manner by the same personnel who will execute the Record Test. The test team should perform a complete end-to-end data run, beginning with test events and going through every step until the data collection process has been operationally validated. All data collection forms, including manual backups, must be validated. (Although data collection forms will have been validated during the Data Collection V&V in Step 3, the Pilot Test will reveal any shortcomings that V&V might not have been able to predict.) In addition, all instrumentation, from stopwatches to computers, must be used. The conduct of Pilot Test should mirror the exact conduct of Record Test. For example, if the test involves a two-shift operation then both shifts must be run during Pilot Test to include data reviews and debriefings at shift changes.

Data collected during Pilot Test may or may not be considered valid. This is particularly true for RAM data. Use of this data must be in accordance with the approved SEP and FD/SC Charter. The Pilot Test data must be comparable and compatible with the Record Test data. Any Pilot Test data used in the Test Data Report must have been obtained under the same test conditions as the Record Test.

The Statistician should analyze Pilot Test data to determine if a learning curve exists by plotting the results of individual trials for comparison.

After completion of the Pilot Test, the test team briefs the Director on the results of Pilot Test execution and data collection and analysis. The Director authorizes the test team to proceed with the test or delay the start of Record Test in order to correct Pilot Test deficiencies.

Record Test

The Record Test is the culmination of all test planning activities; it executes the Test Plan and accurately collects the resulting test data. The Record Test generates daily data results and Situation Reports (SITREP). Data results eventually populate the Test Data Report.

At a minimum, the Record Test daily schedule should include the following best practices:

Hold a Morning Brief

The OTPO, along with the MOIC, begins each test day with a brief to the entire test team to provide an overview of the day's goals, a discussion of quality control issues, and an opportunity for questions, comments, or recommendations. In addition, the Data Manager should provide the OTPO with a daily, minimum Level 4 data summary. The summary should include preliminary results of any Measures exercised to date, RAM timelines and dendritics, and TIR summaries (to include MOIC comments and pre-scoring). By reviewing data daily, the OTPO is better equipped to address any data collection and test execution issues. Without current information, the OTPO cannot adjust test events to ensure that adequate data is collected. If there is no data, there is no test, no matter how well the events are conducted.

Manage Test Deviations

A test deviation is any departure from the approved Test Plan, no matter how minor. A primary responsibility of the OTPO is to ensure that the test is conducted in accordance with the MCOTEA/DOT&E approved Test Plan. Unauthorized deviations from the Test Plan are not permitted.

Conducting the test according to an approved Test Plan eliminates the perception of bias or of rigging the test to

ensure positive results and also ensures that the proper data is collected to answer the Test Plan questions.

However, despite the best intentions of test planners, the reality of test execution often requires some deviation from the Test Plan once the test is underway. To address these potential deviations and retain testing integrity, MCOTEA follows a strict procedure for approving changes to the Test Plan, as described below.

If the test team believes that a deviation from the Test Plan is required during the Pilot Test or Record Test, then the OTPO, with support from the test team, must

- Identify the deviation from the plan
- Identify the effect of the deviation
- Formulate in writing an alternate plan, or document proposed changes to the existing plan
- Obtain approval for the changes before execution from the Director, MCOTEA

Regarding schedule deviations, as the days progress, especially after long days or during lengthier tests, it may be tempting to deviate from the planned schedule (e.g., wait until the next day to debrief data collectors, put off review and comment on TIRs, do away with preparations for the next day's events, put off scheduled survey sessions until later in the test, etc.); however, even minor schedule deviations can be counterproductive and should be avoided unless absolutely necessary. A disciplined approach to maintaining the daily schedule helps ensure overall test success.

Manage Site Visitors

The OTPO should consider planning a VIP day for high visibility programs. A VIP day provides the OTPO with the best control of VIPs and minimizes disruption to the test.

All VIPs or other visitors to the test site must complete a Non-Disclosure Agreement; the only exception to this is DOT&E personnel or their representatives. All visitors must be escorted to ensure that their presence does not interfere with or affect the integrity of testing.

Personnel from the PM Office (to include the developing contractor) must be kept at arm's length from the test participants. While an "open book" test is in the best interest of all parties and PM personnel should be free to witness any or all events, they must not be allowed to coach test participants or influence the test in any way. PM personnel may only participate in a test when the approved Test Plan calls for PM involvement in operation, maintenance, or other support of the system when deployed in combat.

Validate Data Daily

Figure 3-4-3 identifies the levels of data MCOTEA processes. At a minimum, the Data Manager must reduce and order the raw data (Levels 2 and 3) daily. In addition, the Mathematical Statistician, with support from the DM, processes findings and summary statistics (Level 4) each day to support the daily briefs to the OTPO.

Transfer Data to Analysts

The Data Manager should transfer the data (Level 4) electronically to the supporting analysts at regular intervals for review and higher-level processing.

Conduct Daily Hot Wash

Similar to the morning brief, the OTPO and MOIC conclude each day with a hot wash to recap the day's events, capture information for the SITREP, discuss any lessons learned, and provide further time for questions, comments, or recommendations.

Write Daily Situation Reports

The SITREPs must contain, at a minimum, planned and actual data collection, and must highlight any problems with test execution. At each day's end, the OTPO completes and forwards the SITREPs

Seven Levels of Data				
Level	***Description***	***Possible Forms***	***Examples of Content***	***Disposition***
Level 1 ***Raw Data***	Data in its original form. Results of field trials just as recorded. Team Lead: DM	Complete data collection sheets, raw video/audio, original instrumentation output, completed questionnaires, and/or interview notes.	1. All reported target presentations and detection. 2. Clock times of all events. 3. Azimuth and vertical angle.	Accumulated during trials for processing. Not published.
Level 2 ***Reduced Data***	Data taken from the raw form and consolidated. Invalid or unnecessary data points identified. Trials declared "No Test" identified. Team Lead: DM	Confirmed and corrected data collection sheets, video/audio with extraneous material removed, invalid trials filtered out.	1. Record of all valid trials. 2. Start and stop times of all applicable events. 3. Computed impact points of each round flashed. 4. Confirmed interview records	Produced during processing. Not published.
Level 3 ***Ordered Data***	Data that have been checked for accuracy and arranged in convenient order for handling. Operations limited to counting and elementary arithmetic. Team Lead: DM	Spreadsheet, tables, ordered and labeled printouts, edited video/audio.	1. Counts of detections arranged in sets showing conditions under which detections occurred. 2. Elapsed times by type of event. 3. Impact points of rounds by condition under which fired. 4. Interview comments categorized by type.	Provided to the analysts daily. Published as the basic factual findings of the test (i.e., Test Data Report).
Level 4 ***Findings or Summary Statistics***	Data that has been summarized by elementary mathematical operations. Operations limited to descriptive summaries without judgments or inferences. Does not go beyond what was observed in the test. Team Lead: DM/Statistician	Tables or graphs showing totals, means, medians, modes, maximums, minimums, quartiles, percentiles, curves, or standard deviations. Qualitative data in form of lists, histograms, counts by type, or summary statements	1. Percentage of presentations detected. 2. Mean elapsed times. 3. Calculated probable errors about the centers of impact. 4. Bar graph showing relative frequency of each category of comment.	Published as the basic factual findings of the test (i.e., Operational Test Agency Evaluation Report (OER)).
Level 5 ***Analysis or Inferential Statistics***	Data resulting from statistical tests of hypothesis or interval estimation. Execution of planned analysis data. Includes both comparisons and statistical significance levels. Judgments limited to analysts' selection of techniques and significant levels. Team Lead: Statistician	Results of primary statistical techniques such as T-tests, Chi-square, F-test, analysis of variance, regression analysis, and other associated confidence levels. Follow-on tests of hypotheses arising from results of earlier analysis, or fallback to alternate nonparametric technique when distribution of data does not support assumption of normality.	1. Inferred probability of detection with its confidence interval. 2. Significance of difference between two mean elapsed times. 3. Significance of difference between observed probable error and criterion threshold. 4. Magnitude of difference between categories of comments.	Published in system evaluation reports (i.e., OER).
Level 6 ***Extended analysis or operations***	Data resulting from further analytic treatment going beyond primary statistical analysis, combination of analytic results from different sources, or exercise of simulation or models. Team Lead: ORSA	Insertion of test data into a computational (decision) model or a combat simulation, aggregation of data from different sources, curve fitting and other analytic generalization, or other operations research techniques such as application of queuing theory, inventory theory, cost analysis, or decision analysis techniques.	1. Exercise of decision models to determine effectiveness, suitability, and survivability. 2. Computation of probability of hit based on target detection data from test combined with separate data. 3. Determination of whether a trend can be identified from correlation of data from different sources. 4. Delphi technique treatment of consensus of interview comments.	Published as appropriate in system evaluation reports or follow-on reports (i.e., OER).
Level 7 ***Conclusion or Evaluation***	Data conclusions resulting from applying evaluative military judgments to analytic results. Team Lead: OTPO	Stated conclusions as to issues, position statements, and challenges to validity or analysis. Military impact of results.	1. Conclusion as to military worth. 2. Translate quantitative results to military implications.	Published as the basic evaluative conclusions of system evaluation reports (i.e., OER).

Figure 3-4-3 Seven Levels of Data

to the Division Head as well as others as directed.

Release of Test Data

Before analysis is complete, test data is released only with the approval of the Director, MCOTEA. Data release is restricted because conclusions drawn before test results are completely analyzed can be highly misleading. Moreover, an assessment based on an incomplete set of test data can cause biases that are difficult to overcome, even when further information proves the initial analysis to be correct (or incorrect).

Data Reduction

Data reduction, while technically a posttest activity, actually begins during Pilot Test and continues throughout Record Test. Initial analysis may be performed as data is reduced, but these results are of limited value because each subsequent data point obtained has the potential to change the analytic results. Therefore, the test team's primary focus, specifically the MS/DM, is to ensure that test data is reduced and reported *each day*.

Posttest Activities

FD/SC Scoring Conference

MCOTEA convenes the FD/SC Scoring Conference after the Record Test has ended and before the test team leaves the test site. MCOTEA, DC, CD&I, and the Program Manager each provide a representative to the conference; the OTPO represents MCOTEA and serves as chair. (The OTPO may also schedule intermediate scoring conferences during the Record Test, especially during a long test or one with many TIRs.)

Scoring Conference participants use the guidance contained in the system's FD/SC Charter, which was developed early in the test planning process. The conference members review, classify, and then discuss the scoring of all TIRs, which support evaluation of RAM.

The OTPO should ensure the nearby presence of essential personnel to respond to questions or to clarify TIRs. In addition, the OTPO ensures that the following are available for the conference:

- MOIC of the Operating Forces for pre-scored TIRs and comments regarding them
- A summary of TIRs for each member of the conference. (Conference members should review each TIR to date and determine a preliminary score before the conference begins.)
- A summary of maintenance and times (Start Time, Stop Time, and Maintenance Time)
- Copies of the FD/SC Charter
- System description, system mission, mission time, crew correctable maintenance actions, and mef definitions

Conference members score and classify the TIRs by examining the circumstances surrounding each test incident and deciding the classification, chargeability, and hazard/risk assessment for each incident. Refer to the RAM section of volume II for a detailed list of these categories.

Incidents may be left unscored until additional information becomes available to support a scoring decision. Previously scored incidents may be re-examined to consider additional information.

The OTPO documents the results of the Scoring Conference in the minutes. Any conference member may provide a written dissenting opinion on any incident scoring result. The OTPO must include any dissenting opinions in the conference minutes.

Developmental contractors are prohibited from being involved in any way in the performance assessment or evaluation activities of an operational test (OT&E of Defense Acquisition Program 2008). Accordingly, developmental contractors are not invited into the Scoring Conferences as observers or participants. However,

developmental contractors can be requested to present information concerning system design or intended implementation procedures, but they must leave immediately after providing information or answering any questions and before further discussion of TIRs ensues. Only the Director, MCOTEA may release operational test data, including Scoring Conference results. Conference members may not disclose any details of the Scoring Conferences without the Director's approval.

Test Site Closeout

Test Site Closeout involves two areas, Administration and Logistics. In both areas, accountability is of primary importance. Accountability applies to equipment, supplies, and personnel.

To close out the test site, the OTPO/TM are responsible for the following:

- Fuel: all fuel keys issued to support the test are returned to the appropriate POC and receipted for.
- Billeting: All linen is returned; spaces are walked through with billeting manager.
- Messing: Manager is notified that the test is complete and all supporting Marines have departed. Important: Administrative process to reinstate Commuted Rations to eligible Marines has been initiated.
- Supplies: All ServeMart cards turned in, unused supplies returned to MCOTEA, and inventory provided to the S-4.
- MIPRs: Open MIPRs are closed out with the test site fiscal office (may not be complete before leaving, but the fiscal office should know when the test was completed and when fuel and other charges should be terminated.
- Range Control: Range facilities returned to their pretest condition. Complete inspection of the range conducted; everything returned to its original location. Closeout inspection with range control conducted and all access keys, etc., returned. Range control notified at final departure that the test is complete and all personnel are accounted for.
- Communications: COMM notified that the test unit has exited the net and no longer requires assigned frequencies. COMM load frequencies (BEARMAT for example) on MCOTEA communications equipment have been removed.
- SharePoint: Close out of SharePoint site access created for Non-MCOTEA test members.
- OCONUS Communications: Notification to S-4 regarding end of communication services in support of overseas tests
- Classified Material: All classified documents accounted for and returned.
- Data Removal: All test data, including Personally Identifiable Information, removed from support computers and data collection devices other than those that will remain under test team control and are needed for data reduction and reporting.
- Leased Equipment: Any leased equipment (printers, copiers, etc.) returned to vendors, done through face to face turnover to limit MCOTEA's exposure to future claims of damage.
- Supporting Unit Equipment: Supporting unit vehicles, generators, etc. returned to the supporting unit, fueled and POL levels topped off. Joint posttest inspection performed to limit claims of damage. Test team has worked with the MOIC to ensure that all equipment used to support the test is returned Stocklist (SL)-3 complete, clean, and properly stowed.
- Test Item Limited Technical Inspection: Test items inspected before return to Program Manager to capture last minute Test Incident Reports and to document the completeness of the test items (including SL-3 items).
- Transportation of test support equipment arranged for with Transportation Management Office. (Early coordination recommended.) Copies of shipping documents obtained for MCOTEA S-4 and Fiscal section use.
- Personnel: All personnel accounted for and safely returned to home station.

Data Review

The Data Review is held to provide early approval and guidance to the test team, specifically the Mathematical Statistician, on the adequacy and accuracy of the data analysis. The Data Review occurs after the completion of OT data analysis (Level 5) and the FD/SC Scoring Conference (At least 30 calendar days before publishing any assessment or evaluation report). The Scientific Advisor leads a panel that includes the OTAD Head and the Division Head for the Test Division. All members of the test team involved in preparation of the Test Data Report should attend the Data Review, which allows the panel members to discuss their concerns, investigate test data, and review analytical methods. All issues related to data analysis or analytical methods must be resolved before reporting final Measure results.

This page intentionally left blank for two-sided printing.

STEP 5 Operational Test Data Reporting

	Task	Responsible	Provides Input	Task Output	Signature/ Release Authority
Prepare DT Observation Report if observation has been performed					
	Write	TM	OTPO	DT Observation Report	DH
	Request DT Report			DT Report Request Letter	
	Review DT Report		OTPO, MS	General Administration and Management (correspondence) new?	
Prepare Test Data Report for QRA, OA, or OT					
	Write	MS	OTPO, TM, DM	Draft System Assessment TDR or Draft OTDR	Exec
	Review/Comment	OTPO	Military SME, TM, ORSA, MS, DH, OTAD, Exec	CRM	
	Adjudicate CRM			CRB signature-ready Test Data Report	
	Prepare post-CRB Copy			Signature-ready Test Data Report	
	Approve			Signed Test Data Report	Exec
Review and Comment on M&S Verification/Validation Report (if program calls for)		OTPO	ORSA, MS	V&V Report (and associated summary data annexes)	OTAD

Chapter 3-5. Step 5: Operational Test Data Reporting

Background

The outcome of operational testing is the data set, which can be quite large, containing numerous columns and rows of information. As discussed in step 4, Operational Test Execution, the DM and MS have been organizing and reducing raw data daily, as it is recorded, rather than waiting until the end of test.

Test Data Report

Using MCOTEA's standard Test Data Report template, the MS prepares a Test Data Report during posttest activities. The report's purpose is to record any deviations from the Test Plan and to package the data on a CD for an early, unanalyzed look. The report does not evaluate the results or reach conclusions about OE, OS, and OSur.

The Test Data Report is signed by the Director, MCOTEA and sent to DOT&E for programs on oversight. Otherwise the report is released solely at the discretion of the Director, MCOTEA.

Test Data Report Examples

Figure 3-5-1 presents a sample set of raw and reduced data for MOP-1.1 of the ABC system used as an example throughout the 6-step process in this chapter. The table contains 183 rows of data (truncated for presentation) for the force-on-force exchanges used to measure the Target Engagement Cycle. The data in this table is both raw (columns not shaded) and reduced (columns shaded gray).

Figure 3-5-2 presents the 48 trials for Force Superiority (MOE-1). This data was generated using the MOP-1.1 raw and reduced data. The data set is typical of what a test team should expect to have in hand before departing the test site.

Engagement #	Trial Number	Illumination	Target	Distance	Force	BF Detect	BF identify	BF First Shot	BF Time of Flight1	BF Shot1 Hit/Miss	BF Second Shot	BF Time of Flight2	BF Shot2 Hit/Miss	BF TEC	BF End State	Force	RF Detect	RF identify	RF First Shot	RF Time of Flight1	RF Shot-1 Hit/Miss	RF Second Shot	RF Time of Flight2	RF Shot2 Hit/Miss	RF TEC	RF End State	First Detect	First ID	First Shot	Duel Winner		Time to Acq	Time to Fire 1st Rd	t_m	TEC
1	1	N	Ar	Mid	BF	12:00:30 AM	12:01:26 AM	12:01:50 AM	12:02:05 AM	Hit				12:01:36 AM	Dest	RF	12:00:29 AM	12:01:31 AM	12:01:36 AM	12:01:51 AM	Hit				12:01:22 AM	Dest	RF	BF	RF	RF		56	24		96
2	1	N	Ar	Mid	BF	12:04:17 AM	12:05:28 AM	12:05:52 AM	12:06:07 AM	Hit				12:01:50 AM	Dest	RF	12:04:29 AM	12:05:35 AM	12:05:40 AM	12:05:55 AM	Hit				12:01:38 AM	Dest	BF	BF	RF	RF		70	25		110
3	1	N	Ar	Mid	BF	12:06:48 AM	12:07:56 AM	12:08:22 AM	12:08:43 AM	Miss	12:08:48 AM	12:09:09 AM	Hit	12:02:48 AM	Dest	RF	12:06:21 AM	12:07:19 AM	12:07:23 AM	12:07:44 AM	Hit				12:01:23 AM	Surv	RF	RF	RF	RF		69	26		168
4	1	N	Ar	Mid	BF	12:09:27 AM	12:10:29 AM	12:10:55 AM	12:11:02 AM	Hit				12:01:35 AM	Dest	RF	12:09:44 AM	12:10:41 AM	12:10:45 AM	12:10:52 AM	Miss	12:10:59 AM	12:11:06 AM	Hit	12:01:39 AM	Dest	BF	BF	RF	BF		62	26	6	95
5	2	N	FFP	Near	BF	12:09:52 AM	12:10:52 AM	12:11:16 AM	12:11:23 AM	Hit				12:01:31 AM	Dest	RF	12:09:51 AM	12:10:51 AM	12:10:56 AM	12:11:03 AM	Hit				12:01:12 AM	Surv	RF	RF	RF	RF		60	24		91
6	2	N	FFP	Near	BF	12:10:14 AM	12:11:18 AM	12:11:42 AM	12:12:03 AM	Miss	12:12:09 AM	12:12:30 AM	Hit	12:02:18 AM	Dest	RF	12:10:11 AM	12:11:15 AM	12:11:19 AM	12:11:40 AM	Hit				12:01:29 AM	Surv	RF	RF	RF	RF		64	24		138
7	2	N	FFP	Near	BF	12:11:43 AM	12:12:36 AM	12:13:03 AM	12:13:10 AM	Miss	12:13:17 AM	12:13:24 AM	Hit	12:01:53 AM	Dest	RF	12:11:31 AM	12:12:35 AM	12:12:40 AM	12:12:47 AM	Hit				12:01:16 AM	Surv	RF	RF	RF	RF		53	27		113
8	2	N	FFP	Near	BF	12:11:37 AM	12:12:37 AM	12:12:59 AM	12:13:14 AM	Hit				12:01:37 AM	Dest	RF	12:11:53 AM	12:13:02 AM	12:13:08 AM	12:13:23 AM	Hit				12:01:46 AM	Dest	BF	BF	BF	BF		60	22		97
9	3	D	FFP	Mid	BF	12:12:42 AM	12:13:46 AM	12:14:10 AM	12:14:25 AM	Hit				12:01:45 AM	Dest	RF	12:12:40 AM	12:13:38 AM	12:13:43 AM	12:13:58 AM	Hit				12:01:17 AM	Surv	RF	RF	RF	RF		64	24		105
10	3	D	FFP	Mid	BF	12:15:07 AM	12:16:09 AM	12:16:36 AM	12:16:51 AM	Hit				12:02:16 AM	Dest	RF	12:14:35 AM	12:15:38 AM	12:15:44 AM	12:15:59 AM	Miss	12:16:05 AM	12:16:20 AM	Hit	12:01:44 AM	Surv	RF	RF	RF	RF		62	27	6	136
11	3	D	FFP	Mid	BF	12:14:51 AM	12:15:51 AM	12:16:14 AM	12:16:29 AM	Hit				12:01:39 AM	Dest	RF	12:15:14 AM	12:16:16 AM	12:16:21 AM	12:16:36 AM	Hit				12:01:46 AM	Dest	BF	BF	BF	BF		60	24		99
12	3	D	FFP	Mid	BF	12:15:32 AM	12:16:31 AM	12:16:55 AM	12:17:02 AM	Hit				12:01:30 AM	Dest	RF	12:15:46 AM	12:16:47 AM	12:16:52 AM	12:16:59 AM	Hit				12:01:27 AM	Dest	BF	BF	RF	RF		59	24		90
13	4	D	Ar	Far	BF	12:18:48 AM	12:19:47 AM	12:20:11 AM	12:20:18 AM	Hit				12:01:44 AM	Dest	RF	12:18:34 AM	12:19:36 AM	12:19:41 AM	12:19:48 AM	Miss	12:19:53 AM	12:20:00 AM	Hit	12:01:26 AM	Surv	RF	RF	RF	RF		58	24	5	104
14	4	D	Ar	Far	BF	12:21:35 AM	12:22:35 AM	12:22:59 AM	12:23:06 AM	Hit				12:01:32 AM	Dest	RF	12:21:35 AM	12:22:38 AM	12:22:42 AM	12:22:49 AM	Miss	12:22:54 AM	12:23:01 AM	Hit	12:01:26 AM	Dest	RF	BF	RF	RF		60	24	5	92
15	4	D	Ar	Far	BF	12:26:30 AM	12:27:36 AM	12:28:01 AM	12:28:22 AM	Hit				12:01:52 AM	Dest	RF	12:26:53 AM	12:27:53 AM	12:27:58 AM	12:28:19 AM	Hit				12:01:49 AM	Dest	BF	BF	RF	RF		66	26		112
16	4	D	Ar	Far	BF	12:27:07 AM	12:28:09 AM	12:28:33 AM	12:28:48 AM	Hit				12:01:41 AM	Dest	RF	12:27:19 AM	12:28:16 AM	12:28:20 AM	12:28:35 AM	Hit				12:01:28 AM	Dest	BF	BF	RF	RF		62	24		101
17	5	N	FFP	Near	BF	12:29:53 AM	12:30:59 AM	12:31:24 AM	12:31:39 AM	Hit				12:01:46 AM	Surv	RF	12:30:38 AM	12:31:37 AM	12:31:42 AM	12:31:57 AM	Hit				12:02:04 AM	Dest	BF	BF	BF	BF		66	25		106
18	5	N	FFP	Near	BF	12:30:44 AM	12:31:41 AM	12:32:06 AM	12:32:21 AM	Hit				12:01:43 AM	Dest	RF	12:30:38 AM	12:31:32 AM	12:31:36 AM	12:31:51 AM	Hit				12:01:13 AM	Surv	RF	RF	RF	RF		57	25		103
19	5	N	FFP	Near	BF	12:32:58 AM	12:33:59 AM	12:34:25 AM	12:34:32 AM	Hit				12:01:56 AM	Dest	RF	12:32:36 AM	12:33:33 AM	12:33:39 AM	12:33:46 AM	Hit				12:01:09 AM	Surv	RF	RF	RF	RF		62	26		116
20	5	N	FFP	Near	BF	12:33:58 AM	12:34:53 AM	12:35:17 AM	12:35:32 AM	Hit				12:01:34 AM	Dest	RF	12:34:08 AM	12:35:13 AM	12:35:18 AM	12:35:33 AM	Hit				12:01:34 AM	Dest	BF	BF	BF	BF		55	24		94
21	6	N	Ar	Far	BF	12:34:24 AM	12:35:26 AM	12:35:52 AM	12:36:07 AM	Hit				12:01:53 AM	Dest	RF	12:34:14 AM	12:35:13 AM	12:35:18 AM	12:35:33 AM	Hit				12:01:19 AM	Surv	RF	RF	RF	RF		63	26		113
22	6	N	Ar	Far	BF	12:40:23 AM	12:41:31 AM	12:41:57 AM	12:42:18 AM	Miss	12:42:25 AM	12:42:46 AM	Hit	12:02:36 AM	Surv	RF	12:40:10 AM	12:41:09 AM	12:41:13 AM	12:41:34 AM	Miss	12:41:39 AM	12:42:00 AM	Mis	No Hits	Dest	RF	RF	RF	BF		69	25	5	156
23	6	N	Ar	Far	BF	12:42:33 AM	12:43:35 AM	12:43:59 AM	12:44:20 AM	Hit				12:01:51 AM	Dest	RF	12:42:29 AM	12:43:33 AM	12:43:38 AM	12:43:59 AM	Hit				12:01:30 AM	Surv	RF	RF	RF	RF		61	24		111
24	6	N	Ar	Far	BF	12:45:50 AM	12:46:51 AM	12:47:16 AM	12:47:37 AM	Hit				12:01:47 AM	Dest	RF	12:45:51 AM	12:46:53 AM	12:46:58 AM	12:47:19 AM	Hit				12:01:30 AM	Dest	BF	BF	RF	RF		61	25		107
25	7	D	FFP	Far	BF	12:48:16 AM	12:49:11 AM	12:49:36 AM	12:49:43 AM	Miss	12:49:48 AM	12:49:55 AM	Hit	12:01:40 AM	Dest	RF	12:48:20 AM	12:49:24 AM	12:49:29 AM	12:49:36 AM	Hit				12:01:21 AM	Surv	BF	BF	RF	RF		55	25		100
26	7	D	FFP	Far	BF	12:52:27 AM	12:53:33 AM	12:53:58 AM	12:54:13 AM	Hit				12:02:11 AM	Dest	RF	12:52:01 AM	12:52:59 AM	12:53:05 AM	12:53:20 AM	Miss	12:53:26 AM	12:53:41 AM	Hit	12:01:40 AM	Surv	RF	RF	RF	RF		66	25	6	131
27	7	D	FFP	Far	BF	12:53:48 AM	12:54:48 AM	12:55:13 AM	12:55:20 AM	Hit				12:01:40 AM	Dest	RF	12:53:40 AM	12:54:35 AM	12:54:40 AM	12:54:47 AM	Hit				12:01:07 AM	Surv	RF	RF	RF	RF		60	25		100
28	7	D	FFP	Far	BF	12:56:17 AM	12:57:22 AM	12:57:47 AM	12:58:02 AM	Miss	12:58:07 AM	12:58:22 AM	Hit	12:02:05 AM	Dest	RF	12:56:29 AM	12:57:27 AM	12:57:33 AM	12:57:48 AM	Miss	12:57:50 AM	12:58:05 AM	Hit	12:01:48 AM	Surv	BF	BF	RF	RF		65	25	3	125
29	7	D	FFP	Far	BF	12:57:24 AM	12:58:08 AM	12:58:35 AM	12:58:50 AM	Hit				12:01:28 AM	Dest	RF	12:57:23 AM	12:58:28 AM	12:58:32 AM	12:58:47 AM	Hit				12:01:24 AM	Dest	RF	BF	RF	RF		44	27		88
30	8	D	FFP	Near	BF	12:58:55 AM	12:59:57 AM	1:00:21 AM	1:00:28 AM	Hit				12:02:00 AM	Dest	RF	12:58:28 AM	12:59:22 AM	12:59:26 AM	12:59:33 AM	Hit				12:01:05 AM	Surv	RF	RF	RF	RF		62	24		120
31	8	D	FFP	Near	BF	12:59:45 AM	1:00:50 AM	1:01:13 AM	1:01:20 AM	Hit				12:01:34 AM	Dest	RF	1:00:05 AM	1:01:03 AM	1:01:08 AM	1:01:15 AM	Miss	1:01:19 AM	1:01:26 AM	Hit	12:01:40 AM	Dest	BF	BF	RF	BF		64	23	4	94
32	8	D	FFP	Near	BF	1:01:35 AM	1:02:35 AM	1:03:01 AM	1:03:08 AM	Hit				12:01:33 AM	Dest	RF	1:01:54 AM	1:02:53 AM	1:02:58 AM	1:03:05 AM	Hit				12:01:31 AM	Dest	BF	BF	RF	RF		61	25		93
33	9	N	FFP	Far	BF	1:02:30 AM	1:03:32 AM	1:03:56 AM	1:04:03 AM	Hit				12:01:33 AM	Dest	RF	1:02:31 AM	1:03:25 AM	1:03:30 AM	1:03:37 AM	Miss	1:03:42 AM	1:03:49 AM	Hit	12:01:20 AM	Surv	BF	RF	RF	RF		62	24	5	93
34	9	N	FFP	Far	BF	1:03:03 AM	1:03:59 AM	1:04:25 AM	1:04:46 AM	Hit				12:01:53 AM	Dest	RF	1:02:53 AM	1:03:59 AM	1:04:04 AM	1:04:25 AM	Hit				12:01:32 AM	Dest	RF	RF	RF	RF		57	25		113
35	9	N	FFP	Far	BF	1:05:19 AM	1:06:18 AM	1:06:42 AM	1:06:49 AM	Hit				12:01:42 AM	Dest	RF	1:05:07 AM	1:06:12 AM	1:06:17 AM	1:06:24 AM	Hit				12:01:17 AM	Surv	RF	RF	RF	RF		60	24		102
36	9	N	FFP	Far	BF	1:05:36 AM	1:06:34 AM	1:07:00 AM	1:07:07 AM	Hit				12:01:30 AM	Dest	RF	1:05:38 AM	1:06:29 AM	1:06:35 AM	1:06:42 AM	Hit				12:01:05 AM	Surv	BF	RF	RF	RF		58	26		90
37	10	N	Ar	Near	BF	1:07:39 AM	1:08:42 AM	1:09:07 AM	1:09:28 AM	Hit				12:01:50 AM	Dest	RF	1:07:38 AM	1:08:42 AM	1:08:46 AM	1:09:07 AM	Miss	1:09:12 AM	1:09:33 AM	Hit	12:01:55 AM	Dest	RF	RF	RF	BF		63	25	5	110
38	10	N	Ar	Near	BF	1:07:54 AM	1:08:53 AM	1:09:19 AM	1:09:26 AM	Hit				12:01:45 AM	Dest	RF	1:07:41 AM	1:08:43 AM	1:08:48 AM	1:08:55 AM	Hit				12:01:14 AM	Surv	RF	RF	RF	RF		60	26		105
39	10	N	Ar	Near	BF	1:11:27 AM	1:12:24 AM	1:12:50 AM	1:13:11 AM	Miss	1:13:16 AM	1:13:37 AM	Hit	12:02:20 AM	Dest	RF	1:11:17 AM	1:12:19 AM	1:12:25 AM	1:12:46 AM	Hit				12:01:29 AM	Surv	RF	RF	RF	RF		57	26		140
Data truncated here																																			
180	48	D	Ar	Mid	BF	5:41:33 AM	5:42:29 AM	5:42:53 AM	5:43:08 AM	Hit				12:01:42 AM	Surv	RF	5:41:26 AM	5:42:29 AM	5:42:35 AM	5:42:50 AM	Miss	5:42:57 AM	5:43:12 AM	Mis	No Hits	Dest	RF	BF	RF	BF		56	25	7	102
181	48	D	Ar	Mid	BF	5:46:14 AM	5:47:04 AM	5:47:30 AM	5:47:37 AM	Hit				12:01:25 AM	Dest	RF	5:46:11 AM	5:47:14 AM	5:47:19 AM	5:47:26 AM	Hit				12:01:14 AM	Surv	RF	BF	RF	RF		50	26		85
182	48	D	Ar	Mid	BF	5:46:46 AM	5:47:41 AM	5:48:06 AM	5:48:13 AM	Hit				12:01:32 AM	Dest	RF	5:46:41 AM	5:47:38 AM	5:47:44 AM	5:47:51 AM	Miss	5:47:54 AM	5:48:01 AM	Hit	12:01:20 AM	Surv	RF	RF	RF	RF		55	25	3	92
183	48	D	Ar	Mid	BF	5:46:49 AM	5:47:47 AM	5:48:13 AM	5:48:20 AM	Miss	5:48:26 AM	5:48:33 AM	Hit	12:01:43 AM	Dest	RF	5:47:22 AM	5:48:24 AM	5:48:30 AM	5:48:37 AM	Hit				12:01:47 AM	Dest	BF	BF	BF	BF		58	26		103

D=Day • N=Night • Arm=Armor • BF=Blue Force • RF=Red Force • Dtct=Detects • Dest=Destroyed • Surv=Survives • Sht=Shoots • Wns=Wins

No shading = Raw Data • Shaded = Reduced Data

Fig. 3-5-1.
Sample MOP-1.1 Raw & Reduced Data

Fig. 3-5-2.
MOE -1 Sample Data

Trial	Mission Orders Issued (time)	Mission Start (time)	Mission Complete (time)	Illumination	Target Type	Distance	BF PreMission Force Size	BF Readiness Losses	BF Dependability Losses	Init Engagemt BF Size	End Engagemt BF Size	Blue TEC Time (sec)	1 / TEC_{Blue}	RF PreMission Force Size	RF Readiness Losses	RF Dependability Losses	Init Engagemt RF Size	End Engagement RF Size	Red TEC Time (sec)	1 / TEC_{Red}	Force Superiority (Φ_0)	Engagemt Outcome
1	8/1/12 6:29	8/1/12 8:29	8/2/12 22:29	N	Arm	Mid	6	0	0	6	2	117	0.00853	6	1	0	5	2	91	0.01103	1.05	0
2	8/3/12 14:45	8/3/12 16:45	8/5/12 6:45	N	FFP	Near	6	0	2	4	0	110	0.00909	6	2	0	4	3	86	0.01167	0.88	-3
3	8/6/12 6:03	8/6/12 8:03	8/7/12 22:03	D	FFP	Mid	6	0	2	4	0	107	0.00933	6	2	0	4	2	93	0.01070	0.93	-2
4	8/8/12 18:10	8/8/12 20:10	8/10/12 10:10	D	Arm	Far	6	0	1	5	1	102	0.00979	6	1	1	4	1	92	0.01083	1.19	0
5	8/11/12 0:19	8/11/12 2:19	8/12/12 16:19	N	FFP	Near	6	1	0	5	2	105	0.00955	6	2	0	4	2	90	0.01108	1.16	0
6	8/13/12 17:53	8/13/12 19:53	8/15/12 9:53	N	Arm	Far	6	1	1	4	1	122	0.00820	6	0	1	5	3	86	0.01160	0.67	-2
7	8/16/12 0:50	8/16/12 2:50	8/17/12 16:50	D	FFP	Far	6	1	0	5	0	109	0.00919	6	1	0	5	4	88	0.01137	0.90	-4
8	8/19/12 1:50	8/19/12 3:50	8/20/12 17:50	D	FFP	Near	6	0	3	3	0	103	0.00975	6	3	0	3	1	85	0.01174	0.91	-1
9	8/21/12 9:33	8/21/12 11:33	8/23/12 1:33	N	FFP	Far	6	1	0	5	1	100	0.01003	6	0	1	5	4	78	0.01276	0.89	-3
10	8/24/12 10:39	8/24/12 12:39	8/26/12 2:39	N	Arm	Near	6	0	2	4	0	113	0.00888	6	2	0	4	2	94	0.01067	0.91	-2
11	8/27/12 7:55	8/27/12 9:55	8/28/12 23:55	D	FFP	Mid	6	0	0	6	2	106	0.00945	6	1	0	5	3	85	0.01182	1.07	-1
12	8/29/12 14:38	8/29/12 16:38	8/31/12 6:38	D	FFP	Mid	6	2	0	4	2	106	0.00945	6	2	0	4	1	105	0.00949	1.00	1
13	8/31/12 19:12	8/31/12 21:12	9/2/12 11:12	N	Arm	Near	6	2	1	3	0	118	0.00846	6	2	0	4	4	77	0.01301	0.60	-4
14	9/3/12 10:34	9/3/12 12:34	9/5/12 2:34	N	FFP	Mid	6	1	0	5	3	105	0.00951	6	1	1	4	3	89	0.01122	1.15	0
15	9/5/12 18:20	9/5/12 20:20	9/7/12 10:20	N	Arm	Far	6	1	0	5	2	115	0.00873	6	1	0	5	3	101	0.00995	0.94	-1
16	9/8/12 10:05	9/8/12 12:05	9/10/12 2:05	N	FFP	Mid	6	0	0	6	2	121	0.00826	6	1	0	5	5	93	0.01081	1.05	-3
17	9/10/12 19:08	9/10/12 21:08	9/12/12 11:08	D	Arm	Mid	6	1	1	4	0	125	0.00800	6	0	1	5	4	90	0.01111	0.68	-4
18	9/13/12 7:51	9/13/12 9:51	9/14/12 23:51	D	FFP	Far	6	1	0	5	1	116	0.00862	6	0	2	4	2	91	0.01104	1.10	-1
19	9/15/12 15:04	9/15/12 17:04	9/17/12 7:04	N	Arm	Near	6	2	0	4	1	117	0.00854	6	2	0	4	3	89	0.01125	0.87	-2
20	9/18/12 13:32	9/18/12 15:32	9/20/12 5:32	N	Arm	Mid	6	2	0	4	0	119	0.00842	6	1	0	5	4	94	0.01063	0.71	-4
45	11/22/12 3:53	11/22/12 5:53	11/23/12 19:53	N	FFP	Far	6	1	1	4	2	122	0.00821	6	2	0	4	1	120	0.00835	0.99	1
46	11/24/12 18:23	11/24/12 20:23	11/26/12 10:23	N	FFP	Mid	6	1	1	4	2	108	0.00930	6	1	0	5	2	96	0.01045	0.75	0
47	11/27/12 6:47	11/27/12 8:47	11/28/12 22:47	D	Arm	Near	6	2	0	4	1	117	0.00855	6	1	0	5	3	94	0.01064	0.72	-2
48	11/30/12 9:51	11/30/12 11:51	12/2/12 1:51	D	Arm	Mid	6	1	1	4	1	96	0.01045	6	0	2	4	2	87	0.01146	0.95	-1

D=Day • N=Night • Arm=Armor • BF=Blue Force • RF=Red Force

This page intentionally left blank for two-sided printing.

STEP 6 System Evaluation and Reporting

	Tasks	Responsible	Provides Input	Outputs	Signature/ Release Authority
Prepare Assessment/Evaluation Report*					
	Write Report	ORSA	Military SME, OTPO, TM, MS	Draft Assessment/ Evaluation Report	
	Review/Comment	OTPO	Military SME, ORSA, MS, OTAD, DH, Exec	CRM	
	Adjudicate CRM			CRB signature-ready Assessment/Evaluation Report	
	Prepare post-CRB Copy			Signature-ready Assessment/ Evaluation Report	
	Approve			Signed Report	Exec
Prepare M&S Accreditation, if performing					
	Write Accreditation Report	ORSA	Military SME, OTPO, MS	Draft Accreditation Report	OTAD
	Review/Comment on Report	OTPO	Military SME, ORSA, MS, OTAD, DH, Exec	CRM	
	Adjudicate CRM			CRB signature-ready Accreditation Report	
	Approve			Signed Accreditation Report	Exec
Accreditation Decision Letter		OTPO		Follow same process as Accreditation Report	Exec
Write and Upload Lessons Learned		OTPO	Military SME, ORSA, MS, DM, DH, OTAD	Completed Lessons Learned	
Records Management		OTPO			S-1

Chapter 3-6. Step 6: System Evaluation and Reporting

Purpose of Evaluation

The purpose of system evaluation is to answer the evaluation questions (i.e., Issues, COIs, and OE/OS/OSur) contained in the SEP, thereby providing information to decision makers and PMs useful to system design and tradeoff decisions. The necessary input for system evaluation is one or more Test Data Reports, which should naturally flow from test events specified in the TEMP. The ORSA (evaluator) is charged with leading the evaluation.

Evaluation should begin at the lowest levels of indenture (generally the Subtask level) at the early stages of system development. Little benefit exists in delaying evaluation and reporting results until late in the program. As the system matures, the evaluations should progress to higher levels of indenture until reaching the top level of the hierarchy, answering COIs and determining OE/OS/OSur.

Evaluation and Reporting Requirements for OT

After all developmental, live fire, and operational testing is complete, the ORSA leads the evaluation effort by using all available test data and test reports to complete the system evaluation. The MCOTEA test team plays a key role in the evaluation by providing contextual information, explaining any unusual behavior in the data, and providing any other background information pertaining to data taken during any Intermediate Assessments, Operational Assessments, and IOT. The goal of the process at this stage is to help the ORSA understand the conditions under which individual tests were conducted and data was gathered.

The evaluation is designed to accomplish the following:

- Determine if thresholds in the approved capabilities documentation and COIs have been satisfied
- Determine OE, OS, and OSur under realistic operational conditions, including Joint combat
- Assess the impact to combat operations
- Provide additional information on the system's operational capabilities

As part of the system evaluation, the ORSA must include a comparison with current mission capabilities using existing data to help determine measurable improvements brought about by the new system. The cognizant Division will supply data on current mission capabilities. If this isn't possible, the ORSA will consult with the PM, who will propose an alternative strategy for obtaining this information (DODI 5000.02 2008).

Determine Threshold Satisfaction

The ORSA analyzes data from all contractor DT, government DT, LFT&E, modeling and simulation, and MCOTEA's observations, assessments, and operational testing to determine which thresholds have been met and which have not. The OER will address both instances.

Determine Operational Effectiveness

OE is directly related to mission effectiveness, which is represented by MOEs. MOEs are typically associated with specific areas of operational interest, each of which contributes to the system's overall capability to accomplish its mission. OE can only be determined as a result of operational testing.

Determine Operational Suitability

The evaluator determines OS by examining data results from Measures of Suitability throughout program testing and evaluation. Areas of suitability include but are not limited to RAM, logistics supportability, compatibility, interoperability, training, human factors, safety, manpower and personnel selection, transportability, environmental effects, and system

documentation. Data from many of these areas can be accumulated from early program phases, and when evaluated with OT data, helps determine OS.

Determine Operational Survivability

The evaluator uses results from any LFT&E, IA, and CBRN events complemented by data from a modeling and simulation environment in conjunction with DT and OT data to determine OSur. During OT, the system's capability, or lack thereof, is demonstrated using representative tactics and countermeasures for both friendly and opposing forces. The focus of the OSur evaluation is on the capability of the system and the crew to avoid damage, withstand attack, and recover capability in a hostile combat environment without adversely affecting mission accomplishment.

For information systems, the OSur evaluation examines information and data security.

Assess Impact to Combat Operations

This part of the evaluation examines how the system under test contributes to the overall ability of the Marine Corps to conduct combat operations. This assessment, conducted by the test team, may be qualitative or quantitative, and the impact may be small or large. All assessments are supported by data.

Major System Deficiencies

Major System Deficiencies are directly related to the system under test and are generally the failure of the system to attain a required system capability or attain a required threshold value as stated in the capabilities documentation. These deficiencies are identified during IOT, although the potential for a Major System Deficiency may be identified during Integrated Testing. MCOTEA notifies the PM of any potential system deficiencies identified during Integrated Testing.

Operational Deficiencies

During integrated testing and IOT, test personnel may identify issues that affect the performance of the system under test, even though these issues cannot be associated with a specific capability of the system under test. Indeed, operational testing may be the first opportunity to discover these issues.

Operational Deficiencies tend to pertain to interfaces with other systems or to system interactions with the Operating Forces. In some cases, these deficiencies may actually be materiel gaps in operational capability, and in other cases, they may illuminate the need to create or modify TTPs. The test team reports all Operational Deficiencies identified during any phase of testing.

Although Operational Deficiencies are not used in determining OE, OS, and OSur, if an Operational Deficiency is severe enough, MCOTEA may recommend that the system under test not be fielded until the deficiency is addressed.

Types of Evaluation Reports

Evaluations that coincide with major operational test events are termed either OTA Assessment Reports (for EOAs and OAs) or OTA Evaluation Reports (for IOT) or OTA Follow-on Evaluation Reports (for FOTs). All other evaluation reports published before or between major operational test events are termed Intermediate Assessment Reports. Evaluation reports for System Assessments are termed SARs.

MCOTEA sends all evaluation reports to the system's PM and MDA. Major evaluations such as OARs and OERs will often reference IARs as supporting information. No limits exist on the number of evaluation reports that may occur as the system progresses through its development cycle.

Evaluation Process Basics

The evaluation process is relatively straightforward because the standards needed for evaluation were developed in the SEP. The process, regardless of the testing source (developmental or operational), fundamentally compares test results with established standards.

Populations and Samples

Before beginning the comparison it is important to understand the difference between the population and the sample. Populations, such as the total number of helmets in the Marine Corps, are often extremely large and represent the entire universe of objects to be evaluated. A population's size usually makes it impossible, for reasons of cost and practicality, to measure every element of the population. The solution is to draw samples from the population and to generalize from the sample (inference) to the broader population.

Parameters and Statistics

Coinciding with the concepts of populations and samples are the concepts of parameters and statistics. A parameter is any characteristic of the population, while a statistic is a characteristic of the sample (Winer 1971). The parameters evaluated for a system under test are compared with standards derived, in part, from capability documents. Put another way, the standards are what is desired of the population of systems once fielded.

Statistics are used in the evaluation to estimate the value of the parameter. When testing is conducted, data, using samples, is collected, and from that sample statistics are calculated. The statistic is then compared with the standard to determine if expectations for the population are met.

Considering Uncertainty with Confidence Bounds

Comparing results with standards should always account for uncertainty to ensure that the decision maker is accurately informed. The evaluator is usually interested in using the sample collected during testing as an estimate of the true value in the population. An approach used to assess for accuracy of the sample is to calculate boundaries within which the true value is likely to fall. Such boundaries are called confidence bounds (Fields 2005). When comparing the sample result from the test reports with the standard, the evaluator should compare the prespecified confidence bound. The confidence bound takes into account the statistical error and random chance inherent in the testing to express to the decision maker how certain the evaluator is about the answer.

Evaluation at Early Stages of System Development

Early test results generally derive from developmental testing designed to satisfy one or more of the Issues at the lowest levels of indenture in the Evaluation Framework, typically at the Subtask level and below. The process of evaluating these early results begins with receipt of data in a Developmental Test report.

At this early stage of evaluation, aggregation at the lower levels (Subtasks and Tasks) up to mission accomplishment is not necessary. However, if shortfalls are identified in evaluation results it is appropriate to identify the potential ramifications to the next level up in the hierarchy. In the example, the effect of the shortfall is undetermined at this time. The evaluation result, however, has value despite the fact that the nature and direction of the cause/effect relationship has not been firmly established. The evaluation results at this stage identify an area of risk that could ultimately have a negative effect on mission accomplishment. Figure 3-6-1 illustrates the linkage, by Task, of the Firing First Round on Target requirement to the outcome of the Offensive Mission.

Bias in Parameters

Operational test is a valuable opportunity for obtaining realistic estimates for parameters, but not necessarily all parameters.

Operational Availability (A_o) is a parameter that usually carries significant bias in the operational test. A_o is a function of the intervals required for maintenance and the time to maintain the system. Time to maintain is biased because maintainers and parts are co-located with the test unit.

Co-location from a test perspective is sound because it enables greater system availability during OT. However, co-location also creates the detrimental effect of inflating A_o. In reality, parts and maintainers are often separated by time and distance, which delays repairs and reduces the A_o.

The example below is also a proxy measurement being used to satisfy a Subtask. This proxy Measure, single shot probability of hit, falls squarely in the category of developmental testing. This example also illustrates integrated developmental and operational testing. In the example, developmental testing has been integrated in the independent evaluation of the ABC system by satisfying the information needs for Subtask accomplishment and threshold satisfaction.

If the shortfall in performance, such as in the probability of hit example, is accepted by the MDA without employing a fix or mitigation strategy, then the evaluation of this capability is concluded and the answer stands "as is." Should fixes be required of the system, then retesting is required to ensure that the fix was successful and that the modification has not affected other functions. In software testing this is called regression testing. Retesting and regression testing may require updates to the TEMP, depending on the program's need to modify schedule, test events, and resource changes.

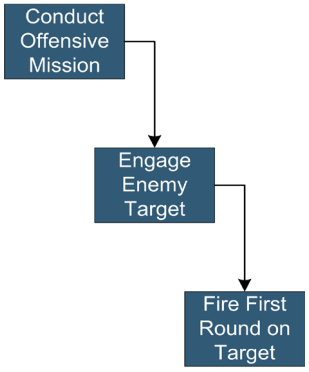


Figure 3- 6-1. Example of Linking Subtasks to Tasks

Evaluation at Later Stages of System Development

At later stages of system development, evaluation addresses Issues at the Task level or suitability Issues at the major component or system level. (See the example of Evaluating at Later Stages).

Example of Evaluating at Early Stages

This example illustrates test data, test statistics, and subsequent evaluation.

Subtask

I-X.X Is the single shot probability of a hit greater than 0.70?

Measure

Single Shot Probability of Hit

Threshold: > 0.70

Null Hypothesis: $\leq p$ 0.70

Alternative Hypothesis: p > 0.70

Confidence Level: α= 0.20

Test Data and Test Statistics

The developmental test results in the tables to the right were documented in the test report of the ABC system.

	Distance		
Trial	1250	2500	3750
1	Hit	Hit	Hit
2	Hit	Hit	Miss
3	Miss	Hit	Miss
4	Miss	Miss	Miss
5	Hit	Hit	Hit
6	Miss	Hit	Hit
7	Hit	Hit	Miss
8	Miss	Hit	Miss
9	Hit	Hit	Miss
10	Hit	Hit	Miss
11	Hit	Hit	Hit
12	Hit	Hit	Miss
13	Miss	Hit	Hit
14	Hit	Hit	Miss
15	Hit	Hit	Miss
16	Miss	Hit	Miss
17	Hit	Miss	Miss
18	Hit	Hit	Hit
19	Hit	Miss	Miss
20	Hit	Hit	Hit
21	Hit	Miss	Hit
22	Miss	Hit	Miss
23	Hit	Hit	Hit
24	Hit	Miss	Hit
25	Hit	Miss	Hit
26	Hit	Miss	Miss
27	Miss	Hit	Miss
28	Hit	Hit	Miss
29	Hit	Hit	Hit
30	Hit	Hit	Hit

Evaluation of Test Results in the IAR

I-X.X Is the single shot probability of hit greater than or equal to 0.70? Answer: No.

	Distance			
	1250	2500	3750	Overall
Hits (x)	22	23	13	58
Attempts (N)	30	30	30	90
Point Est.	0.73	0.77	0.43	0.64
80% LCB ($\hat{p}$)	0.64	0.68	0.34	0.60

According to the rationale for the requirement in the ABC Capability Production Document, "The ABC's single shot probability of hit should be high enough to enable a first round defeat of an armored target out to maximum effective range in a combat environment." Based on the sample data from the developmental testing, MCOTEA is 80 percent confident that the true population proportion for probability of hit is 0.60 or greater. Because the lower confidence bound is below the threshold value of 0.70, MCOTEA does not have sufficient certainty to conclude that the requirement has been satisfied at this time. The potential exists that this requirement's shortfall to have a negative effect on the target engagement task performed by the crew.

Example of Evaluating at Later Stages

This example illustrates the importance of threshold conditions in conjunction with performance.

Task

I-X. Can the Light Armored Reconnaissance crew's destroy enemy targets in less than or equal to 60 seconds after identification?

Measure

Target Engagement Cycle

Threshold: ≤ 60 seconds

Null Hypothesis: $\bar{x} \geq 60$ seconds

Alternative Hypothesis: $\bar{x} < 60$ seconds

Confidence Level: $\alpha = 0.20$

Statistics	
sample mean	69
standard deviation	15.0
Upper Confidence Bound	70.6
alpha	0.2

Test Data

The operational test results in the tables are from a fictional Operational Assessment of the ABC system.

Test Data		
Trial	Tank	BMP
1	50	62
2	67	46
3	58	69
4	56	54
5	68	68
6	106	46
7	51	67
8	56	74
9	69	60
10	33	74
11	84	57
12	76	82
13	58	91
14	69	83
15	72	86
16	59	75
17	60	69
18	90	73
19	71	100
20	64	86
21	93	71
22	66	61
23	54	72
24	86	77
25	54	78
26	58	81
27	58	108
28	68	54
29	46	71
30	65	65

Test Statistics

Evaluation of Test Results

I-X. Can the ABC crews destroy enemy targets in less than or equal to 60 seconds after identification?
Answer: No.

> According to the rationale for the requirement in the ABC Capabilities Production Document, "the ABC shall have the ability to successfully destroy armored fighting vehicles in less than 60 seconds. The threshold value of 60 seconds represents an operationally significant decrease in the kill-chain time from target detection to target kill." Because the upper confidence bound is greater than the threshold value, MCOTEA cannot conclude with reasonable certainty that the system is capable of meeting it's target engagement threshold value at this time. An analysis of the sample data indicates that performance is not uniform across all conditions. Performance against tanks was better than against BMPs, although neither condition met the threshold expectations.

ANOVA: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Tank	30	1,966	66	229
BMP	30	2,160	72	205

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	622.8528	1	622.8528	2.871625	0.095517	1.680443
Within Groups	12,580.15	58	216.8991			
Total	13,203	59				

As a system progresses in its development, MCOTEA may perform an Operational Assessment as a pre-IOT event.

Evaluating COIs

After evaluating screening criteria, the ORSA compares the observed value for the measured effects to the expected value in the System Evaluation Plan. The ORSA also inserts test data into the analytic model developed specifically for the COIs. The principle task here is to conduct sensitivity analysis. The ORSA uses the analytic model from the SEP to assess the relative contribution of system performance and suitability towards achievement of the observed effect. The following example of an IOT performed on the ABC System assumes that all earlier stages of evaluations have been successfully completed and the screening criteria have been satisfied.

The ABC system provides the using unit with mobile anti-armor capabilities during offensive and defensive MAGTF operations. In this example the ABC System is replacing a legacy system having the same mission and achieving the same effect, albeit not as well as the ABC System's requirements. Given this information, the ABC System has two COIs and MOEs.

Example of Analysis of COI Results

The following discussion illustrates only a small part of the analysis that should occur to explain evaluation results. From here the evaluator continues to thoroughly explore the data results and prepares to inform the decision maker of the evaluation's conclusions.

COI-1. Can a Light Armored Reconnaissance (LAR) platoon equipped with anti-tank variants destroy a like-size armored enemy force during offensive operations? To answer this question the LAR platoon had to demonstrate that it could achieve a force superiority value greater than or equal to 1.55 to be found fully mission capable in offensive operations. Force superiority is a measure of combat power brought to bear against an enemy force. A force superiority value above one means the force is superior while a value below one means the force is inferior.

The LAR platoon was able to achieve an average force superiority value of 0.94 as noted in the equation in figure 3-6-2, which is below the threshold value meaning the platoon was not a superior force. Low force superiority explains why the LAR platoon was only able to win 6 and tie 7 of the 48 total mission engagements.

Some of the shortfall in force superiority can be attributed to the enemy force. The enemy force used during testing was more robust than what was used to set the threshold value which explains why they won 35 of the 48 engagements. When the differences in the expected performance of the enemy force versus the actual performance is accounted for and removed, the force superiority for the LAR platoon rises to 1.34 using the equation in figure 3-6-3. However, this value still falls short of the threshold.

Fig. 3-6-2. Average Force Superiority Value

$$Force\ Superiority_{unadjusted} = \frac{m_{blue}}{m_{red}}\sqrt{\frac{a_{blue}}{a_{red}}} = \frac{4.3}{4.3}\sqrt{\frac{0.009}{0.011}} = 0.94$$

Fig. 3-6-3. Difference Between Expected and Actual Performance Removed

$$Force\ Superiority_{adjusted} = \frac{m_{blue}}{m_{red}}\sqrt{\frac{a_{blue}}{a_{red}}} = \frac{4.3}{4}\sqrt{\frac{0.009}{0.006}} = 1.34$$

Where m_{blue}= initial blue force size at outset of engagement
m_{red}= initial red force size at outset of engagement
a_{blue}= effective firing rateBlue=1/target engagement cycleBlue
a_{red}= effective firing rateRed= 1/target engagement cycleRed

MCL on the y-axis of the Graph

The translation of force superiority into mission capability level is illustrated in figure 3-6-5. The red triangle and yellow circle identified in the figure marks the unadjusted and adjusted force superiority values. The adjusted value indicates the LAR platoon was partially mission capable when using the anti-tank system.

Calculating MCLs for COIs

Using the COI results from the analytic models, the evaluator calculates the MCL for each COI using the decision model from the SEP. The output of this process is a value on the MCL scale between 0 and 100 for each COI.

To find MCL, in this example the last two rows of figure 3-6-4 are depicted in a piecewise linear function (fig. 3-6-5). The table contains results from the analytic model, depicted by Pmission on the x-axis and MCL on the y-axis of graph. The process must be repeated for each COI if multiple COIs are contained in the SEP.

Each of the 48 missions conducted by the LAR platoon included a single force-on-force engagement. The 48 missions were run under a variety of conditions during testing to include both day and night; against both armored and fixed fighting position targets; and at near, middle, and far distances. Figure 3-6-6 contains subsample averages of the unadjusted force superiority results under the various mission conditions. The results in the table indicate small differences in force superiority, none of which are considered statistically significant.

With typical factors ruled out as a cause for the lower than expected observed results the next step is to examine the constituent components of force superiority, namely effective firing rate and force size.

Effective Firing Rate. Effective firing rate is the reciprocal of the time to complete the target engagement cycle (i.e., 1/TEC), or more simply the time to kill a target. The faster a unit can complete the target engagement cycle the more lethal it is considered. Figure 3-6-7 (next page) illustrates the ordered process steps to

Parameter	Description	Lowest Possible	Current Capability	Threshold Capability	Objective Capability	Test Results
$R(t_m)$	mission reliability	0.40	0.55	***0.75***	0.90	**0.74**
k	probability of failure detection	0.90	0.90	***0.90***	0.90	**0.89**
$M(t_r)$	probability of repair given a specified turnaround time	0.50	0.65	***0.85***	0.95	**0.39**
M_o	operational maintainability	0.25	0.50	***0.85***	0.90	**0.50**
t_a	time to acquire the target	20	20	***20***	20	**60**
t_1	time to fire the first round	180	120	60	20	**25**
t_f	time of flight of the projectile from weapon to target	22	22	22	22	**22**
$P(H)$	single shot probability of hit	0.70	0.70	0.70	0.90	**0.78**
$P(K/H)$	probability of kill given a hit	1.00	1.00	1.00	1.00	**1.00**
t_m	time to fire a round given the preceding round was a miss	180	5	5	5	**5**
t_h	time to fire a round given the preceding round was a hit					
$P(H_2/H_1)$	probability of second shot hit given preceding round hit					
$P(H_2/M_1)$	probability of second shot hit given preceding round missed	0.90	0.90	0.90	0.90	**0.95**
$E[TEC]$	expected time for the target engagement cycle	289	171	111	65	**110**
$POR(ta)$	operational readiness given a specified turnaround time	0.67	0.81	***0.94***	0.99	**0.83**
D	dependability	0.55	0.78	***0.96***	0.99	**0.87**
m_{blue}	expected blue force size	2	4	***5***	6	**4.3**
a_{blue}	blue force effective firing rate	0.003	0.006	0.009	0.015	**0.009**
$-a_{blue}m_{blue}$	attrition rate of red forces	-0.007	-0.023	-0.045	-0.092	**-0.039**
Φ (x-axis)	Force Superiority	0.38	1.00	1.55	2.43	**1.34**
MCL (y-axis)	Mission Capability Level	0	50	80	100	**69**

Figure 3-6-4. Results of the Analytic Model in Tabular Form

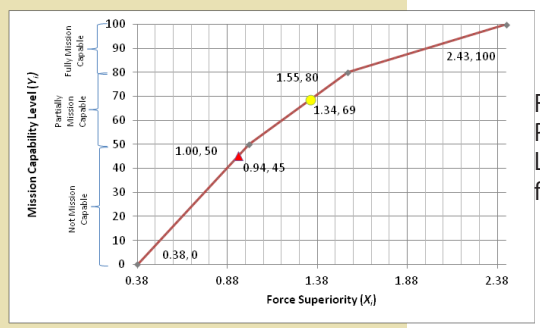


Figure 3-6-5. Piecewise Linear Function for MCL

Subsample Force Superiority Averages				
	Illumination			
	Day		Night	
Target Distance	Armor	FFP	Armor	FFP
Near	0.885	0.871	0.946	0.946
Mid	0.985	1.058	0.885	0.978
Far	0.964	0.983	0.982	0.941

Figure 3-6-6. Subsample Averages of Unadjusted Force Superiority Results

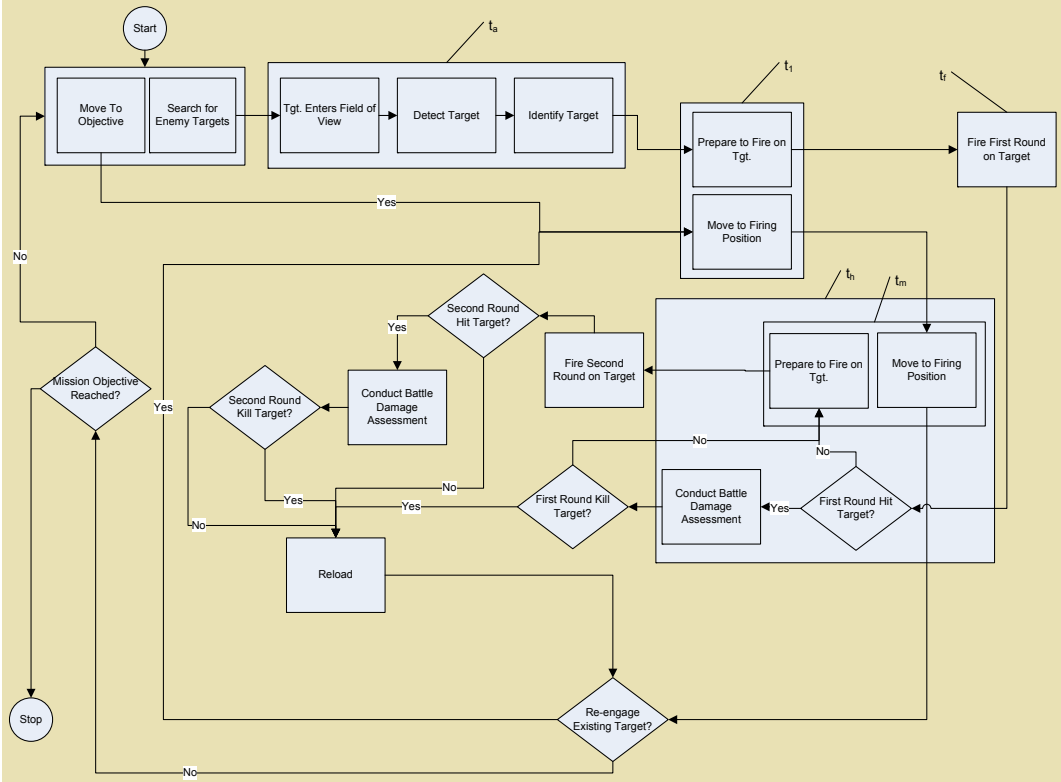


Figure 3-6-7. Ordered Process Steps to Complete the Target Engagement Cycle

Simplifying assumptions were made during testing that make certain steps of the target engagement cycle process not applicable in the evaluation. During force-on-force testing, live rounds were not used; therefore, a hit was considered a kill. The kill diamonds in the figure were affected by the assumption. Any time a hit was recorded, the diamonds became a single path vice a decision node (i.e., hit equals a kill = yes). The analytical equation from the evaluation plan was modified to reflect this assumption and shown in figure 3-6-8. The simplifying assumption also resulted in a modification to the expected value. The updated expected value for the target engagement cycle and the tested results are shown in figure 3-6-9.

Figure 3-6-8. Analytical Equation Modified to Reflect Assumption

$$E[TEC] = t_{acq} + t_1 + t_f + \frac{t_m + t_f}{1-(1-P_h)^2}(1-P_h)$$

Where t_{acq} = time to acquire the target
t_1 = time to fire the first round
t_f = time of flight of the projectile from weapon to target
t_m = time to fire a round given the preceding round was a miss
P_h = probability of hit

Figure 3-6-9. Updated Expected Value for the Target Engagement Cycle and the Tested Results

Parameter	Description	Threshold	Test Results
t_{acq}	time to acquire the target	20	**60**
t_1	time to fire the first round	60	**25**
t_f	time of flight of the projectile from weapon to target	22	**22**
$P(_h)$	single shot probability of hit	0.70	**0.78**
t_m	time to fire a round given the preceding round was a miss	5	**5**
$P(_{h1}/_{m1})$	probability of second shot hit given preceding round missed	0.90	**0.95**
E[TEC]	expected time for the target engagement cycle	111	**110**

As seen in figure 3-6-9 the target engagement cycle for the LAR platoon overall was not a causal factor in poor mission results. It took on average about 5 seconds more to acquire and fire the first round than was expected. However, the first and second round probabilities of hit offset the time increase, resulting in a lower average target engagement cycle time.

Force Size. The second component of force superiority is force size, which is the numerical count of forces at the outset of the engagement, in this case the number of vehicles. This area of testing is where the LAR platoon equipped with the anti-tank system fell short of expectations. The expected force size for the engagements was an average of 5 vehicles, but an average of 4.3 vehicles was demonstrated. With a diminished force size the LAR platoon could not routinely overtake the enemy during mission engagements. For a vehicle to be counted at the outset of an engagement it must be both operationally ready (ready to start the mission) and dependable (failure-free operation during missions).

Operational Readiness. At the beginning of a mission fewer vehicles were operationally ready than expected. The demonstrated operational readiness was 0.83, which is lower than the expected 0.94. The process of determining if a vehicle was operationally ready is illustrated in figure 3-6-10.

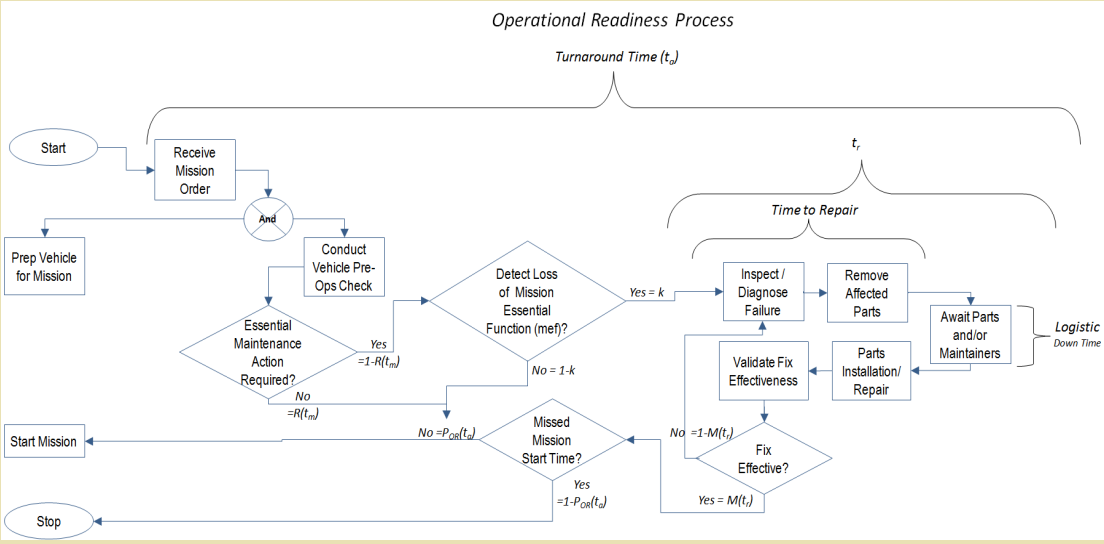


Figure 3-6-10. Process for Determining Operational Readiness

The low probability of affecting repairs to a system in the turnaround time between missions can be traced to a subcomponent of operational readiness. The mean Down Time noted by the horizontal brace labeled "t_r" was the root cause for the reduced number of operationally ready vehicles.

Between each mission were 12 hours that could be used for repairing vehicles that had failed during the preceding mission or were discovered to have a failure during pre-mission vehicle inspections. The expected likelihood that a vehicle could be turned around during this 12-hour time was 0.85; however, the tested likelihood was only 0.39 due to the 24.4-hour mean Down Time observed. For the turnaround time to have been met, the mean Down Time would need to have been equal to or less than 6.3 hours.

Dependability. Anti-tank vehicle dependability is the second reason LAR platoons had fewer than expected vehicles at the outset of engagements. The demonstrated dependability was 0.87, which is below the expected value of 0.96. Dependability is the combination of mission reliability (likelihood of failure-free operation) and operational maintainability (probability that when a failure occurs, it will be repaired in a time not exceeding the allowable Down Time). The dependability process is depicted in figure 3-6-11.

Mission reliability was close to meeting expectations at 0.74 (threshold ≥ 0.75) for a 38.5-hour mission duration and therefore was not the dominant value contributing to low dependability.

In figure 3-6-11, the subprocess representing operational maintainability noted by the horizontal brace labeled "t_r" is the culprit for low dependability. No more than 30 minutes was allowed for a repair when loss of a mef occurred during a mission. On average, 42 minutes were needed to restore a mef during missions, resulting in an operational readiness estimate of 0.87 compared to the threshold of 0.96. To have satisfied the requirement for operational maintainability, the average Down Time would need to have been equal to or less than 18 minutes.

Figure 3-6-12 summarizes the thresholds and results for both operational readiness and dependability parameters.

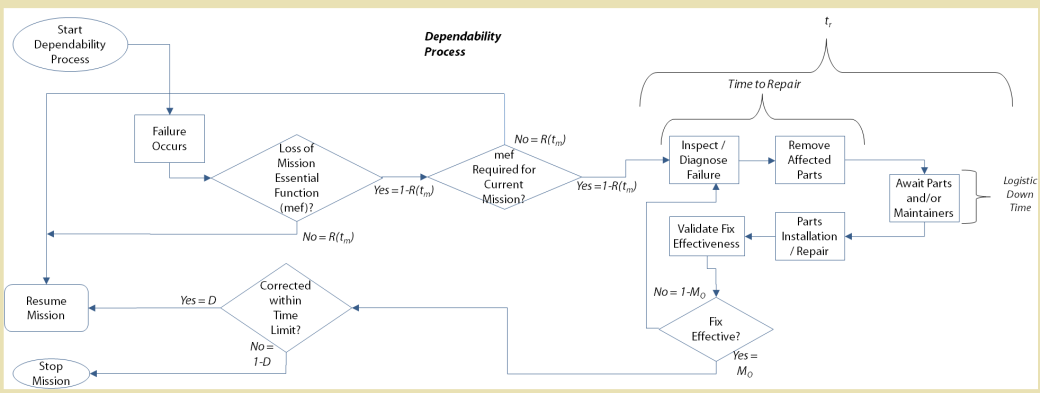


Figure 3-6-11. Process for Determining Dependability

Parameter	Description	Threshold	Test Results
$R(t_m)$	mission reliability	0.75	0.74
k	probability of failure detection	0.90	0.89
$M(t_r)$	probability of repair given a specified turnaround time	0.85	0.39
M_o	operational maintainability	0.85	0.50
$POR(t_o)$	operational readiness given a specified turnaround time	0.94	0.83
D	dependability	0.96	0.87
MDT_b	Mean Down Time Between Missions (hours)	6.3	24.4
MDT_d	Mean Down Time During Missions (hours)	0.3	0.7

Figure 3-6-12. Operational Readiness and Dependability Results

OE/OS/OSur Conclusions

Determining OE

The final step in the evaluation is to arrive at the top-level answers, i.e., determining if the system is OE. OE is the first answer that must be computed based on the MCL scores of the subordinate COIs. Systems with multiple COIs must have their answers aggregated using the weights from the SEP. Although not fully illustrated here, the process consists of taking the MCL score (MCLi) for each COI and multiplying by the COI weights (wi) from the SEP.

The next step is to sum the weighted scores to arrive at an aggregate score across all missions. When the evaluation contains only one COI, the weight defaults to 100 percent, thereby making the MCL score the same as the aggregate scores computed below. The resultant score determines if the system is OE or Not OE using the formulas in figure 3-6-13:

Aggregate MCL Score	Conclusion
$80 \leq \sum_{i=1}^{n} w_i(MCL_i) \leq 100$	OE
$0 \leq \sum_{i=1}^{n} w_i(MCL_i) < 80$	Not OE

Figure 3-6-13. Defining MCL Scores as OE/ Not OE

Example of Arriving at MCL Score

A sample data set for the ABC Attack System is seen in figure 3-6-14. The example assumes the ABC System has two COIs weighted at 50 percent each. The MCL score for each COI is in the third column, while the weighted MCL score is in the fourth.

COI=*i*	w_i	MCL_i	$w_i(MCL_i)$
1	50%	69.0	34.5
2	50%	87.0	43.5
$OE = \sum w_i(MCL_i)$			78.0

Figure 3-6-14. Sample Data Set with Two COIs

In the example, the overall weighted score is 78, translating into a conclusion of Not OE. The shortfall can be traced immediately to lower than expected capability in COI-1. As previously shown in the analysis, the shortfall results from less than adequate operational readiness and dependability, both of which are elements of Operational Suitability.

Both COIs are equally weighted, so some trade-offs can be made between mission capability, but the shortfall is so great in COI-1 that the performance in COI-2 cannot provide enough counterbalance for the overall system to be OE.

To see just how sensitive the answer is to the relative weighting, figure 3-6-15 shows that a minimum 22 percent change in weighting values (COI-1: 50 percent to 39 percent and COI-2: 50 percent to 61 percent) would be required to change the outcome to OE.

COI=*i*	w_i	MCL_i	$w_i(MCL_i)$
1	39%	69.0	26.8
2	61%	87.0	53.2
$OE = \sum w_i(MCL_i)$			80.0

Figure 3-6-15. Sensitivity to Relative Weighting

Assuming the weights were qualitatively derived, the relative change could be compared to variances in the estimated weight from stakeholders to see if this shift is within the margin of error.

Transferring Decision Model Answers to the OER

The numerical values representing OE and MCL are a means to an end, not the end itself. They represent a method to arrive at a conclusion consistently. The OE and MCL values should be present in the analytic annex, but a plain language translation should be used in the main body of the OER. For example, plain language for the example used in this chapter would be as follows: "The Anti-Tank System is not Operationally Effective. The Marines using the system were unable to achieve the minimum level of effect using the system in both types of required missions. The Marines using the system in defensive missions were able to exceed expectations but were not able to meet expectations in offensive missions. Defensive mission capability is not sufficient to offset offensive mission shortcomings."

Determining OS and OSur

Having determined OE, the evaluator now determines OS and OSur. To simplify the example, use of multiple COIs and determination of OSur are dropped here, but the procedure for examining OSur remains the same. Using the OE score as a point of reference, the evaluator determines OS and OSur with the same analytic model used to analyze OE results.

In the ABC system example, the evaluator already knows that the system is not achieving sufficient effect, evidenced by the MCL score of less than 80. The next step is to trace the source of the problems to one or more root causes.

Performing Sensitivity Analysis

The data set used to arrive at the conclusions for OS and OSur is the same as that for OE. However, more detail is required to isolate the cause of the shortfall in effect to Performance, Suitability, and/or Survivability. The evaluator must understand the constituent components of the evaluation model to arrive at causality. To answer this, the evaluator performs sensitivity analysis on the results to see the parameters' influence on the OE outcome. Computations of the analytic model are redone by first setting all performance parameters to Threshold values and setting all OS parameters to Tested Results.

The result from this sensitivity analysis indicates that when considering OS by itself, the system falls below the minimum score of 80. Based on this process the evaluator can determine if Suitability is a root cause contributor to poor effectiveness results.

Sensitivity analysis on the performance parameters is accomplished the same way as for OS. The evaluator reruns the computations of the analytic model by first setting all OS parameters to Threshold values and setting all performance parameters to Tested Results.

Transparency of Evaluation

The preceding example illustrates the mechanism for deliberately and systematically arriving at a series of evaluation conclusions. This process, while lengthy, is sufficiently transparent to allow outsiders to examine and replicate results in an independent setting. The process is also a useful tool for decision makers and engineers when deciding on improvements to current capabilities. The transparency of the process and its analytic nature also lend themselves to future evaluations that might occur to see how the system

Annex A: Data Depiction

Why Depict?

MCOTEA depicts data for several reasons—to create visual interest, to support the text, and of primary importance, to explore the data. Graphics are instruments for reasoning about quantitative information. For example, graphs can be used to evaluate changes over space or time, compare ideas, and provide a tool for evidence-based reasoning. The following pages provide guidance for MCOTEA staff, in particular ORSAs, for creating appropriate depictions of data.

Exploring the Data

Exploratory data analysis can maximize insight into data sets, detect outliers and anomalies, and test underlying assumptions. For example, graphing reveals patterns in the data that would not be apparent from a table or spreadsheet.

Further exploration can also aid in determining the distribution of the data, which will help to determine valid methods of statistical analyses. For example, when comparing test data against a theoretical normal distribution, graphing, along with goodness of fit tests, will help determine whether the data is normally distributed. A graphical and statistical analysis of the data distribution is also required for Reliability equations. For example, if a Reliability equation assumes an exponential distribution, graphing (and goodness of fit tests) will help validate that assumption.

How Often to Depict?

A good ratio to strive for in technical documents is 25 percent depiction (graphs, tables, diagrams, and images) and 75 percent text. In the main body of MCOTEA's reports, which are targeted at the 05/06-level audience, the graphs should match the overall level of the text. The reports' technical annexes are appropriate for graphs that are more analytical, such as distribution plots.

However, even these more technical graphs should present the data in a manner that allows the reader to quickly and unambiguously grasp what the data means.

Which Depictions to Use?

The type of depiction needed for a document depends on the data and the point trying to be made. The following questions are helpful in trying to decide on a depiction method:

- Are categories of data being compared?
- Are trends or correlations between two or more variables visible?
- Are trends depicted over time?
- Is data distribution visible?
- Is Reliability data being depicted?
- Are survey results being depicted?
- Are OE, OS, and OSur results being depicted?

Types of Depiction

Graphs are used to display data efficiently, meaningfully, and unambiguously to supplement and support the text of the document. They reinforce and clarify the text by telling a story pictorially. Distribution graphs can include histograms, line graphs, and probability plots. Line graphs and histograms are easy to read and are helpful in depicting outliers and skewness as well as the distribution of the data. Probability plots, which can include Q-Q plots, are a more powerful approach to comparing distributions, but require more skill to interpret.

The information contained in this section is from the NIST/SEMATECH e-handbook of Statistical Methods (2010). General guidelines for graphing are as follows:

- **Histograms** are similar to bar graphs

> What is to be sought in designs for the display of information is the clear portrayal of complexity. Not the complication of the simple; rather the task of the designer is to give visual access to the subtle and the difficult—that is, the revelation of the complex.
>
> *–Edward Tufte (1996)*

except each bar represents numbers that are grouped and form a continuous range from left to right. Histograms can be used to depict the range and distribution of the data and the presence of outliers.

- **Line graphs** are scatter plots with lines connecting the data points. Line charts are appropriate for displaying how data changes over time. Often, the dots will be connected to illustrate this, but if a logical connection does not exist, the dots should not be connected. To avoid scaling effects, a rectangular plot with the x-axis about 1.5 times as long as the y-axis is appropriate.
- **Bar charts** display the relationship between categorical variables (x-axis) and quantitative variables (y-axis). For more than eight categories, use a rotated bar chart. Stacked bar graphs should be used with caution as it is difficult to make comparisons. If a stacked bar graph is needed, the category that requires comparisons should appear on the bottom.
- **Probability plots** help to determine if the data follows a given distribution. If the data forms a somewhat straight line on the plot, then it follows a normal distribution. Any data that does not appear on the line represents a departure from normal distribution.
- **Q-Q plots** are a type of probability plot that verify if two similar sets of data can be fit with the same distribution. Q-Q plots can test many different aspects of the data and can also assess goodness of fit graphically and quantitatively with a probability plot correlation coefficient.
- **Boxplots** are good for depicting the median and upper and lower quartiles. Some boxplots also depict outliers. Boxplots can be used for comparing the distribution of the data of two or more groups and are especially good for non-parametric data since means are not appropriate parameters for non-normal data.
- **Scatter plots** are one of the most efficient graphs for depicting data and are used to detect trends or correlations between two quantitative variables. The x-axis shows the independent variable and the y-axis shows the dependent variable. Regression lines quantitatively describe the linear relationship between the two variables.

Tables usually outperform graphs in reporting small data sets and are valuable for reporting exact numerical values. It is difficult to call attention to a series of data points in a table of numbers; graphing the data points is an effective way to highlight them.

Pie charts should be avoided because they do not allow easy comparisons of data and make it difficult to discern differences in the magnitude of each slice. They also use a large amount of ink to depict a relatively small amount of data.

In General

When depicting data in any document, use strict rules of integrity to guard MCOTEA's reputation as an independent and unbiased evaluator. In addition, adhere to the following guidelines:

- Depictions should be clear and concise.
- Unnecessary chart decorations, heavy lines, overuse of color, etc. , waste space in the depiction or are a distraction. All ink should be used efficiently to aid in conveying what the numbers mean.
- Avoid the use of 3-D charts, which add clutter and distort the data.
- Clearly define what the numbers represent on the graph.
- Clearly label axes: spell out acronyms and abbreviations on the labels.
- Keep gridlines faint or delete them altogether.
- Limit numerical labels on the y-axis to avoid clutter. Consider labeling each data point with the value if a small set of numbers is depicted.
- Show error bars whenever possible. Use the caption or the graph itself to inform the audience of the type of error depicted.
- Legends and chart titles should be embedded into the chart to maximize the size of the area used for displaying the data. Legends to the side of a graph can shrink data depiction space and should be avoided.

Microsoft Excel® vs. Word®

Pro: Microsoft Excel uses an easier automated results population from SQL or other databases. The user is also able to add locked-in drop tables to reduce variations.

Con: Converting an Excel spreadsheet to .pdf requires reformatting for compilation. Also, the tools and capabilities within Excel are not always familiar to general users.

Pro: Microsoft Word tables are more easily manipulated than Excel tables. Word also provides the user with a more familiar toolset. Converting documents to .pdf does not require reformatting.

Con: Microsoft Word does not offer automatic population for tables, which could lead to version control issues. Word also provides limited formula support and no drop table support.

Conclusion: when data will be kept solely on a CD (no hard copy required), Excel may be the better choice. When lengthy data must be printed, Word may be the better choice.

- Avoid cherry-picking data. Use all available data if possible. Defend the reasoning for not using all data within the text of the report.
- Axes should be consistent across comparisons and encompass the full range of the data. If the full range of data is not depicted on the axes, explain why.
- Graphs should not be used to decorate a few numbers. If a point can be made sufficiently with words, then a graph is not needed.
- Keep design variation constant to maintain the integrity of data depiction variation (e.g., don't vary the axes intervals or the vertical or horizontal scales).
- Ensure that the x-axis is about 1.5 times longer than the y-axis to avoid exaggerating the data.
- Do not include a title. Figure numbers will be added beneath the graphic.
- Ideally, graphics should be kept in their own folder and submitted separately from the text. Ensure that graphics are numbered for editorial placement. Ensure that all graphics are referenced in the text.

MCOTEA strives for consistency throughout test program documentation. Formats for all documents are similar in terms of font choice, outlining convention, table formats, and basic page layout (see the MCOTEA SharePoint for templates). Graphics should also be consistent in terms of originating software, format, and content.

Working with Graphics in MCOTEA Documents

System Evaluation Plan

The SEP contains numerous formulas as well as standard graphics.

Formulas

- Create all formulas using equation-making software such as Equation Maker in Word 2007 or MathType. Use lower case letters for subscripts.
- Set the font to Cambria Math italic and the font size to 10.
- Use the one-line table (in the SEP template) to center the formula on the page directly beneath the text that leads to it.
- Number the formula in the table's right-hand column in parentheses.

Mission Capability Level Value Function

The MCL is used to evaluate OE/OS/OSur for all systems. The graphics that support this evaluation are standard and are available in the SEP template folder. For additional information regarding the MCL Value Function, see chapter 3-1.

Test Plans

Test Plans contain a number of standard tables (as found in the template) but few unique graphics apart from the tables. Graphics that do appear will generally be maps or Trial Conduct diagrams unique to each program.

- All large maps, photos, or diagrams should be compressed before inserting in Word to minimize document size. All images should be saved in .jpg format.
- .pdf graphics should not be inserted into Word. If a .pdf must be used, the graphic cannot contain typos or other errors.
- Diagrams should be drawn in Visio, if available, or Word. In either case, be certain to group the diagram when it is finished. To group a diagram, hold Shift while clicking on the separate parts or Select All if available. When all parts have been selected, right click and select "Group."

Reports

Test Data Reports tend not to contain graphics other than tables. However, Evaluation Reports do contain numerous graphics due to data analysis.

When formulas are used in reports, follow the SEP guidelines. Formulas from the SEP can be copied into the OER to save time.

Sample Depictions, Good and Bad

Do not include a title; it will appear in the figure description

Keep background plain (no color or heavy lines)

Ensure that the x-axis is about 1.5 times longer than the y-axis

Figure 3-6-16. Sample Line Graph

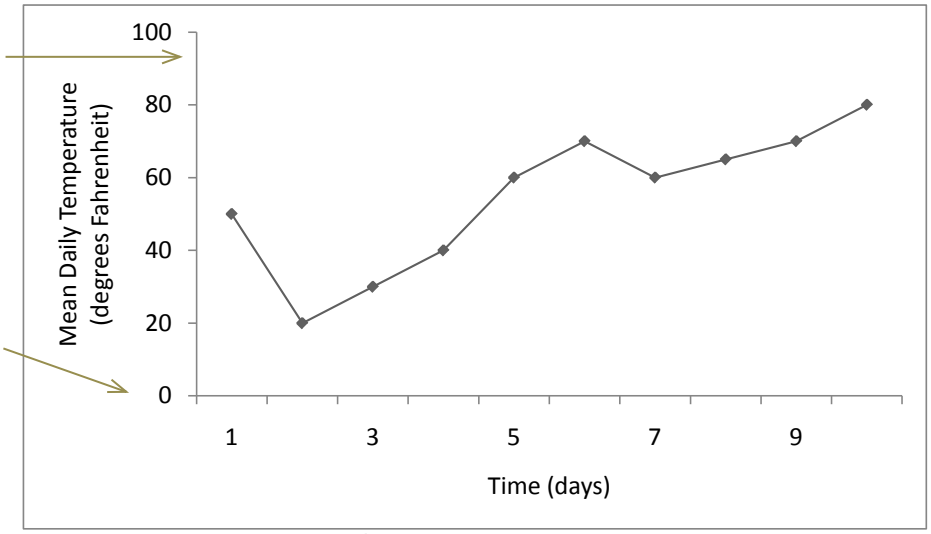


Figure X. Change in Temperature Over Time

A line graph (figure 3-6-16) is similar to a scatter plot; however, instead of using a regression line to show the relationship between the variables, the points are connected by a line to show how the data changes over time. The x-axis is two times as long as the y-axis to avoid inadvertent exaggeration of information.

Figure 3-6-17 depicts a two-axis column line chart displaying two sets of data using three axes. This graph is easy to read because a clearly defined legend is at the top and the graph contains no distractions from the data being presented.

Gray out lines

Gray out borders and set at .25 points

Fig. 3-6-17. Sample Two-Axis Column Line Chart

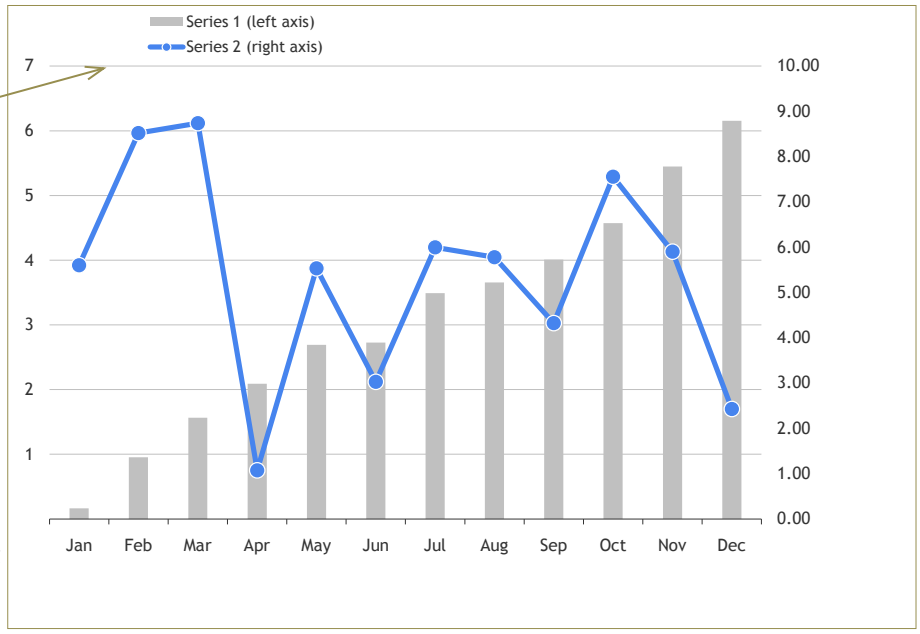


Figure X. Average and Observed Change Over Time

Bar charts (figure 3-6-18) are used to compare categorical and quantitative data. For example, they are often used to compare categories, depict survey results, and show the distribution of data. Error bars should be included in bar charts.

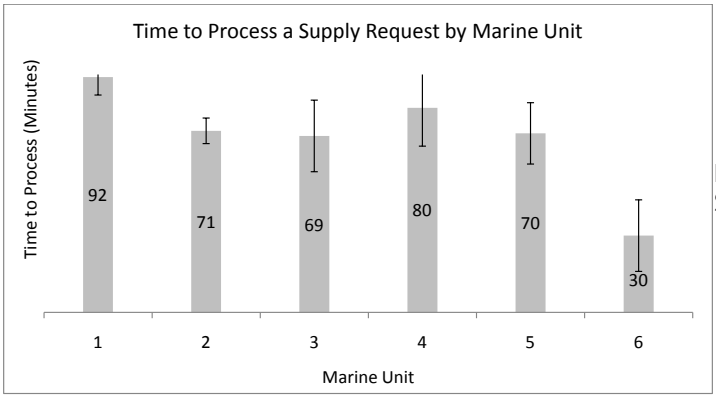


Figure X. Time to Process a Supply Request by Marine Unit

Fig. 3-6-18.
Sample Bar Chart

In a scatter plot (figure 3-6-19), the independent variable appears on the x-axis and the dependent variable appears on the y-axis. A regression line often appears on a scatter plot to show a correlation in the data. However, the data must be evaluated to determine if a correlation is intended.

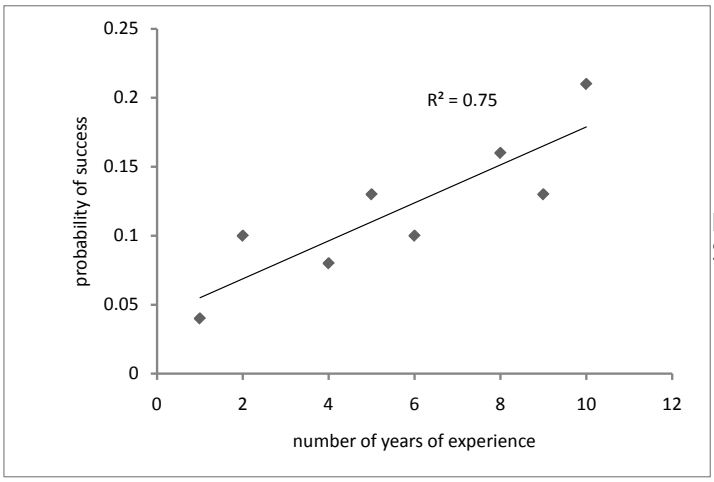
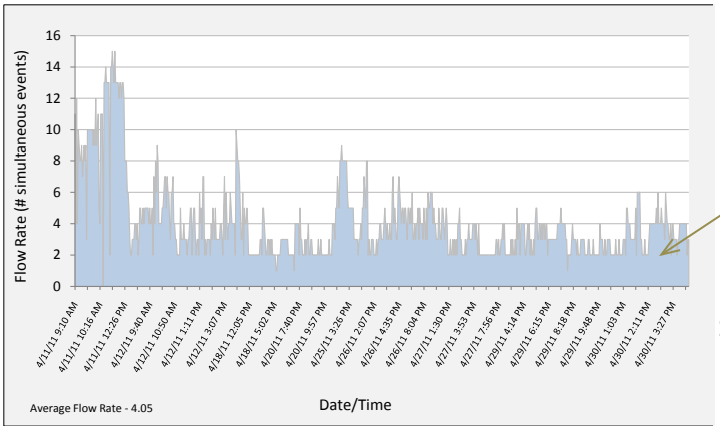


Figure x. Probability of Success by Number of Years of Experience

Fig. 3-6-19.
Sample Scatter Plot

Figure 3-6-20 exemplifies the use of color when color is useful in depicting data. (Color by itself is not necessary and may in fact create distraction.) When choosing colors for a depiction, muted colors allow the audience to focus on the data rather than the color scheme. Also, it is best to refrain from using red, green, and yellow, as these colors create a stop light effect, which is not appropriate for evaluative documents.



Use a muted color so as not to distract from the data

Figure x. Flow Rate by Day/Time

Fig. 3-6-20.
Sample Color

What Not to Do

No need to bold.
(Does not aid comprehension)

Figures 3-6-21 and 22 are not ideal for use in technical documents. Three-dimensional presentation does not clarify data and in fact can obscure important features of the data.

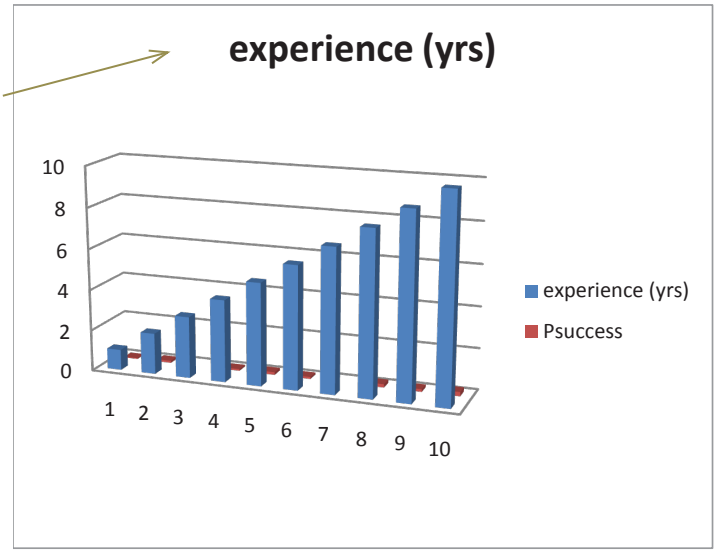


Fig. 3-6-21.
Sample Three-dimensional Graph

Pie charts do not allow easy comparison between pieces of data. In addition, too many design elements, such as the color and three-dimensional presentation, interfere with interpretation.

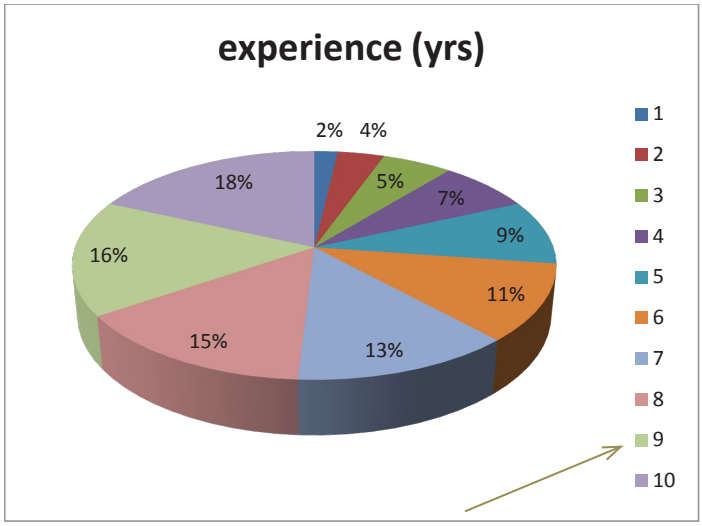


Fig. 3-6-22.
Sample Pie chart

Clearly define what the numbers represent

This page intentionally left blank for two-sided printing.

4 Documentation

Standardized Documentation Approach
General Template Descriptions and Usage Guidance
General Documentation Process Guidance
Standard Documentation Timelines
CRB Document Approval Process
Cover Page Marking Guidance

Chapter 4. Documentation

MCOTEA produces a variety of documentation throughout the life of a program. Each step of MCOTEA's test and evaluation process generates at least one document product, ranging from relatively simple letters to comprehensive test plans and evaluation reports. This chapter explains the nature of each document, its schedule, author, and content.

MCOTEA's Standardized Approach to Documentation

MCOTEA produces nearly all of its test and evaluation documentation in a repeating, standardized format that supports scientific and technical reporting. Repeating the format allows each document to "feed" the next one, creating ease of use, consistency, and traceability throughout a program's T&E history. MCOTEA T&E documents, in typical order of creation, are as follows:

Early Test and Evaluation Planning

- **System Evaluation Plan (SEP)**. Pre-MS B; sets forth the evaluation plan the program will follow for the duration; prepared by the ORSA.
- **FD/SC Charter**. Written by the ORSA with MCSC and DC, CD&I.
- **Test Concept**. Working document prepared as slides for briefing purposes, developed by the mathematical statistician after the SEP and before the TEMP.
- **Feasibility of Support**. Naval message outlining requirements for test personnel and facilities; generated by OTPO/S-3

DT Observation

- **DT Observation Plan**. MCOTEA's plan for observing DT events, written by the Test Manager.
- **Observation Report**. Documents the adequacy of DT execution with no judgment or conclusion; written by the Test Manager, usually before test results are available.
- **Intermediate Assessment Report**. MCOTEA's report of system progress, written by the ORSA/OTPO; and addressed to the PM and the MDA.

Operational Test Plans

If used, these plans require a SEP as their basis (very few programs will require the full list of Test Plans). They are written by the Test Manager.

- **Early Operational Assessment Test Plan (EOATP)**. Specifies test logistics and the detailed planning of test trials at the Issue/Subtask level.
- **Operational Assessment Test Plan** (OATP). Specifies test logistics and the detailed planning of test trials at the Issue/Task level.
- **Initial Operational Test Plan (IOTP)**. Specifies test logistics and the detailed planning of test trials at the COI/Mission level and lower levels as required.
- **Follow-on Operational Test Plan (FOTP)**. Specifies test logistics and the detailed planning of test trials of any post-IOT events.
- **Multi-Service Operational Test Plan (MOTP)**. Specifies test logistics and detailed planning for MCOTEA's participation in Multi-Service testing.

Operational Test Reports

- **Test Data Report**. Packages the test data from the test event (before analysis) and is written by the DM/MS.
- **Operational Test Agency Assessment Report (OAR)**. Evaluation report that follows an EOA/OA; stops short of OE/OS/OSur; does not support a milestone decision; OTPO/MS/ORSA prepare.
- **Operational Test Agency Evaluation Report (OER)**. Documents final system evaluation after IOT; provides OE/OS/OSur designation; OTPO/MS/ORSA prepare.
- **Operational Test Agency Follow-On**

Evaluation Report (OFER). Used after FOT&E as the final evaluation report; OTPO/MS/ORSA prepare.

- **Operational Test Agency Milestone Assessment Report (OMAR)**. Focuses on a specific acquisition milestone, either A, B, or C. Content varies based on the milestone being supported. The OMAR assesses and summarizes risk/progress towards meeting system requirements and risk/progress towards a determination of effectiveness, suitability, and survivability. For example, a Milestone A OMAR contains very little test information while a Milestone B OMAR is essentially a risk assessment.
- **System Assessment Plans, System Assessment Test Plans, System Assessment Test Reports, and System Assessment Reports**. All documents follow the regular templates, tailored for less detail.

Joint or Multi-Service Test Events

- **Documents (and schedules) conform to those of the lead Service**. If MCOTEA is the lead, documents continue to conform to standard plan and report templates, with the addition of annexes for Joint contributions.

M&S Accreditation Process

The Accreditation Plan and Report follow MIL-STD 3022. The Accreditation Decision Letter is a standard naval letter, as explained in the VV&A chapter in vol. II of this manual.

Base Templates

MCOTEA uses base templates for plans and reports. These documents are based on a prescribed set of paragraphs organized the same way for all programs. Exceptions are the TEMP and FD/SC Charter, which follow different formats because they are not produced solely by MCOTEA.

Using Citations in Text

In preparing T&E documentation, authors must abide by MCOTEA's requirement for citation. The credibility of technical documentation is based in

net/mike/docs/TrafficEngineering.pdf

The Reference list should not include sources used in a general sense, such as educational background. Formulas used in SEPs or Evaluation Reports do not need to be referenced if they are MCOTEA-derived; if they are taken from another source, such as a textbook or website, they do need to be referenced.

CRB Document Approval Process

All documents proceed through MCOTEA's chain of command for approval, and most T&E-related documents require the Director's signature. All program documentation must be edited before entering the approval process. When constructing a program's POA&M, the OTPO must include time for the CRB to receive and review documents.

The lead time for submitting documents to the CRB for initial review is seven calendar days. The CRB is scheduled through the Deputy's office. The CRB reviews the draft document for technical content and adherence to MCOTEA process, format, and standards. After CRB approval and any required changes, the document is prepared for the Director's review.

Document Change Process

Documents such as SEPs, TEMPs, and Test Plans may need periodic updating after they are signed as major program milestones are achieved. See the S-1 for change process guidance.

General Guidance for Using Templates

Templates are stored on the MCOTEA SharePoint site. Standard timelines and approval requirements are shown in figure 4-1. See the S-1 for additional guidance on document development.

Cover Pages

The covers of all MCOTEA documents contain a Distribution Statement, which is based on the document's content and purpose. A complete explanation of cover markings is contained in Annex A of this chapter.

Executive Summaries

Most MCOTEA documents are short enough that an Executive Summary is not required. However, a summary should be included when the main body of a document exceeds four pages. Executive Summaries are automatically included with an OER, regardless of length.

The following paragraphs are used in an Executive Summary: Purpose, Background, Scope, Conclusions, and Recommendations (include top three only).

The summary *must not exceed one page* in length and should not carry any information or ideas that are not contained in the main document itself. The best way to write an Executive Summary is to finish the main document first, then copy and paste key ideas from the paragraphs with the same headers noted above into the summary.

Graphics

Guidance for creating original graphics to be used in MCOTEA documents is contained at the end of chapter 3-6. Graphics or photos coming from other sources should be large enough (~1 MB) to reproduce well.

Annexes

The template for each document lists any required annexes. To support consistency among MCOTEA documents and to keep them as streamlined as possible, no additional annexes should be included without CRB concurrence.

Figure 4-1. Standard Documentation Timeline

Task	Responsible Entity	Schedule/Planning Factor (includes signature)	CRB Approval Required (Y/N)	Release/Signature Authority
SEP/SAP (including FD/SC Charter and RAM Evaluation Plan, if applicable)	ORSA	Program entry + 120 calendar days (CDs) (80 working days (WDs)). Time limit is for ACAT I only; all others will be produced in less time.	Y	Director signs.
Letter: DT Plan Late/Missing	OTPO	TEMP identified date or 30 CDs (20 WDs) before DT event if not specified.	N	Division Head signs and sends to PM.
Observation Plan	TM	7 CDs (5 WDs) from receipt of final Developmental Test (DT) Plan	N	Division Head signs.
Observation Report	TM	14 CDs upon return from DT event (10 WDs)	N	Division Head signs.
Letter: DT Data Late/Missing	OTPO	TEMP identified date or 30 CDs (20 WDs) after DT event completion if not specified.	N	Division Head signs and sends to PM.
IAR/SAR	ORSA	45 CDs (30 WDs) after receipt of Data. IARs can be distributed individually or aggregated after last required DT event, depending on program. IARs may be timed for use at Gate Reviews.	Y	Director signs and sends to PM/MDA.
Feasibility of Support Message	OTPO/S-3	90-180 CDs (60-120 WDs) before test*. This is a standard naval message.	N	Director concurs.
Test Concept (internal planning brief)	MS	The earlier of either 200 CDs prior to OT, or prior to submission to PM as MCOTEA TEMP input.	Y	Director concurs. (DOT&E concurs if Oversight program).
Test and Evaluation Master Plan (including Test Concept)	OTPO	60 CDs (40 WDs) or as designated by T&E WIPT.	Y	Director signs (DOT&E approves if Oversight program).
OTRB	OTPO	90 CDs (60 WDs) before test* begins for MCOTEA Test.	N	Director concurs.
Test Plan	TM	75 CDs prior to test* for IOT/FOT and EOA/OA (or before test for other events).	Y	Director signs (DOT&E approves if Oversight program).
OTRR Brief	OTPO	30 CDs (20 WDs) before test* in support of OT.	N	Director concurs.
Test Data Report	MS	9 CDs (7 WDs) after test completion.	Y	Director signs.
Data Review Meeting	MS	30 CDs prior to publishing any Assessment or Evaluation report.	Y	CRB concurs.
OAR/OER/OFER/OMAR	ORSA	45 CDs (30 WDs) after test.	Y	Director signs. Sent to ACMC. Copy to PM/MDA (DOT&E for Oversight).
Lessons Learned	OTPO	45 CDs (20 WDs) following signature of final report.	N	OTAD concurs.
Accreditation Plan	ORSA	30 CDs before V&V activity.	Y	OTAD signs.
Accreditation Report	ORSA	30 CDs before OTRB.	Y	OTAD signs.
Accreditation Decision Letter	OTPO	30 CDs before OTRB.	Y	Director signs.

* The word “test” is defined as all activities that include New Equipment Training, Pilot Test, and Record Test.

Editorial References

MCOTEA abides by a number of standard editorial references, such as the Navy Correspondence Manual (for letters and memos), the Government Printing Office Style Manual (for government style issues), and the Chicago Manual of Style (for technical guidance on topics such as citing references or setting up mathematical equations). MCO 5216.20 provides additional Marine Corps-specific guidance on style and usage.

Annex A: Marking Cover Pages

Distribution Statements

All MCOTEA T&E documents are marked with an appropriate distribution statement (DOD 1987). The reference provides policies and procedures for marking technical data for release and dissemination without additional approvals or authorizations. If applicable, all MCOTEA T&E documents are also marked with an appropriate export control warning in accordance with DODD 5230.25. No MCOTEA document is distributed without first undergoing a proper security classification review and assignment of a distribution statement.

The Division Head or appropriate staff lead is responsible for determining the distribution code and applicability of an export control warning for all programs assigned.

Method

MCOTEA uses the guidance contained in this section to select an appropriate distribution statement. Contractor Sensitive documents are always either B or E.

Proper Marking of Documents

The document originator is responsible for ensuring that the appropriate markings are applied. The distribution statement must be displayed conspicuously on electronic documents. For standard written or printed material, the distribution statement appears on each front cover and title page. See the examples in figure 4-2. The majority of MCOTEA's documents will use Distribution Statements C or D.

If the technical information is not prepared in the form of an ordinary document and does not have a cover or title page (such as forms, spreadsheets, and charts), the applicable distribution statement shall be stamped, printed, written or affixed by other means in a conspicuous position.

Defense Technical Information Center

Some reports must be submitted to the Defense Technical Information Center (DTIC). See the S-1 for additional guidance.

DTIC Submission Process

MCOTEA submits reports electronically in .pdf as part of the Records Management Process. The responsible Division or staff section prepares a .pdf copy of the report with all proper markings on the cover page and submits that and a .pdf copy of the SF298 (Submission Form) to S-1. The S-1 is responsible for establishing and maintaining a DTIC account and electronically submitting each report via the DTIC website. See the S-1 for additional guidance.

Cover Page

MCOTEA documents should have a completed cover page that includes all necessary information identified in the templates provided by the S-1.

Definition of Technical Data

Recorded information related to experimental, developmental, or engineering works that can be used to define an engineering or manufacturing process or to design, procure, produce, support, maintain, operate, repair, or overhaul material. The data may be graphics and pictures, text in specifications or related performance or design type documents, or computer printouts. Examples of technical data include research and engineering data, engineering drawings, and associated lists, specifications, standards, process sheets, manuals, technical reports, catalog-item identifications, and related information and computer software documentation.

Figure 4-2. Distribution Statement Examples

DISTRIBUTION A. Approved for public release; distribution is unlimited. (*Note – Documents recommended for Public Release must first be reviewed in accordance with DoD Directive 5230.9)

DISTRIBUTION B. Distribution authorized to U.S. Government Agencies only (fill in reason) (date of determination). Other requests for this document shall be referred to Marine Corps Operational Test and Evaluation Activity, 2032 Barnett Ave, Quantico, VA 22134-5014.

DISTRIBUTION C. Distribution authorized to U.S. Government Agencies and their contractors (fill in reason) (date of determination). Other requests for this document shall be referred to Marine Corps Operational Test and Evaluation Activity, 2032 Barnett Ave, Quantico, VA 22134-5014.

DISTRIBUTION D. Distribution authorized to DoD and U.S. DoD contractors only (fill in reason) (date of determination). Other requests for this document shall be referred to Marine Corps Operational Test and Evaluation Activity, 2032 Barnett Ave, Quantico, VA 22134-5014.

DISTRIBUTION E. Distribution authorized to DoD Components only (fill in reason) (date of determination). Other requests for this document shall be referred to Marine Corps Operational Test and Evaluation Activity, 2032 Barnett Ave, Quantico, VA 22134-5014.

DISTRIBUTION F. Further dissemination only as directed by Marine Corps Operational Test and Evaluation Activity, 2032 Barnett Ave, Quantico, VA 22134-5014.

DISTRIBUTION X. Distribution authorized to U.S. Government Agencies and private individuals or enterprises eligible to obtain export-controlled technical data in accordance with DoD 5230.25 (date of determination). Controlling DoD office is Marine Corps Operational Test and Evaluation Activity, 2032 Barnett Ave, Quantico, VA 22134-5014.

EXPORT CONTROL WARNING AND DESTRUCTION NOTICE. In addition to the Distribution Statement verbiage, the following Export Control Warning and Destruction Notice verbiage must also be listed if the document contains technical data that is export controlled (or if documentation is not available stating otherwise):

DESTRUCTION NOTICE – For classified documents, follow the procedures in DoD 5220.22-M, Industrial Security Manual, Section 11 -19 or DoD 5200.1-R, Information Security Program Regulation, Chapter IX. For unclassified, limited documents, destroy by any method that will prevent disclosure of contents or reconstruction of the document.

WARNING - This document contains technical data whose export is restricted by the Arms Export Control Act (Title 22, U.S.C., Sec 2751, et seq.) or the Export Administration Act of 1979, as amended, Title 50, U.S.C., App. 2401 et seq. Violations of these export laws are subject to severe criminal penalties. Disseminate in accordance with provisions of DoD Directive 5230.25.

This page intentionally left blank for two-sided printing.

Marine Corps Operational Test and Evaluation Activity

Operational Test & Evaluation Manual

Ancillary Topics

Third Edition
Volume II

This page intentionally left blank for two-sided printing.

1 Live Fire T&E

Volume II, Chapter 1. Live Fire Test and Evaluation

Background

Public law mandates that major weapon system and munitions programs, as well as product improvements to those programs that are likely to significantly affect the vulnerability or lethality of those programs, undergo a realistic Live Fire Test and Evaluation (LFT&E) program.

Simply put, LFT&E is the realistic testing of platforms or munitions against real threats expected to be encountered in combat. The basis of the evaluation for LFT&E characterizes the system under test against the current and future threat environment.

This section provides the Marine Corps process for LFT&E programs. It presents the basis for determining whether an LFT&E program is required for a given system, delineates the two types of LFT&E programs, outlines the key steps in developing an adequate LFT&E strategy, and describes the key building blocks of LFT&E.

A realistic LFT&E building block program represents the best alternative to "actual" combat in assessing the system's performance; however, with the lack of actual combat data a disciplined and realistic approach to assessing the vulnerability and lethality of weapons systems must be articulated. A well-planned and well-structured LFT&E program reduces the potential for surprises on the battlefield.

An early, active, well-planned, well-managed, and well-executed LFT&E program is essential to understanding the system. It is also essential for supporting decisions regarding the system's acquisition as well as the development of tactics, techniques, and procedures for its proper operational employment. A properly structured and integrated LFT&E program, within the overall context of the system T&E strategy, will enable design changes to be incorporated into the system at the earliest possible date, thereby reducing the need for expensive retrofit programs.

Objective

LFT&E supports a timely and thorough system vulnerability/lethality assessment during development and subsequent production phases. It should demonstrate the weapon system's or munition's ability to provide battle-resilient survivability or lethality. LFT&E should provide insights into

- The principal damage mechanisms and failure modes for the platform/target occurring as a result of the munition/target interaction
- Techniques for reducing personnel casualties or enhancing system survivability/lethality

Data that emerges can be used to support cost-effectiveness trade-offs to predict the optimal "mix" of vulnerability/lethality enhancement measures as early as possible in the acquisition cycle.

The primary emphasis of LFT&E is testing under realistic combat conditions as a source of personnel casualty, system vulnerability, and system lethality information to ensure potential design flaws are identified and corrected before full-rate production. The LFT&E program should assess a system's vulnerability/lethality performance relative to the expected spectrum of battlefield threats; it is not constrained to addressing specific design performance goals or threats. LFT&E should also assess the battle damage assessment and repair (BDAR) capabilities to enhance system survivability.

Requirement for LFT&E

Public law requires LFT&E on "covered" systems before proceeding beyond low-rate initial production (LRIP). A "covered" system is defined as a system which provides protection to users in combat or is considered a "major" system. A system shall be considered a "major" system if one of the following categories is met:

- Total expenditures for research, development, test, and evaluation for the system are estimated to be more than $365 million (in FY00 dollars)
- Total expenditures for procuring the system is estimated to be more than $2,190 million (in FY00 dollars)
- The Secretary of Defense or the Secretary of the Navy designates it as such (special interest).

A "major" munitions program meets one of the above criteria or has plans to acquire more than 1,000,000 rounds.

Specifically, the legislation requires side by side vulnerability LFT&E if a wheeled or tracked armor vehicle is to replace an existing vehicle; requires LFT&E for all covered systems and major munition and missile programs; and requires LFT&E for product improvements to major systems (modification or upgrades). Figure 1-1 depicts the process for determining a system's LFT&E requirement and addresses both new systems and changes (modifications, upgrades, or follow-on blocks) to existing systems. Additionally, recent Defense Authorization Acts have included language that specifically calls for the LFT&E of equipment not normally subject to LFT&E, e.g., Personal Protection Equipment (PPE) such as helmets and tactical vests. PPE items are now covered under LFT&E as "special interest" programs. LFT&E programs are subject to DOT&E oversight.

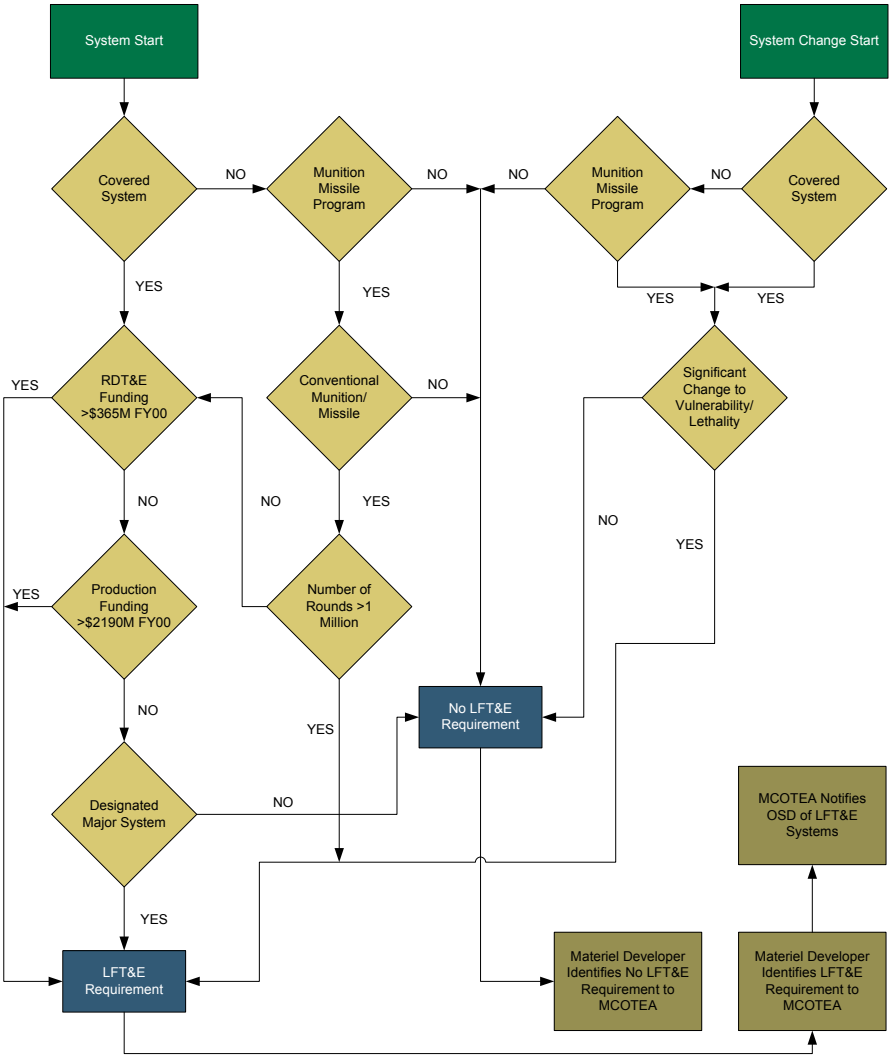
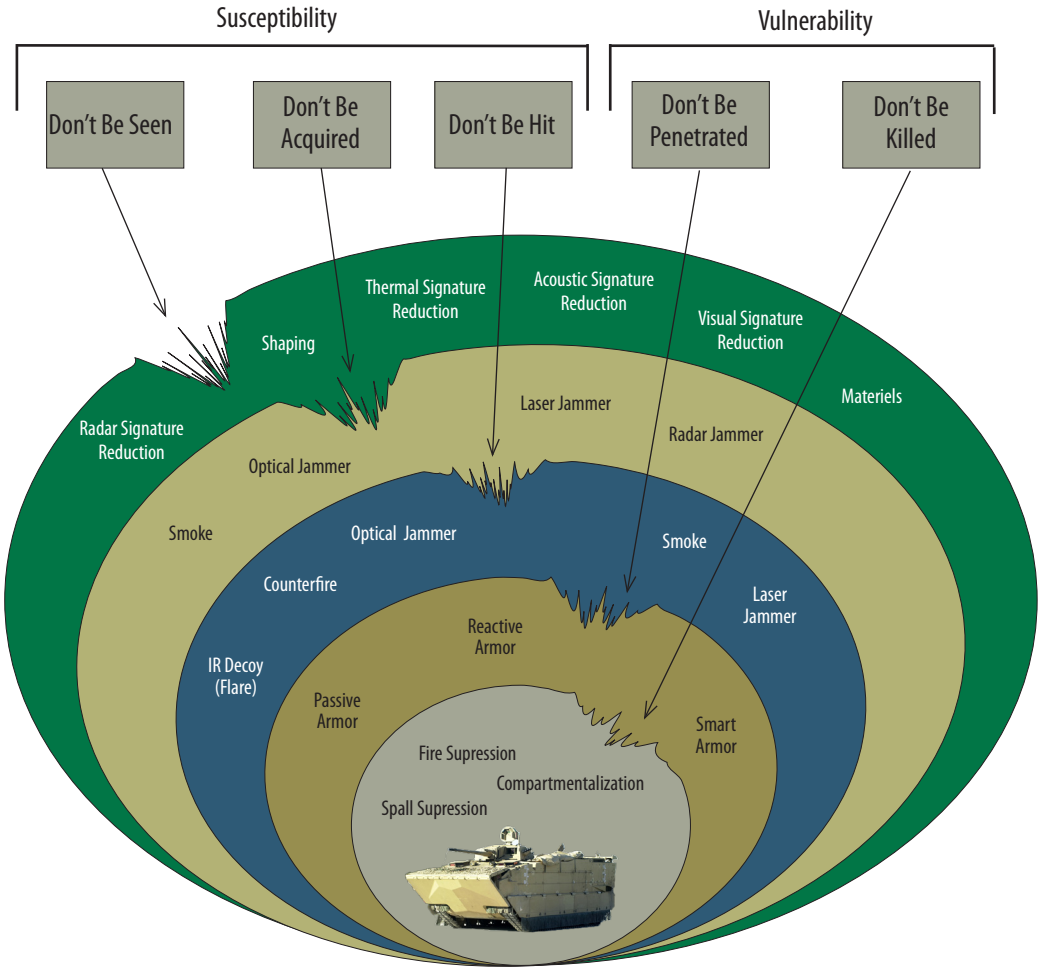


Figure 1-1. LFT&E Requirements Flow Chart

Figure 1-2. Elements of Survivability



Vulnerability LFT&E

LFT&E comprises two major components: vulnerability and lethality. Vulnerability LFT&E focuses most specifically on the system's response once a threat affects the system, i.e., penetration and kill, which is depicted by the inner layers of figure 1-2.

Penetration

Penetration involves the actual defeat of the platform protection system, normally the armor. Armor systems are designed to meet a protection specification, which is normally delineated as the defeat of a certain round or munition. For example: "The vehicle will provide protection against the 7.62 mm round at zero degrees of elevation at any azimuth at the muzzle velocity." However, LFT&E is not specification-focused testing. LFT&E addresses all realistic threats likely to be encountered on the battlefield; as such, the platform is subject to "overmatching" threats. While preliminary specification-based validation testing will confirm baseline requirements compliance, LFT&E evaluates other rounds and determines the conditions and distances from which these other rounds are able to penetrate the platform. Overmatching is the term used to descriLive Fire Testbe testing against realistic, real-world threats that are known to exceed the baseline requirements. In this example, although the overmatched weapon would be expected to penetrate the armor, the test is performed to determine the level of functionality the system retains, as well as the number and nature of injuries incurred after penetration. This data can then used to adjust the system to mitigate the effects of the overmatched threats. This helps address one of the goals of Vulnerability LFT&E—the

Vulnerability LFT&E deals with the "...testing for vulnerability of the system in combat by firing munitions likely to be encountered in combat (or with a capability similar to such munitions) at the system configured for combat, with the primary emphasis on testing vulnerability with respect to potential user casualties...and combat performance of the system" (Major systems and munitions programs 2008).

characterization of the platform's armor system's overall resistance to penetration.

Kill

The "kill" of a system or embarked crew/personnel refers to the resultant damage from a threat penetration. After a round penetrates a system, several damage mechanisms affect platform-critical functionality, such as mobility, firepower, communication, etc. There are also several distinct and concomitant damage mechanisms that affect personnel survivability (Force Protection). Vulnerability LFT&E examines how a platform mitigates post-penetration damage mechanisms such as behind armor debris (BAD), spall, ballistic penetration (the round itself), secondary projectiles, toxic fumes, shock and acceleration, and fire. Another goal of Vulnerability LFT&E is to characterize a system's loss of functionality and embarked crewmen/personnel incapacitation after the platform has been breached by a threat.

Lethality LFT&E

Lethality LFT&E, the less common of the two types of LFT&E programs, is concerned with the system's offensive capabilities. Lethality LFT&E deals with the "...testing for lethality by firing the munition or missile concerned at appropriate targets configured for combat (Major systems and munitions programs 2008)." Lethality is the weapons system's ability to cause the loss of, or the degradation in, the target system's ability to complete its designated mission. In requirements documents, lethality is normally delineated in the form of a "target set" or "target list." This target set outlines the required targets and the desired effect on each target. For example: The platform "will suppress infantry in the open at 1,000 meters" or "will destroy Light Armored Vehicles at 800 meters." The major components of lethality, shown in figure 1-3, are accuracy and terminal effects. Typically, the effectiveness of accuracy, (seeing, acquiring, and hitting the target) is resolved during OT&E as part of the system's Operational Effectiveness. Terminal effects (penetrating and killing the target), the inner two circles in figure 1-2, are examined during Lethality LFT&E. Ultimately, an end to end mission profile using real rounds against real threat targets is normally conducted as part of IOT&E lethality testing. Lethality is referred to as "Vulnerability LFT&E in reverse." As such the parameters of "penetrate" and "kill" are presented in both.

LFT&E Management

While the details of each element of an overall LFT&E program must be decided on a case by case basis, this chapter presents the general approaches and lessons learned from previous successful Marine Corps LFT&E programs and should prove beneficial to those involved in future LFT&E

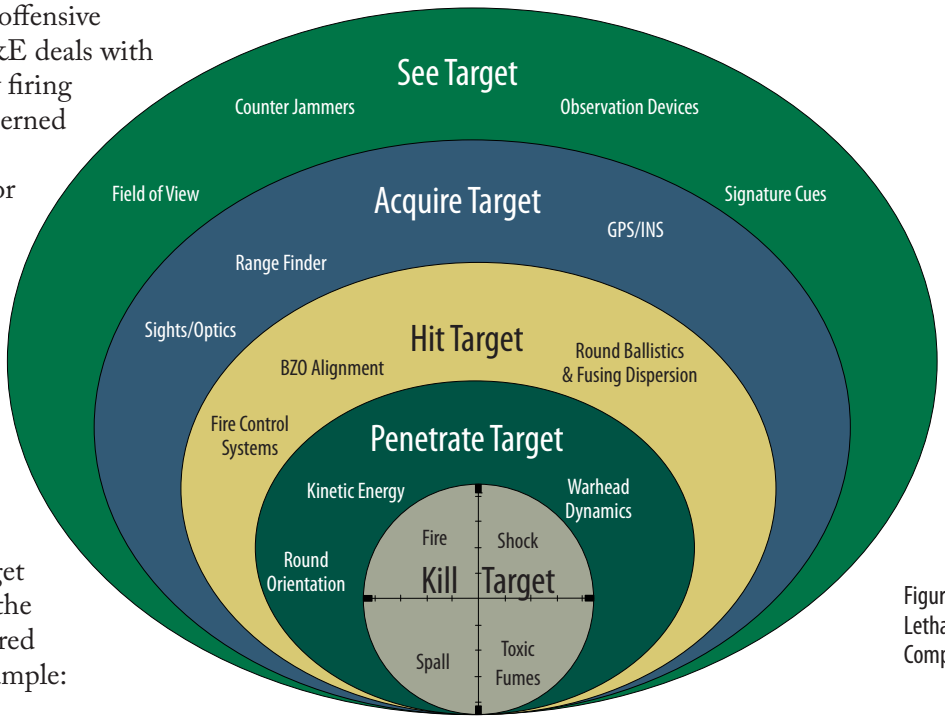


Figure 1-3. Lethality Components

programs. Figure 1-4 depicts the basic elements of the overall LFT&E process from the initial strategy definition to the writing of the final LFT&E report. Before documenting issues to support LFT&E strategy development, the LFT&E analyst must obtain the COIs for the system from MCOTEA's test team. These COIs form the basis for the critical LFT&E issues.

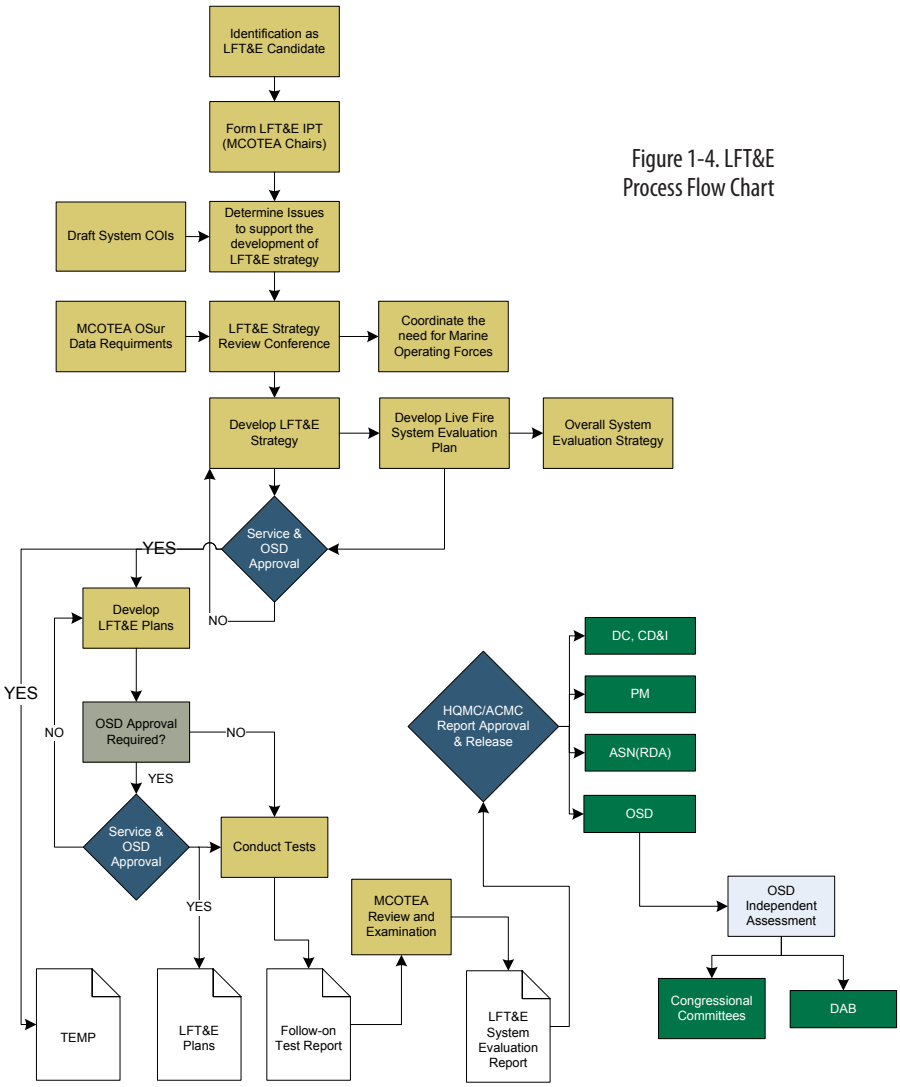


Figure 1-4. LFT&E Process Flow Chart

The "Strategy Review Conference" depicted in figure 1-4 constitutes stakeholder concurrence with the overall LFT&E strategy.

Although current legislation only requires LFT&E in certain cases, it provides a means of ensuring that Marines using the system in combat are protected to the highest degree possible. The scope of LFT&E needs to be addressed in a comprehensive LFT&E strategy, incorporated into the appropriate documentation, and provided to Marine Corps leadership for guidance and approval. According to the Secretary of the Navy, the Naval materiel developer is responsible for executing LFT&E (Secretary of the Navy 2008). The materiel developer is responsible for completing the system's LFT&E and ensuring that the LFT&E Report is completed and submitted prior to a Full-Rate Production decision. SECNAVINST 5000.2 delineates that the materiel developer must submit the LFT&E Report to DOT&E via MCOTEA. The materiel developer has the following options available when addressing the inherent requirement to execute LFT&E:

- Task MCOTEA to execute and report on LFT&E; historically this is the preferred technique to conduct LFT&E within the Marine Corps. In this arrangement, MCOTEA chairs the LFT&E IPT.
- Execute LFT&E with MCOTEA oversight; historically this option has been taken with minor "special interest" LFT&E programs. In this arrangement MCOTEA co-chairs the LFT&E IPT
- Task an outside technical agent/agency to conduct LFT&E on behalf of the materiel developer; this is the least preferred method and involves both MCOTEA oversight and the inclusion of an outside agency, which may or may not have the requisite experience in LFT&E. MCOTEA typically acts as co-chair of the LFT&E IPT in this arrangement.

The system's proposed acquisition strategy and overall evaluation strategy should include live fire testing requirements with supplementary and complementary data to be drawn from DT and OT. The system's

mature LFT&E strategy and resource requirements should be included in the system's TEMP. The program's LFT&E IPT develops and produces the LFT&E strategy. The LFT&E IPT produces and reviews all LFT&E documents. Typically there are several core members of the LFT&E IPT including representatives from

- PM/Materiel Developer
- MCOTEA
- Marine Corps Intelligence Activity
- DC, CD&I
- Technical test agency (normally Aberdeen Test Center for most Marine Corps ground systems, and the M&S agent (normally Army Research Lab (ARL))
- DOT&E

MCOTEA Vulnerability Process

Live fire consists of a range of testing and evaluation that begins with preliminary component, subsystem, and/or system-level tests and culminates in Full-Up System Level (FUSL) tests of system vulnerability and lethality. FUSL live fire testing satisfies the statutory requirement for "realistic survivability testing" or "realistic lethality testing." The LFT&E program includes all vulnerability/lethality T&E phases and associated modeling and analysis efforts that support the live fire evaluation.

Vulnerability LFT&E focuses on protection against lethal mechanisms and minimizing damage to the crew and hardware given an impact or breach by a lethal mechanism. In addition, vulnerability LFT&E addresses recoverability from combat damage. Critical issues for Vulnerability LFT&E address the following key areas:

- Crew/Occupant vulnerabilities (Force Protection)
- System and hardware vulnerabilities (Vehicle Survivability)
- BDAR capabilities

Figure 1-5 contains MCOTEA's LFT&E process and milestones within a generic LFT&E vulnerability program. The LFT&E IPT chair executes the checklist shown in figure 1-5; however, because each program is unique, certain elements will not apply to every LFT&E strategy. Regardless of the scope for the LFT&E program, this process serves as a guide to effectively incorporate live fire testing into the system's overall T&E strategy.

Figure 1-5. MCOTEA Live Fire Vulnerability Process

Test Concept Development	☐ Review program Documentation ☐ Review requirements/capabilities Documents ☐ Review System Survivability Specifications ☐ Form LFT&E IPT o Designate Chair (MCOTEA either Chair or Co-Chair) o ID Core members o Develop/Approve IPT Charter ☐ Obtain updated COIs from MCOTEA OA ☐ Determine Level of M&S Needed
Strategy Review Conference	☐ Present Draft LFT&E concept 1. Armor Validation Scope 2. Armor Characterization Scope 3. Armor Exploitation 4. Ballistic Hull & Turret (BH&T) Scope 5. Component Candidates for Component Ballistic Testing 6. Determine screening criteria 7. System Level test scope 8. Controlled Damage Testing (CDT) scope 9. FUSL scope 10. M&S scope ☐ Assign agencies to conduct LFT&E events ☐ Coordinate need for Marine operating forces with Force Synchronization Conference (normally BDAR participants) ☐ Present Draft LF Critical issues ☐ Present Draft LF Strategy ☐ Obtain BDAR concept plan from PM/Material Developer ☐ Coordinate with MCOTEA OA to ensure all LFT&E OSur, data requirements will be met ☐ Present Draft Live Fire System Evaluation Plan ☐ Update/Present funding/resource profile
TEMP Development	☐ Approve LF Strategy ☐ Approve LF Critical issues ☐ Submit M&S VV&A plan ☐ Submit M&S Requirements ☐ Synchronize PM/Material Developer Armor Specification compliance with LF Strategy Coordinate TEMP inputs with the MCOTEA System Evaluator pertaining to components and test assets required for LFT&E ☐ Submit Live Fire System Evaluation Plan
LFT&E Test Plan Development	☐ Develop, approve, and distribute the following plans via the LFT&E IPT 1. Armor Characterization 2. Component Ballistic Testing (CBT) 3. Armor Exploitation 4. BH&T 5. BDAR 6. System Level Test 7. CDT 8. FUSL ☐ Develop and submit M&S Accreditation Plan to Dir, MCOTEA or designated representative ☐ Track BDAR development & insure BDAR elements are addressed in all applicable LFT&E test events.
Execution of LFT&E Building Block events.	☐ Observe LFT&E events ☐ Track and review M&S VV report ☐ Report results to MCOTEA OA
FUSL TRR	☐ Insure test Asset availability ☐ Arrange for system testing for Marine participants who will conduct Post shot functionality testing and BDAR ☐ Receive preliminary M&S Accreditation
FUSL Test Execution	☐ Observe and Monitor FUSL conduct ☐ Obtain Pre-shot predictions (normally from M&S stakeholder) ☐ Compare Pre-shot predictions with actual outcomes ☐ Provide updates/SITREPs to Dir, MCOTEA ☐ Report results to MCOTEA OA
Conduct DAT activities	☐ Convene Damage Assessment Team ☐ Receive final M&S Accreditation
Produce LFT&E Report	☐ Collect and Review all applicable DT and LFT&E Reports. Report results to MCOTEA OA ☐ Receive and Review MUVES S-2 Model effort output ☐ Submit final draft of USMC LFT&E Report for MCOTEA Content Review Board ☐ Resolve any Ballistic requirements for OT&E OER ☐ Publish and Route USMC LFT&E Report (Report needs to be delivered to DOT&E 45 days prior to a Full Rate Production decision)

Building Block Approach

The building block approach helps build upon the system's sequential LFT&E.

This information, especially early in the life of a program, helps shape and improve system design. The main building blocks in a Vulnerability LFT&E program are listed chronologically but can be repeated if necessary and are defined in the following sections.

Armor Validation

Armor Validation is normally a DT, materiel developer-conducted set/series of tests executed at the armor coupon level to determine if the armor solution meets its technical specifications. Coupon testing involves testing an isolated piece of armor on its own when not incorporated in to the overall system. While a DT event, the LFT&E IPT will observe this test and receive copies of the test reports. The LFT&E IPT may require additional coupon tests that examine the environmental, multihit, and durability characteristics of the armor. Depending on the platform's intended operational environment, additional coupon tests may be required to ascertain the limit of resistance to penetration for designated overmatching threats. The results of coupon testing will be used to build the baseline resistance to the penetration module in the M&S suite and to ensure the vendor's specification compliance. Typically, the Marine Corps uses the Army's Modular Unix-based Vulnerability Estimation Suite (MUVES) S-2 M&S tool for LFT&E.

Armor Characterization

Armor Characterization identifies the Armor's BAD characteristics, which is often referred to as the "spall cone angle." Typically, several overmatching threats are examined and the BAD data is then transferred to the MUVES S-2's BAD module. Armor Characterization also defines the armor's dynamic deflection properties. Dynamic deflection testing helps the materiel developer identify the "safe" distance behind the armor. This influences the placement of critical components and seats for occupants. MCOTEA's active involvement in LFT&E typically begins with this step and continues through to the end of the system's LFT&E.

Armor Exploitation

Armor Exploitation characterizes an armor system's resistance to penetration. Instead of testing coupons, the integrated armor solution (Pre-MS B prototype) is examined. The areas of interest are usually armor seams, armor interface points, through bolts, and locking mechanisms embedded on or in the armor.

Ballistic Hull and Turret

Ballistic Hull and Turret (BH&T) testing is typically performed on a Technology Development Phase prototype to verify system-wide ballistic protection requirements (usually underbody/under wheel/under track blast requirement). This is typically an LFT&E event, but the materiel developer heavily influences the test event design. Additionally, BH&T may be used to conduct preliminary end to end Fire Extinguishing System testing across the entire platform. From an asset allocation standpoint, the BH&T asset is often used for both BH&T and Armor Exploitation. If significant vulnerabilities are discovered during these two test phases, and design improvements are made to mitigate these vulnerabilities, BH&T is then typically repeated on the follow-on design (normally a post-MS B platform).

Component Ballistic Testing

Component Ballistic Testing (CBT) examines the critical component's ballistic properties within a platform. The data from this testing provides information on the specific component's vulnerability and also provides the Probability of Component Dysfunction (PCD). The PCD for specific components is then loaded into MUVES S-2. During the CBT Phase a Criticality Analysis and Damage Assessment List are produced by ARL and DC, CD&I respectively. The Critical Analysis addresses

the system component's engineering and functional hierarchy, while the Damage Assessment List addresses a specific component's critical functionality "value" (communications, mobility, firepower) of the overall platform. Both the Critical Analysis and the Damage Assessment List are inputs to MUVES S-2.

System-Level Testing

System-Level (SL) testing examines system-wide response to threat interactions while accounting for threat tactics. Typically, vulnerabilities that were uncovered early in the LFT&E process are revisited to determine if design improvements have mitigated known vulnerabilities. SL testing also verifies and validates BDAR procedures. This testing is normally conducted on a late Engineering and Manufacturing Development Phase prototype. SL testing also allows the PM to verify any system-level ballistic requirements/specifications prior to MS-C. The evaluation from SL concomitantly affords the materiel developer with timely input to help focus the Critical Design Review before committing to an LRIP design.

Controlled Damage Testing

Controlled Damage Test (CDT) is a precursor FUSL event that occurs on the asset before the full-up shots begin. CDT, in a non-destructive format, looks to verify the Critical Analysis and update any changes to the Critical Analysis in MUVES S-2 prior to FUSL.

Full-Up System-Level Testing

Full-up System-Level Testing (FUSL) testing involves a complete, production-representative platform with all ancillary, support equipment, fuel, and ammunition onboard. MUVES S-2 is used to conduct pre-shot predictions to estimate embarked personnel incapacitation and damage to the vehicle. The FUSL pre-shot predictions are compared to the actual damage incurred to improve the fidelity of the model. The Damage Assessment Team assesses actual damage and personnel injury and incapacitation data from a FUSL event and subsequently distributes the information to the LFT&E IPT.

Marine Corps Lethality Process

Lethality LFT&E addresses both the ability to perforate or breach the target and to inflict significant damage to the target and/or its crew and occupants. Generally, the following lethal abilities are critical:

- Accurately engage a threat system (often evaluated using DT and OT data)
- Perforate or breach the threat system's protection
- Significantly degrade the threat system's combat/mission functions
- Injure/incapacitate the crew/occupants

Building Block Approach

The main building blocks in a Lethality LFT&E Program are listed chronologically and defined in the following sections.

Qualification Testing

Qualification testing is typically a DT, materiel developer-conducted set/series of tests executed to qualify the munition for service, to safety certify the munition, and to gain initial data on its capabilities.

Munition Terminal Effects Characterization

Munition Terminal Effects Characterization is typically a DT, materiel developer-conducted set/series of tests executed from a fixed firing point against representative armor coupons and surrogate targets to determine if the munition meets its technical specifications. While a DT event, the LFT&E IPT will observe this testing and receive copies of the test reports. The LFT&E IPT may require additional tests that examine capabilities of the munition against realistic target sets expected to be encountered in combat. The results of this testing are used for M&S

purposes to build the baseline terminal effects/penetration module in the M&S suite. Typically, the Marine Corps utilizes the Army's MUVES S-2 model for Lethality LFT&E. Ultimately, the purpose of this testing will be to see if an accurately delivered munition delivers the desired terminal effect.

System-Level Testing

System Level testing is normally an LFT&E event that looks to characterize the lethality of a munition when it is delivered by a host platform against realistic targets.

Engagement TTPs are typically developed during this test series/phase.

End-to-End FUSL Testing

End-to-End FUSL Testing involves the real munition, with real Marine operators, with its intended delivery system, engaging realistic targets at tactically relevant distances to characterize its operational lethality. This testing is often executed within the context of an Operational Mission Profile.

Figure 1-6 illustrates the MCOTEA process and milestone events within a generic LFT&E lethality program. Since each program is unique, certain elements will not apply to every strategy. Regardless of the scope for the LFT&E program, this process serves as a guide to effective incorporation of live fire testing into the overall test and evaluation strategy.

Vulnerability LFT&E

Breaks threats into major and minor

- A major threat kills or reduces the effectiveness of a large percentage of the systems in the force effectiveness evaluation or has a high density in the force
- all other threats are considered minor

PM provides funding

Lethality LFT&E

The MCIA representative to the LFT&E IPT plays a key role in identifying potential realistic targets to the LFT&E IPT

Based on the target driving the design (usually the most difficult target to kill given a hit)

Identifies threat target requirements and availability

PM provides funding and acquires targets

LFT&E Key Elements

Each live fire program contains several critical elements, defined in further detail in this section.

Scope of LFT&E

The following questions should be considered in order to prepare a properly scoped LFT&E strategy:

- What are the technical and operational characteristics of the concepts, technology, and requirements for the system?
- How do they differ from the system being replaced (where appropriate)?
- Which threats are to be considered in the LFT&E?

The threats considered in LFT&E should be based on a review of the System Threat Assessment (STA); the densities of various classes of threat weapons and countermeasures in organizations likely to be encountered; and the frequency that various threats kill or are killed by the system from force effectiveness analysis supporting program decisions or planning studies.

LFT&E Strategy

The LFT&E strategy is the most important element of the LFT&E process. An LFT&E concept should be prepared and approved as early as possible in the acquisition cycle, with the goal of producing a viable LFT&E Strategy by MS B. The LFT&E System Evaluator has the lead in preparing and obtaining approval for the strategy in coordination with the LFT&E IPT. The ACMC

approves the strategy for the Marine Corps before it is sent (via the TEMP) to DOT&E for approval. If consensus cannot be reached on the LFT&E scope, or if program constraints limit compliance with required reporting dates, the ACMC will be approached to help resolve the issue. The LFT&E strategy is the foundation of the LFT&E section of the TEMP and all subsequent planning documents (Live Fire System Evaluation Plan (LFSEP), Event Design Plan (EDP), Pre-Shot Prediction Report, and the Test Plan). The strategy should be detailed enough to adequately project resource requirements; schedules for major T&E efforts; and to trigger long lead time planning, procurement of threats/surrogates, and modeling.

How LFT&E results will be evaluated is formulated during the system vulnerability/lethality examination and during the definition of critical LFT&E issues. After strategy development, the evaluation process is finalized the and the details are articulated in the LFSEP and LFT&E EDP. The evaluation must crosswalk all vulnerability/lethality testing and complementary modeling and assessment with LFT&E issues. The following aspects of the evaluation process must be examined when developing the LFT&E strategy:

- Possibly using M&S to address evaluation issues pertaining to system vulnerability or lethality, crew casualties, and logistics supportability.
- Planning building block-level vulnerability tests to assess the protective system of the item under test's ability (for example, armor and optics) to withstand impacts by threat missiles and projectiles, and to examine the ability of critical components (for example, ammunition compartments) to withstand damage from a threat warhead or projectile that breaches the protective system. Early system development LFT&E will focus on the component/subsystem level to address vulnerability issues and upgrade and develop the system vulnerability model. The FUSL vulnerability LFT&E conducted against a full-up (combat-loaded) production or production-representative system is generally the last in the series of live fire tests conducted.
- Lethality LFT&E must be planned to assess the system's ability to damage target-critical components and injure/incapacitate the crew. During the early weapons system development, the testing will usually focus on the warhead's or penetrator's ability to breach the threat target's protective system. During pre-qualification and qualification testing, impact conditions will be firmly established for the missile or projectile and the ability of the warhead or penetrator to breach the target's protective system will be refined. The End-to-End FUSL lethality life five testing is the last live fire testing phase and is conducted against a full-up

Figure 1-6. MCOTEA's Live Fire Lethality Process

Test Concept Development	☐ Review program Documentation ☐ Review requirements/capabilities Documents ☐ Review System Lethality Specifications and Target list ☐ Form LFT&E IPT o Designate Chair (MCOTEA either Chair or Co-Chair) o ID Core members o Develop/Approve IPT Charter ☐ Obtain updated COIs from MCOTEA OA ☐ Determine Level of M&S Needed
Strategy Review Conference	☐ Present Draft LFT&E concept ☐ Determine screening criteria ☐ Assign agencies to conduct LFT&E events ☐ Present Draft LF Critical issues ☐ Present Draft LF Strategy ☐ Present Draft Live Fire System Evaluation Plan ☐ Coordinate need for Marine operating forces with Force Synchronization Conference ☐ Update/Present funding/resource profile
TEMP Development	☐ Approve LF Strategy ☐ Approve LF Critical issues ☐ Submit M&S VV&A plan ☐ Synchronize PM/Material Developer Lethality Specification compliance with LF Strategy ☐ Coordinate TEMP inputs with the MCOTEA OA pertaining to components and test assets required for LFT&E ☐ Submit Live Fire System Evaluation Plan
LFT&E Test Plan Development	Develop, approve, and distribute the following plans via the LFT&E IPT o Qualification Testing o Munition Terminal Effects Characterization o System Level Test Plan o End-to-End FUSL Test Plan (normally executed during IOT&E) ☐ Develop and submit M&S Accreditation Plan to Dir, MCOTEA or designated representative
Execution of LFT&E Building Block events.	☐ Observe LFT&E events ☐ Track and review M&S VV report ☐ Report results to MCOTEA OA
End-to-End FUSL TRR	☐ Insure test Asset availability ☐ Arrange for system testing for Marine participants who will conduct End-to-End FUSL Testing ☐ Receive preliminary M&S Accreditation
FUSL Test Execution	☐ Observe and Monitor FUSL conduct ☐ Provide Pre-shot predictions ☐ Compare Pre-shot predictions with actual outcomes ☐ Report results to MCOTEA OA ☐ Provide updates/SITREPs to Dir, MCOTEA
Conduct Target DAT activities	☐ Convene Damage Assessment Team (DAT) for Target analysis Receive final M&S Accreditation
Produce LFT&E Report	☐ Collect and Review all applicable DT and LFT&E Reports ☐ Report results to MCOTEA OA ☐ Submit final draft of USMC LFT&E Report for MCOTEA Content Review Board ☐ Resolve any Ballistic requirements for OT&E IER ☐ Publish and Route USMC LFT&E Report (Report needs to be delivered to DOT&E 45 days prior to a Full Rate Production decision)

(combat-loaded) threat target. However, the extent of target functionality and application of combat load may be impacted by availability of assets and specific T&E requirements. If it is not possible to obtain a realistic threat target, FUSL lethality LFT&E must use the best available surrogate threat targets. The scarcity of lethality LFT&E targets and their cost may dictate that these targets not be fully combat-loaded with live munitions to preclude catastrophic loss.

- Vulnerability models are also used to estimate the spare parts and time required to repair combat damaged components. FUSL vulnerability LFT&E provides valuable inputs for refining these estimates. In addition, rapidly returning damaged systems to battle requires accurately accessing the damage and applying field-expedient repairs. Again, FUSL vulnerability LFT&E provides the materiel developer and TECOM with valuable training opportunities to refine and develop field-expedient repair methods and to identify tools and materials required to execute these repairs.

Live Fire System Evaluation Plan

Specifically, the LFSEP provides the crosswalk between the evaluation issues and the data requirements. Additionally, the data sampling plan and analysis techniques are specified to ensure the logic of the evaluation is understandable. The LFSEP will identify MOPs and MOEs associated with the issues developed in the strategy. The LFSEP includes a section describing the types of threats or targets that the system is expected to encounter during the operational life of the system and the key characteristics of the threats/targets that affect system vulnerability/lethality. Any T&E limitations or shortfalls and their impact will be discussed.

The Event Design Plan (EDP) contains guidance on the conditions and data requirements for use in the development of the Test Plans. The EDP is concerned with the higher-level issues of interest in constructing the Test Plan such as vulnerable/physical areas to examine, impact angles, and whether or not to examine seams. The EDP also describes statistical analyses, criteria, models, system comparisons, and how they support the evaluation. The EDPs provide the tester or analyst with the details on what data is required from a particular test or evaluation event. The EDP will detail the decision process for foreseeable changes in the test design. If an unexpected change in the test design is required, the change to the EDP is submitted to the Director, MCOTEA for approval 90 days prior to test initiation and is subsequently forwarded to DOT&E.

Threats

An integral part of the LFT&E strategy development is identifying the threat target (lethality LFT&E) and munition (vulnerability LFT&E) requirements. These requirements need to be identified early in the acquisition cycle to allow for possible long lead times for procurement. It is very likely that some of the required threat munitions will not be available for LFT&E. It is also likely that intelligence data on some munitions may be limited. Therefore, LFT&E may be conducted using threat munitions based on postulated technology options derived from intelligence assessments. This will require surrogates in lieu of "real" threats. The rationale for threat surrogate selection must be detailed in the LFT&E strategy.

The rationale for selecting surrogate threat projectiles for vulnerability LFT&E is to match physical performance characteristics of the projected threat. For kinetic energy projectiles, penetration into rolled homogeneous armor (RHA); muzzle velocity and impact velocity; and penetrator material, length, and diameter are typical key parameters. For shaped charge warheads, typical parameters are penetration into RHA, impact velocity, and warhead diameter and explosive type. Availability and cost of surrogate

projectiles may also drive the selection. Typically, U.S. projectiles and warheads will be selected as surrogates. The projectiles and warheads selected as threat surrogates must be submitted, along with supporting rationale, to the Director, MCOTEA for approval.

Shotlines

The attack conditions and the munition/target impact location (that is, shotline) must be identified for each shot to provide the appropriate information required the address critical LFT&E issues. The shotline selection methodology that will be used is described in the LFT&E Strategy, whereas the specific shotlines selected and the rationale for their selection must be included in the EDP for the specific test series. There are two types of shots: engineering and random. Engineering shots provide information and data to address specific vulnerability or lethality issues for a specific threat. Random shots are selected from the combat distribution of impact conditions (direction, location, and range) for the threats of interest. The minimum number of engineering shots should be selected first to address the vulnerability and/or lethality critical issues. Next, the number of random shots required for each threat weapon should be selected. Random shots should be reviewed to determine if any remaining engineering shots are duplicated or if a critical issue is satisfied by a random shot. Those remaining engineering shots duplicated by a random shot should be eliminated.

The LFT&E program should be planned independent of constraints and then efforts must be made in developing and approving the strategy to obtain relief from schedule and resource constraints. The most likely outcome of this process is compromise and negotiating strategies that meet the spirit and intent of the law within existing or modified constraints. The following parameters must be selected and specified:

- Posture (offense or defense)
- Range (based upon offense or defense)
- Angle of attack (stratified into equal probability intervals to ensure sampling over all possible attack angles with small sample sizes)
- Target side (left or right)
- Hull or turret
- Horizontal dispersion
- Direction of horizontal dispersion (left or right)
- Vertical dispersion
- Direction of vertical dispersion (up or down)

LFT&E Reporting

The LFT&E Report documents the live fire vulnerability/lethality evaluation and contains the assessment of the critical issues and conclusions concerning the vulnerability/lethality and battlefield damage assessment and system repair (vulnerability live fire programs only). The LFT&E Report addresses the test objectives, issues, and criteria as defined in the LFSEP, EDPs, and BDAR Support Plan. It discusses the crosswalk between results and the evaluation and specifies any limitations relative to the analysis. The LFT&E Report objectively addresses all aspects of the system vulnerability/lethality based on the likelihood of occurrence on the battlefield. Not all vulnerabilities identified in Vulnerability LFT&E can be fixed. Constraints on system funding, system weight, and other aspects necessitate the ranking of the identified vulnerabilities from the perspectives of likelihood of occurrence on the battlefield and the degree of system degradation given an occurrence. The final LFT&E report provides this information to the user and to the PM for resolution.

This page intentionally left blank for two-sided printing.

2 RAM

Volume II, Chapter 2. Reliability, Availability, and Maintainability (RAM)

Reliability, Availability, and Maintainability, collectively known as RAM, is a critical component of OS and is inextricably linked to OE. The link between effectiveness and the individual components of RAM can be explained as follows:

- Weapon systems that are not ready for use (Availability) when needed prevent effects from occurring. Weapon systems can be unavailable because there aren't enough to go around, including spares to keep pace with operational demand; or, the systems assigned to a unit are undergoing a maintenance action to restore or preserve functionality.
- A weapon system that malfunctions (Reliability) when operating affects the Marine's ability to achieve a desired effect. A weapon system that malfunctions requires repairs to either correct the malfunction or prevent its reoccurrence.
- Systems undergoing repairs (Maintainability) are unable to be used when called for at any random, given point in time. This situation is exacerbated when repair actions are difficult or time-consuming, causing a reduction in system Availability.

To understand RAM one must first understand the basic definitions and mathematical expressions. What follows are extracts from various DOD publications.

Availability (A)

Availability is defined as a measure of the degree to which an item is in an operable and committable state at the start of a mission when the mission is called for at a random point in time. Availability is the parameter that translates system Reliability and Maintainability characteristics into an index of effectiveness. It is based on the question, "Is the equipment available in a working condition when it is needed?"

The basic mathematical definition of Availability is defined by the equation:

$$Availability = A$$

$$= \frac{Up \quad Time}{Total \quad Time}$$

$$= \frac{Up \quad Time}{Up \quad Time + Down \quad Time}$$

Where

Up Time = Time the system is available to perform designated mission.

Down Time = Total Time–Up Time = Time system is unavailable for tasking.

Actual assessment of Availability is accomplished by substituting the time-based elements defined above into various forms of this basic equation (DOD 1982).

Material Availability (A_m)

Am measures the percentage of systems in operational use—providing a meaningful snapshot of the overall efficiency of the program elements (design, support structure, use profiles, planned and unplanned maintenance downtimes, and so on) to provide the necessary capability to the warfighter or end user. Am is not a substitute for operational readiness metrics (such as Operational Availability (Ao), Mission Reliability, Mission Capability Rate). Am provides the trade space between acquisition and support costs related to the system design and support approach. Am applies to all end items acquired throughout their life cycle, while operational readiness metrics apply to end items in the operational environment only—excluding float/spare systems, systems at depot for overhaul or repair, systems that have not been operationally assigned, and so on.

When a system capability that includes

planned float/spare systems is fielded, Am is defined by the following equation:

$$A_M = \frac{Number \quad of\ Operational \quad End \quad Items}{Total \quad Number \quad of \quad End \quad Items \quad Acquired}$$

Assessment of the achieved Am involves determining the number of operational end items (i.e., those ready for tasking) divided by the total number of end items acquired at the time the sample is taken.

When a system is being fielded without float/spares then all acquired end items are put into operational service and remain there unless maintenance is required. Under these conditions, the following equation is used:

$$A_m = \frac{Up \quad Time}{Up \quad Time + Down \quad Time} = \frac{MTBM}{MTBM + MDT}$$

Where

MTBM = Mean Time between maintenance actions requiring removal of system from operational use

MDT = Average system Down Time expected given the anticipated support structure (RAM-C Report Manual 2009).

Operational Availability (A_o)

A_o covers all segments of time that the equipment is intended to be operational. Up Time includes Operating Time plus nonoperating (Standby) Time (when the equipment is assumed to be operable). Down Time includes preventive and corrective maintenance and associated administrative and logistics lead time. All are measured in clock time using the following formula:

$$A_m = \frac{Up \quad Time}{Up \quad Time + Down \quad Time} = \frac{MTBM}{MTBM + MDT}$$

Where

Total time = Up Time + Down Time

OT = Operating Time = when the equipment is in use, further defined as the time during the accomplishment of a mission profile when the system is turned on and actively performing at least one, if not all, of its functions.

ST = Standby Time = not operating but assumed operable in a specified period, further defined as Up Time when a system is not committed to accomplishing a specific mission profile.

TPM = Total Preventive Maintenance Time = scheduled maintenance time per specified period.

TCM = Total Corrective Maintenance Time = unscheduled maintenance time per specified period.

ALDT = Administrative and Logistics

Down Time = time spent waiting for parts, administrative processing, maintenance personnel, or transportation per specified period.

This relationship is intended to provide a realistic measure of equipment availability when the equipment is deployed and functioning in a combat environment. A_o is used to support operational testing assessment, life cycle costing, and force development exercises. One significant problem associated with determining A_o is that it becomes costly and time-consuming to define the various parameters. Defining ALDT and TPM under combat conditions is not feasible in most instances. Nevertheless, the A_o expression does provide an accepted technique of relating standard Reliability and Maintainability

elements into an effectiveness-oriented parameter (DOD 1982).

Achieved Availability (A_a)

Aa is frequently used during development testing and initial production testing when the system is not operating in its intended support environment. Excluded are operator before-and-after maintenance checks and standby, supply, and administrative waiting periods. Aa is much more a system

$$A_a = \frac{OT}{OT + TCM + TPM}$$

hardware-oriented measure than is Ao, which considers operating environment factors. It is, however, dependent on the preventive maintenance policy, which is greatly influenced by non-hardware considerations. All times are measured in clock time using the formula:

Inherent Availability (A_i)

A_i is useful in determining basic system operational characteristics under conditions which might include testing in a contractor's facility or other controlled test environment. Likewise, A_i becomes a useful term to describe combined Reliability and Maintainability characteristics or to define one in terms of the other during early conceptual phases of a program when, generally, these terms cannot be defined individually. Since this definition of Availability is easily measured, it is

$$A_i = \frac{OT}{OT + TCM} = \frac{MTBF}{MTBF + MTTR}$$

frequently used as a contract-specified requirement.

A_i defines system availability with respect only to Operating Time and Corrective Maintenance and can be expressed using the formula:

Where

MTBF = Mean Time Between Failures= the average time during which all parts of the item perform within their specified limits, during a particular measurement period under stated conditions (DOD 2005).

MTTR = Mean Time To Repair = includes diagnostic time (time to detect and isolate failure); time to repair (in-place repair or removal and replacement); and time required to validate the repair (e.g., functional check) (DOD 2005).

Under these idealized conditions Standby and Delay Times associated with Scheduled or Preventive Maintenance can be ignored as well as Administrative and Logistics Down Time. As is evident from this definition, A_i provides a very poor estimate of true combat potential for most systems because it provides no indication of the time required to obtain required field support. This term should normally not be used to support an operational assessment (DOD 1982).

Reliability

Reliability measures the probability that the system will perform without failure over a specified interval under specified conditions. Reliability must be sufficient to support the warfighting capability needed in its expected operating environment.

Considerations of Reliability must support both A_o and A_m. Reliability may be expressed initially as a desired failure-free interval that can be converted to a failure frequency for use as a requirement (DOD 2009).

Two very different system Reliability design objectives exist. One is to enhance system effectiveness; the other is to minimize the burden of owning and operating the system. The first objective is addressed by means of Mission Reliability, the second by means of Material or Logistics-related Reliability. Measures of Mission Reliability address only those incidents that affect mission accomplishment. Measures of Logistics-related Reliability address all incidents that require a response from the logistics system (DOD 1982).

Mission Reliability

Mission Reliability is the probability that a system will perform mission essential functions for a period of time under the conditions stated in the mission profile. Mission Reliability for a single shot type of system, i.e., a pyrotechnic device, would not include a time period constraint. A system with high Mission Reliability has a high probability of successfully completing the defined mission. A typical Mission Reliability metric is Mean Time Between Operational Mission Failure (MTBOMF) for systems with a continuous Reliability requirement. MTBOMF is defined as a measure of operational mission Reliability for the system; it is the average time between operational mission failures that cause a loss of the system's "mission" as

$$MTBOMF = \frac{TOT}{\#\ of\ \ OMFs}$$

defined by the customer. This parameter may include both hardware and software "failures." This parameter also includes failures that are generally attributed to human errors during operation and maintenance that cause failures (DOD 2005). MTBOMF can be expressed by the equation (MOA on MOT&E 2009):

Where

TOT = Total Operating Time = summation of all periods of Operating Time when the equipment is in use. Operating Time is further defined as the time during the accomplishment of a mission profile when the system is turned on and actively performing at least one, if not all, of its functions.

OMFs = Operational Mission Failures = any incident or malfunction of the system that causes (or could cause) the inability to perform one or more designated mission essential functions (TRADOC/AMC 1987).

Therefore, a Mission Reliability analysis must include the definition of mission essential functions (DOD 1982).

Mission Essential Functions

Mission Essential Functions (mef) are the minimum operational tasks that the system must be capable of performing to accomplish the assigned mission. Descriptions of mission essential functions should be in operational terms that relate to mission requirements. The equipment operator should be able to readily identify the loss of a mission essential function (DOD 1982).

Mefs have both a qualitative and quantitative aspect. Qualitatively mefs are brief statements, usually infinitives, that declare why the given equipment is needed, what its purpose is. Typical mefs include "to move," "to shoot," and "to communicate." Quantitative mefs are followed up with quantitative information to describe the break point between satisfactory and unsatisfactory performance of the function. Using the "to move" qualitative example the quantitative aspects might be to "travel at 30 mph on cross-country under full load," (TRADOC/AMC 1987).

Material Reliability

Material Reliability (R_m), also known as Logistics Reliability or Basic Reliability, is a characteristic of the final system design. All indicated and recorded failures, even those that do not affect successful completion of the mission, eventually result in some corrective action (repair). Corrective action often includes some

$$MTBF = \frac{1}{\lambda}$$

level of repair or inspection to mitigate the

λ =System failure rate=the total number of failures within an item population, divided by the total time expended by that population, during a particular measurement interval under stated conditions.

failure. Repairs in this case can consist of

removal and replacement, in-place repair, or some combination thereof for the failed item (DOD 2005).

R_m is defined by the Mean Time Between Failures (MTBF) of the system:

Where

Logistics-related Reliability measures, as indicated above, must be selected so that they account for or address all incidents that require a response from the logistics system. Logistics-related Reliability may be further subdivided into Maintenance-related Reliability and Supply-related Reliability. These parameters respectively represent the probability that no Corrective Maintenance or the probability that no unscheduled supply demand will occur following the completion of a specific mission profile (DOD 1982).

Maintainability

Maintainability, along with Reliability, is one of the two major system characteristics that combine to form Availability. Maintainability and maintenance are not the same. Maintainability is a design consideration whereas maintenance is the consequence of the design (DOD 1982). Maintainability is defined as the probability that an item can be retained in, or restored to, a specified condition in a given time when maintenance is performed by personnel having specified skill levels, using prescribed procedures and resources, at each prescribed level of maintenance and repair. Maintainability is a function of the design (DOD 2005).

Maintenance

Maintenance is the term used to define all actions required to retain an item in, or restore it to, a specified condition. This includes diagnosis, repair and inspection. Maintenance can be further subdivided into Preventive and Corrective Maintenance.

Preventive (Scheduled) Maintenance

Preventive Maintenance (PM) is defined as systematic inspection, detection, and correction of incipient failures either before they occur or before they develop into major defects. Adjustment, lubrication, and scheduled checks are included in the definition of preventive maintenance (DOD 1982).

Preventive Maintenance actions are considered Scheduled Maintenance Actions (SMA). SMAs are services or repairs performed at intervals measured by calendar time, by use (hours of operations, rounds fired, etc.), or by condition (wear limits, low battery power, depleted lubrication, etc.). To qualify as an SMA the maintenance must be prescribed by an equipment publication and enough latitude in the time to perform the maintenance must exist that it can be done in a slack period between missions (TRADOC/AMC 1987).

A system undergoing Preventive Maintenance can be considered either available or unavailable depending on the effect of the maintenance action on the system's ability to perform mefs. Preventive Maintenance that inhibits the accomplishment of a mef causes the system to be unavailable. An example would be a routine brake inspection on a vehicle whose mef is to move. If the wheel of the vehicle has to be removed to inspect the brakes during a Preventive Maintenance check, then that vehicle loses the ability to perform the move mef. Therefore, during the Preventive Maintenance period the system is considered unavailable until such time as the vehicle is capable of performing all of its mefs.

Corrective (Unscheduled) Maintenance

Corrective Maintenance (CM) is defined as that maintenance performed on a non-scheduled basis to restore equipment to satisfactory condition by correcting a

malfunction. All CM actions are considered unscheduled maintenance actions. The importance of the repair action will often dictate when the repair takes place. Repair actions can be performed immediately, deferred until after the mission but before the next mission, or deferred until a time when the system is not required to be available. Deferring maintenance until a scheduled period of Down Time does not change the corrective action from unscheduled to scheduled. The existence of a failure or malfunction is what determines whether or not the maintenance action is unscheduled vice scheduled.

Depending on the urgency of the repair actions and the impact to mefs, CM actions can be classified in three basic categories: Operational Mission Failures (OMF), Essential Maintenance Actions (EMA), and Unscheduled Maintenance Actions (UMA).

Operational Mission Failures

Operational Mission Failures (OMF) are incidents that require immediate resolution in order to resume or perform a mission. When a system is undergoing corrective action as a result of an OMF the system is considered unavailable. If the malfunction is such that the repair can be deferred, and the mission can be continued, then the incident is not considered an OMF. A special case of OMFs is called crew-correctable maintenance actions (CCMA) (TRADOC/AMC 1987).

Crew-Correctable Maintenance Action

CCMAs are optional, but when used they are defined as those minor interruptions of the mission which the crew overcomes by quick, local action. CCMAs are resolved by the crew using only the system's onboard tools, repair parts, and spares. Crew action need not be maintenance, but can be simply a powering down and powering up of the equipment. The amount of time allowed to a CCMA before the incident becomes a more serious stoppage (i.e., an OMF) depends on the mission. Within a given system different CCMA times may be allowed for the different mefs according to the function and its urgency to the mission.

CCMAs may occur multiple times in a mission. The occurrence of multiple CCMAs in a single mission may have the aggregate effect comparable to an OMF. In other words, when too many minor interruptions occur, the net effect can be a failure of the mission. The cumulative effect should be defined in advance in the FD/SC Charter (TRADOC/AMC 1987).

Essential Maintenance Actions

EMAs are incidents in which the malfunction, or the deviation from specification, has to be corrected for complete mission readiness. At times special conditions exist or alternative methods or components for carrying out a mission are present that make what is otherwise considered an OMF to be an EMA. As an example, if the headlights are broken during daylight-only missions, then the incidents are considered EMAs vice OMFs (TRADOC/AMC 1987).

Unscheduled Maintenance Actions

All maintenance actions not otherwise the result of OMFs, CCMAs, or EMAs are considered UMAs.

Maintenance Metrics

Maintenance metrics are based on a few key observations that are predominately based on time, personnel, and parts/spares used. From this information Mean Time To Repair (MTTR), Maximum Time To Repair (MaxTTR), and Maintenance Ratio (MR) can be derived. The other component of maintenance relates to the logistics aspect, and is Administrative and Logistics Down Time.

Mean Time to Repair

MTTR, also called Mean Corrective Maintenance Time, is the total Corrective Maintenance Down Time accumulated during a period divided by the total number of Corrective Maintenance Actions completed during the same period. MTTR includes

- Diagnostic Time (time to detect and isolate failure)
- Time to repair (in-place repair or removal and replacement of the failed item)
- Time required to validate the repair (e.g., functional check) (DOD 2005)

MTTR is commonly used as an on-equipment measure but can be applied to each maintenance level individually. MTTR is expressed by the following formula:

$$MTTR = \frac{TCM}{\#CMActions}$$

MTTR does not account for frequency of corrective maintenance items or for the number of man-hours expended; therefore, MTTR is not a good measure of maintenance burden (DOD 1982). An appropriate measure for maintenance burden is Maintenance Ratio (MR).

Maximum Time to Repair

MaxTTR is the maximum Corrective Maintenance Down Time within which either 90 or 95 percent (as specified) of all Corrective Maintenance Actions can be accomplished. MaxTTR is useful in special cases where the system has a tolerable Down Time. An absolute maximum would be ideal but is impractical because some failures will inevitably require exceptionally long repair times (DOD 1982).

Maintenance Ratio

MR is the cumulative number of man-hours of maintenance expended in direct labor during a given period of time, divided by the cumulative number of end-item operating hours (or rounds or miles) during the same time. MR is expressed with the formula:

$$MR = \frac{maintenance\ man\ hours}{operating\ hours}$$

The MR is expressed at each level of maintenance and summarized for all levels of maintenance combined. Both Corrective and Preventive Maintenance are included. Man-hours for off-system repair of replaced components and man-hours for daily operational checks are included for some classes of systems. MR is a useful measure of the relative maintenance burden associated with a system. It provides a means of comparing systems and is useful in determining the compatibility of a system with the size of the maintenance organization (DOD 1982).

Administrative and Logistics Down Time

Administrative and Logistics Down Time (ALDT) is the time spent waiting for parts, administrative processing, maintenance personnel, or transportation per specified period (DOD 1982). During ALDT active maintenance is not being performed on the downed piece of equipment (TRADOC/AMC 1987).

Linking Metrics to Time-Based Models

The Measures of Suitability previously presented are inextricably linked to Operational Mode Summary and Mission Profile for the system and time categorizations.

All systems have some form of time characterizations, which identify the state a system is in at any given time. For example, spare systems sitting in a warehouse may be in Inactive Time, while systems assigned to operational units would be in Active Time. An individual system may be employed on a specific mission profile putting it in Mission Time, while another might be in between missions and thus currently in Standby Time.

Many of the times categories have already been identified as part of the RAM metrics

such as Operating Time, Corrective and Preventive Maintenance Down Times, Administrative and Logistics Down Time, Standby Time, etc. Figure 2-1 depicts the discrete time categories that can be used to classify time for a system (TRADOC/AMC 1987).

A system can be in only one time category at a time, although some exceptions exist that the figure does not depict. The time categories are used to identify the state of any given system on a timeline as depicted in figure 2-2 (next page).

The timeline illustrates a single-system, single-mission example. In the example, the system starts the timeline in standby time and remains in that time status until such time as a mission are called for at the random given point in time. Mission time for this system begins with pre-operations checks, although it's worth noting that for this fictional system post-operations checks are not considered part of the mission. Following pre-operations checks the system is placed in alert time, a special case of standby time where the system is committed to a specific mission, is considered operable, but is not currently operating. When the operators are given the command to begin operating the system the time categorization changes from alert time to operating time. During operating time at least one or more mefs are in use.

This example further illustrates two types of unscheduled maintenance, the arrival of each occurs during operating time.

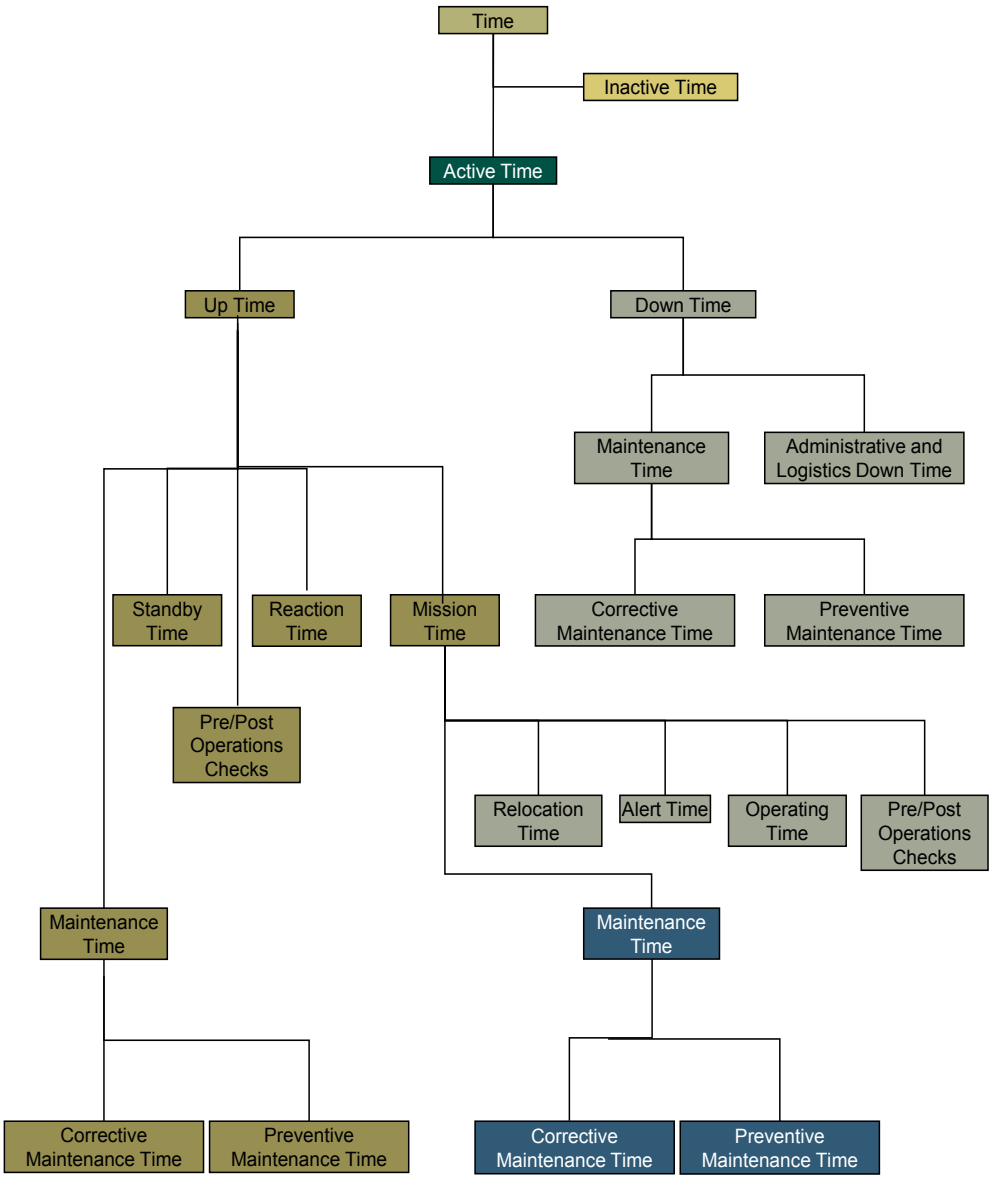
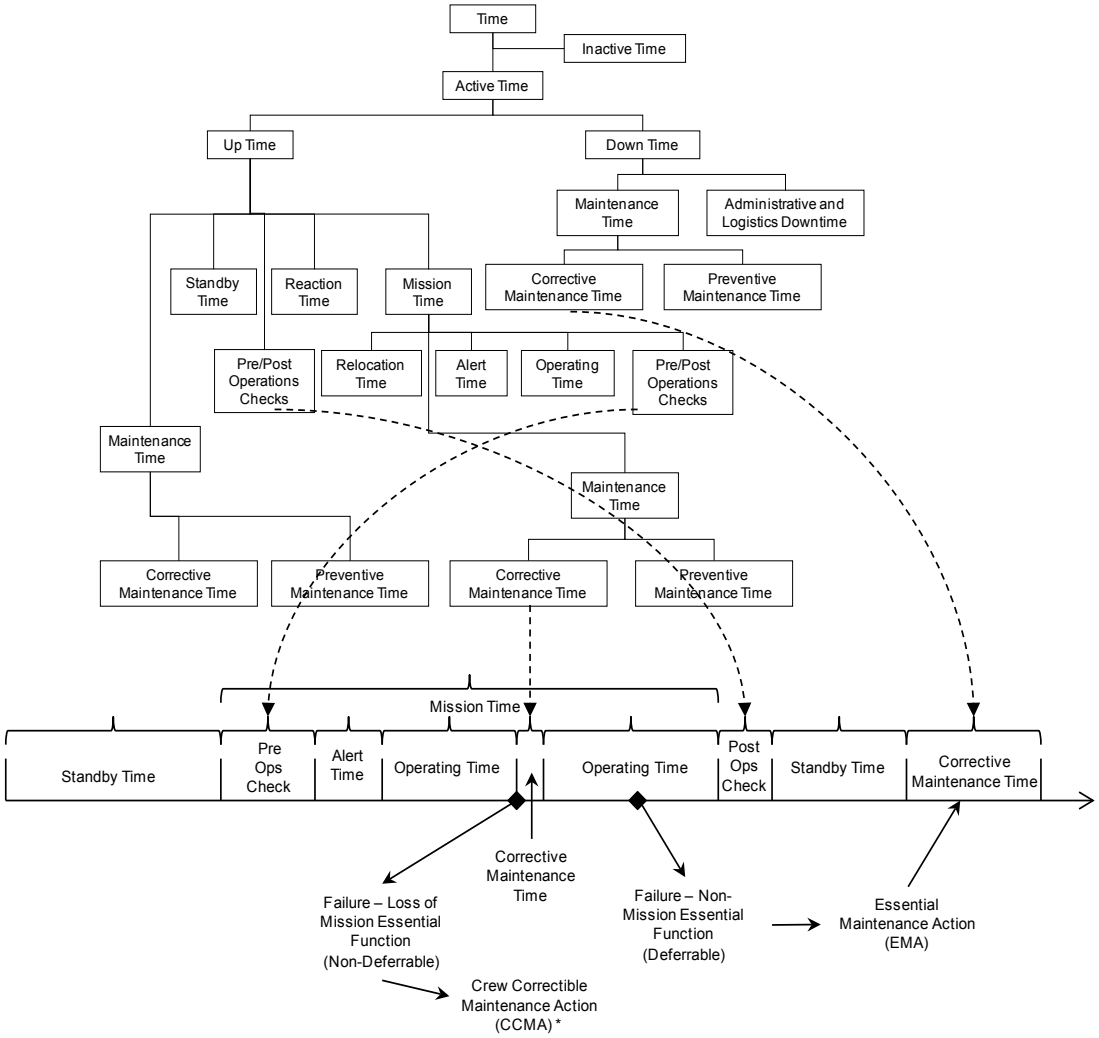


Fig. 2-1. Time Classification Dendrite

In this example the arrival of the failure causes the loss of a mef. Upon arrival of the failure, the crew immediately begins crew correctable maintenance actions and restores the necessary functionality, thus preventing the mission from becoming a complete failure. The second failure that arrives does not cause the loss of a mef because a redundancy exists that prevents the mission from being a complete failure; therefore, the crew takes no immediate action to repair/restore the loss due to the failure. Ultimately the maintenance action is deferred until the completion of the mission.



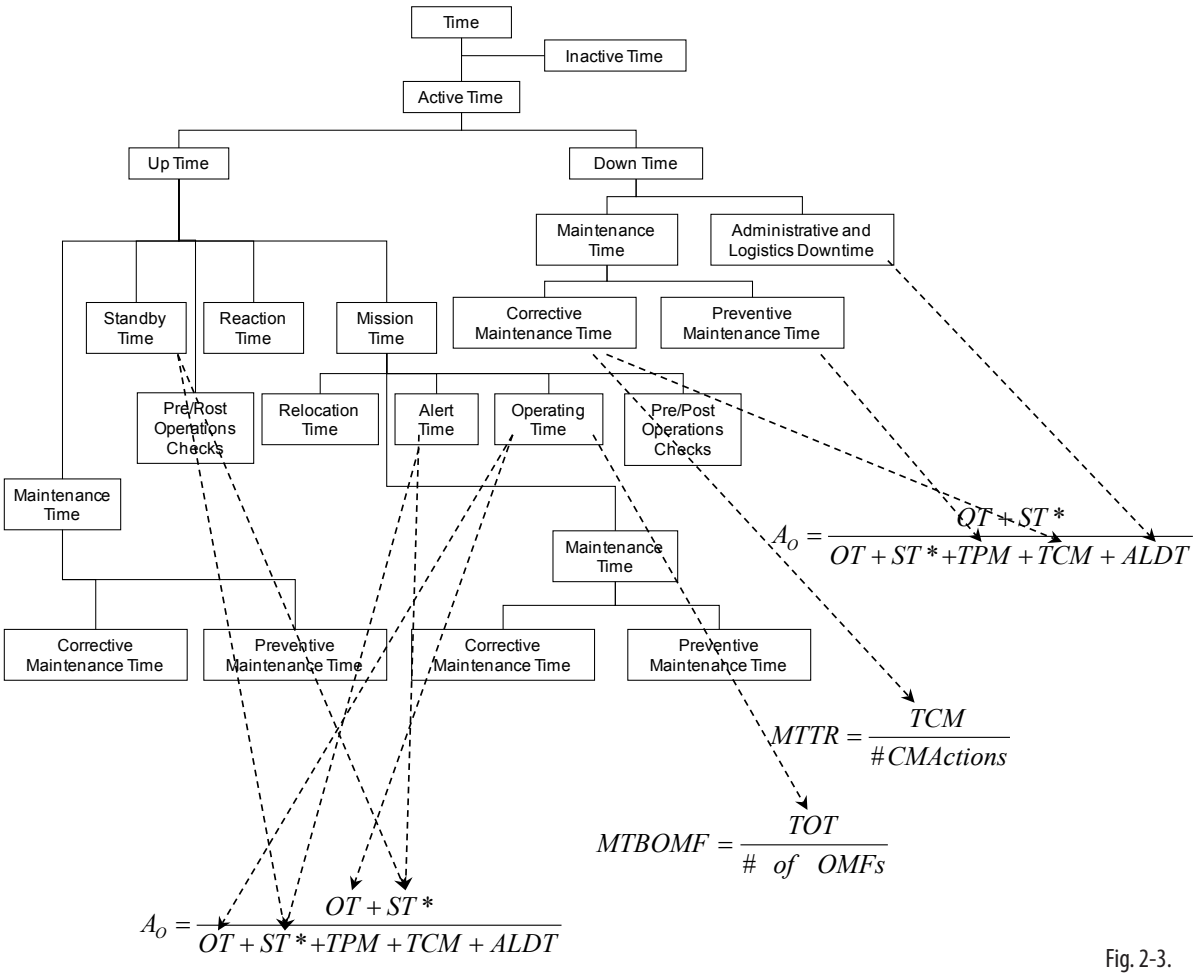
* Fixed by the crew using onboard tools, equipment, and spares within the specified time limit.

Fig. 2-2. Sample System Timeline

In this example, the failure is considered an EMA. Therefore, the restoration of the functionality must take place prior to the start of the next mission. The time spent restoring the functionality is considered as part of downtime under CM.

Figure 2-3 illustrates the links between time categories and RAM metrics.

As illustrated in the figure specific time categories like operating time, standby time, alert time, and downtimes associated with preventive and corrective maintenance, administration, and logistics feed directly into the equations.



* Alert Time is a special case of Standby Time. In Alert Time a system is committed to a specific mission, considered operable, but not actually operating during that time period.

Fig. 2-3.
Links Between Time Categories and RAM Metrics

Operational Mode Summary and Mission Profile

The combat developer must articulate the mix of ways the system performs its operational role in an Operational Mode Summary and Mission Profile (OMS/MP). An integral part of the analysis is the determination of the frequency of task performance, the conditions under which they are performed, and the standards which constitute acceptable performance. This description of tasks, frequency, conditions, and standards forms the basis for the OMS/MP.

Operational Mode Summary

The OMS describes the relative frequency of the various missions, which systems will be involved in those missions, and the types of environmental conditions to which the system will be exposed during the system life cycle (DOD 2009). The contents to look for in an OMS are as follows:

- General statement of broad missions that the equipment will be expected to perform on the battlefield.
- Separately addresses both wartime and

peacetime use.

- Addresses special conditions of use, such as high-intensity wartime usage.
- Expected number of occurrences, operating time, and calendar time of each mission or the percentage of the systems involved in each mission.
- Expected breakdown of environmental conditions in which the entire fleet of systems is expected to be used.
- States the total Operating and Alert Time associated with each mission (i.e., time that the system is required to be operable and committed on a specific mission, even if not operating) (TRADOC/AMC 1987).

When elements of the OMS are not present a clarification to the combat developer should be initiated to address the lack of information.

Tables 2-1 and 2-2 illustrate a fictitious example of a typical wartime OMS for the XYZ system. The OMS lists the types of missions and number of missions. In addition, the OMS lists the quantities of operating, alert, and clock time associated with each mission type. The more detailed breakdown of a mission can be found in the Mission Profile. The OMS also describes the operating envelope in terms of environment for the system.

Mission Profile

The Mission Profile describes the tasks, events, durations, frequency, operating conditions, and environment of the system for each phase of a mission (DOD 2009). The Mission Profile also defines a time-phased description of the operational events and environments an item experiences from beginning to end of a specific mission. (TRADOC/AMC 1987). The contents to look for in an MP are as follows:

- Profiles should be based on typical scenario for the system.

Table 2-1. Fictitious OMS for the XYZ System

XYZ Missions	Operating Time (a)	OT+Alert Time (b)	Calendar Time (c)	No. of Missions (d)	Total OT (a) x (d) = (e)	Total OT+AT (b) x (d) = (f)	Total CT (c) x (d) = (g)
Covering Force	16 hr	16 hr	18 hr	2	32 hr	32 hr	36 hr
Forward Line of Troops Defense *	68 hr	72 hr	72 hr	10	680 hr	720 hr	720 hr
Deep Strike	16 hr	20 hr	20 hr	1	16 hr	20 hr	20 hr
Counter Attack	25 hr	30 hr	30 hr	2	50 hr	60 hr	60 hr
Total	N/A	N/A	N/A	15	778 hr	832 hr	832 hr
*Detailed breakdown can be found in the example for Mission Profile.							

Table 2-2. Fictitious Environmental OMS for the XYZ System

Climate Environment	– Temperature: Hot 20%, Basic 60%, Cold 15%, Severe 5% – Humidity Range: 15% - 95% – Movement Terrain: Primary 10%, Secondary 35%, X-Country 55%
Weather Environment	– Precipitation Type: Rain, Light Snow
Terrain Conditions	– Soil: Clay, Loam, Sand – Vegetation: Coniferous Forest – Slope: 0% to 10% (over 50% of area)

- State specific amounts of operation (e.g., hours, rounds, miles, and/or cycles) for each mef within the mission.
- Should be consistent with future doctrine and tactics.
- Information should be provided on a timeline, a summarization, or other type of format.
- Environmental conditions for each mission.

When elements of the MP are not present a clarification to the combat developer should be initiated to address the lack of information. Mission Profiles are related to the Operational Task Analysis (OTA) previously mentioned in other sections of the manual. When Mission Profiles exist for a system they should be used as the basis for the OTA.

Tables 2-3 and 2-4 illustrate a fictitious example of a typical wartime mission profile for the XYZ system's defensive mission. The example illustrates the type of information needed for a thorough understanding of the mission profile. Table 2-3 identifies the tasks to be performed during Operating Time, including the number of occurrences, time allotted, and the cumulative time. Table 2-4 identifies the expected environmental conditions. Information resources for populating the environmental table are found in the MIL-STD-810 series and the Universal Naval Task List (MCO 3500.26 series). The details of each table are complementary but not identical. For example, the movement terrain is not the same as the OMS and the MP. The difference can be explained in that the OMS is an aggregate of all mission profiles, each of which varies in movement.

Linking Time-Based Models to Failure Definition/ Scoring Criteria

Table 2-3. XYZ System Defensive Mission Profile

XYZ Defensive Mission Tasks	Number of Occurrences	Operating Time for Each Task	Total Operating Time
Movement	12	30 min	6.0 hr
Set-up and Pre-Ops Checks	12	20 min	4.0 hr
Search and Surveillance	80	30 min	40.0 hr
Target Acquisition	36	15 min	9.0 hr
Track	24	5 min	2.0 hr
Fire (Air)	9	200 Rounds at 100 rds/ min (2 min)	0.3 hr
Fire (Ground)	28	400 Rounds at 50 rds/ min (8 min)	3.7 hr
Tear Down	12	15 min	3.0 hr
Total *	N/A	N/A	68.0 hr

*For the mission, all time that the system is not operating is required as alert time.

Table 2-4. XYZ System Defensive Profile (Environmental)

Climate Environment	– Temperature: Hot 20%, Basic 60%, Cold 15%, Severe 5% – Humidity Range: 15% - 95% – Movement Terrain: Primary 20%, Secondary 30%, X-Country 50%
Weather Environment	– Precipitation Type: Rain, Light Snow
Terrain Conditions	– Soil: Clay, Loam, Sand – Vegetation: Coniferous Forest – Slope: 0% to 10%

Following the development of the CONOPS and OMS/MP, the combat developer must decide what minimal operational tasks the system must be able to perform in order to accomplish its mission, as well as what the associated mefs are in order to identify and classify potential failures. This information is documented in the FD/SC. Refer to chapter 3-2 of Volume I for information about tailoring the FD/SC Charter.

Incident Classification

The first step in the process for a given test incident is to determine classification. In the classification step, the members of the scoring conference determine if an incident is related to Reliability or Maintainability of the equipment as it will be expected to be used in the field environment. See figure 2-4, Test Incident Scoring Flowchart, at the end of this chapter.

No Test

Incidents that are judged not pertinent to RAM parameters are classified as No Test. Incidents classified as No Test include:

- Pre-test Checkout. Any incident observed during the designated burn-in, pre-test inspection, or other pre-test activity is classified as No Test. The Test Plan must specify the length of the burn-in period (the number of miles, rounds, or hours) to permit a determination of when the pre-test period has ended.
- Equipment Modification. Maintenance done to install a hardware kit or to incorporate a redesigned component is classified as No Test. However, if the replaced part was not functioning when it was being replaced, that malfunction will be scored on its own merits. A subsequent malfunction of the installed part will also be scored on its own merits.
- Test-Peculiar Incident. An incident caused by someone not acting as a test player (crew member or maintainer), or by equipment not part of the system being tested is classified as No Test. An example of this is an engineering evaluation and the maintenance done in furtherance of that evaluation. This classification also includes malfunctions to or caused by test instrumentation. However, an incident caused by test-peculiar equipment (equipment used in the test in lieu of the equipment to be fielded) will be scored under its own merits because if the test planners have introduced equipment for the purposes of the test, they have judged it to be an adequate substitute for the equipment to be fielded; hence its failures are to be regarded as representative of the failures of the equipment to be fielded.
- Daily Checks and Services. These are checks and services, performed by the operator (or by the crew, if applicable) using only repair parts and On-Equipment Material (OEM) in accordance with the equipment publication before, during, or after the operation of the equipment. Checks and services that meet these conditions are classified as No Test.
- Test-Directed Abuse. An incident in which the tester directs the deliberate abuse of the system (e.g., a test to over-stress the performance limits of the system), whether called for by the test plan or not. However, damage to the system willfully caused by the operator or maintainer and not directed by the tester will be scored under its own merits.
- Non-RAM Oriented. This is a catch-all term to capture those incidents in which a TIR has been prepared, but which have no bearing on the RAM assessment. Example are suggested improvements; reports on inadequate test procedures; reports on unacceptable replacement parts, provided they were discovered before or during installation; reports on the equipment's consistent inability to meet performance specifications even though no actual malfunction has occurred; suggested human factors improvements; and recommended changes to the system support package.

Crew Correctable Maintenance Action (Optional)

The second step in the classification process is to determine if the incident

was crew correctable. If the incident was a malfunction that was correctable by the crew, within the specified time limits, using only the system's onboard tools, repair parts, and spares, then the incident should be scored as a CCMA.

- If the system and its mission are such that the classification of CCMA is not useful in characterizing RAM of a system, then step 2 may be omitted from the FD/SC.
- During a test incident, there will often be test-peculiar time, i.e., time taken by the test administrators, as distinguished from the test players, for analysis and diagnoses. This test-peculiar time should be excluded from the maintenance times.
- The crew maintenance times are usually excluded from maintainability parameters; however, this should be reviewed to determine applicability.

Operational Mission Failure

The third step in the classification process is to determine if the incident was an OMF. If the incident was a malfunction that caused, or could have caused, the inability to perform one or more mefs, it should be scored as an OMF. In addition, if the incident is a critical or catastrophic hazard to personnel or equipment, it should be scored as an OMF.

- If maintenance is needed to restore the loss of a mef, then the OMF will also be scored as an EMA and an UMA.
- If the malfunction is caused by another, simultaneous malfunction, the latter will be scored an OMF and the former will be regarded as a secondary failure and will not be scored.
- If the malfunction is such that the repair can be deferred and the mission can (safely) be continued, the incident is not scored an OMF. It will be scored on its own merits under succeeding steps.
- If the system has two components or assemblies, one of which is redundant to the other at all times, an OMF is not scored unless both are down at the same time. However, if the redundancy is not full time, a failure of the primary component is generally scored an OMF regardless of the status of the backup item at the time of the incident. Exceptions to this rule can be made on a case-by-case basis if the redundancy is nearly full-time.
- The recurrence of CCMAs within a limited period of time may warrant the classification of a group of incidents as an OMF. For example, "The recurrence of two or more CCMAs within an hour, or four (or more) with an 8-hour mission will be classified as an OMF."
- Critical or catastrophic hazards are defined in MIL-STD-882 series, System Safety Program Requirements.

Essential Maintenance Action

All EMAs are also classified as UMAs. For some systems that lack redundant features and for which the performance is not affected by "special conditions" the classification of EMA can be omitted.

Unscheduled Maintenance Actions

Any incident classified in steps 2-4, or any maintenance that does not qualify as a Scheduled Maintenance Action (SMA). In other, words any maintenance that does not qualify as an SMA the maintenance must be prescribed by an equipment publication; and, there must be enough latitude in the time for the performance of the maintenance that it can be done in a slack period between missions.

Incident Chargeability

The following is a description of each chargeability category.

- Hardware. This category includes not only malperforming hardware but also personnel-related incidents that are attributable to the hardware's design. For example, if the device has an exposed on/off toggle switch that is easily tripped inadvertently, an unintended power down of the equipment may be charged to the hardware vice the crew. Hardware chargeability may be further broken down into Government-furnished hardware and contractor furnished hardware.

- Software. This category applies to contractor and Government-furnished software that malfunctions. Similar to hardware, personnel-related incidents that are attributable to the software design may be charged to the software vice the crew.
- Care should be taken to distinguish between genuine software reliability problems and simply improperly designed software incapable at any time of executing a given task.
- Care should also be taken in defining what software is part of the system under test and what software is peripheral events (associated). Application software is usually treated as "support equipment."
- Crew
- Maintenance Personnel
- Manuals. These are incidents that are attributable to misleading, incorrect, or nonexistent, but needed, information.
- Training. These are incidents that are attributable to misleading, incorrect, or nonexistent, but needed, information.
- Support Equipment. These are incidents caused by special and common tools and Test Measurement Diagnostic Equipment, spares, repair parts, the associated software, and sometimes power sources.
- Accidents. This category includes only those accidents that are not occasioned by the design of the system. Incidents that should not be accounted for as accidents are those due to inadequate training, inadequate warning in the manual, and/or careless operation. These would be captured under manuals, training, or crew.
- Unknown. These are incidents that cannot be charged to one of the above categories. This category is sometimes helpful in the characterization of communications networks in which there are "spontaneous remissions" of a malfunction. The unknown category has the potential for misuse, therefore it should be used as a last resort in chargeability (TRADOC/AMC 1987).

Hazard Severity Assessment

The hazard severity categories are as follows:

- Catastrophic (I): Could result in death, permanent total disability, loss exceeding $1M, or irreversible severe environmental damage that violates law or regulation.
- Critical (II): Could result in permanent partial disability, injuries, or occupational illness that may result in hospitalization of at least three personnel, loss exceeding $200K but less than $1M, or reversible environmental damage causing a violation of law or regulation.
- Marginal (III): Could result in injury or occupational illness resulting in one or more lost work days, loss exceeding $10K but less than $200K, or mitigatible environmental damage without violation of law or regulation where restoration activities can be accomplished.
- Negligible (IV): Could result in injury or illness not resulting in a lost work day, loss exceeding $2K but less than $10K, or minimal environmental damage not violating law or regulation (DOD 2000).

Hazard Probability Assessment

The hazard probability categories are as follows:

- Frequent (A): Likely to occur often in the life of an item, with a probability of occurrence greater than 0.1 in that life.
- Probable (B): Will occur several times in the life of an item, with a probability of occurrence less than 0.1 but greater than 0.01 in that life.
- Occasional (C): Likely to occur some time in the life of an item, with a probability of occurrence less than 0.01 but greater than 0.001 in that life.
- Remote (D): Unlikely but possible to occur in the life of an item, with a probability of occurrence less than 0.001 but greater than 0.000001 in that life.
- Improbable (E): So unlikely, it can be assumed occurrence may not be experienced, with a probability of occurrence less than 0.000001 in that life (DOD 2000).

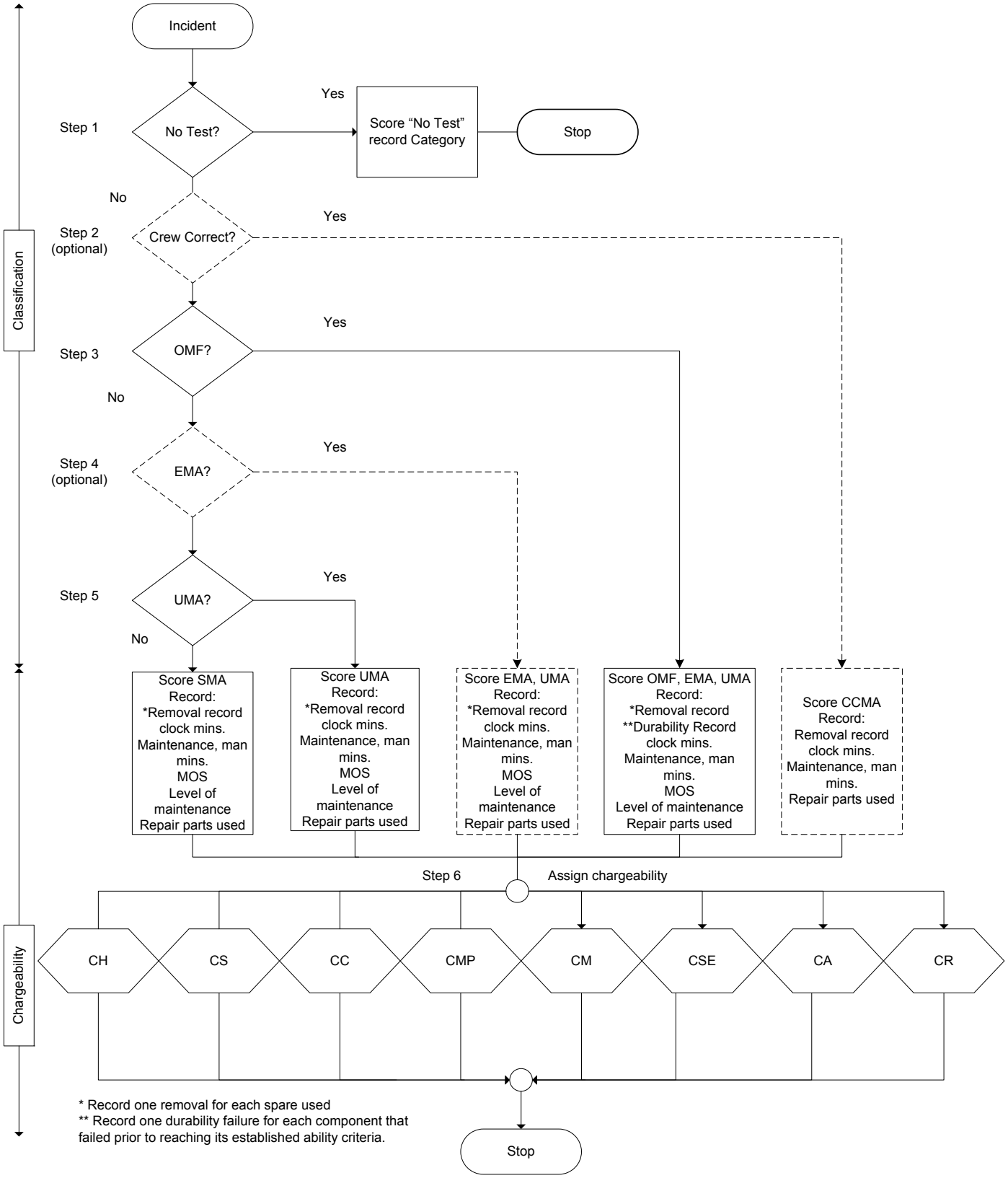


Fig. 2-4.
Test Incident Scoring Flowchart

3 M&S and VV&A

Chapter VII-3. Modeling and Simulation and the Verification, Validation, and Accreditation Process

This chapter provides an introduction to the topic of Modeling and Simulation (M&S), MCOTEA's purpose in using M&S, and the process for its verification, validation, and accreditation (VV&A). Figure 3-1 provides an overview of definitions the reader will need for this chapter.

Figure 3-1 . Quick-look Definitions

Model	Any physical, mathematical, or otherwise logical representation of a system, entity, phenomenon, or process.
Simulation	Imitates the operation of a real-world process over time. Composed of one or more models and governed by a set of assumptions about system operation.
Simulator	Device used to artificially duplicate real world conditions. Typically used as a training device.
Stimulator	Causes real-world response to simulated inputs.
M&S Developer	Individual, group, or organization that develops or modifies an M&S in accordance with design requirements and specifications. Can also be responsible for executing the Configuration Management Plan and the V&V Plan.
M&S Proponent	Organization that ensures that the M&S satisfies the requirements, develops the V&V Plan, performs V&V, develops reports, ensures CM, and ensures sufficient information for M&S accreditation. If MCOTEA resources VV&A, MCOTEA is the proponent. For M&S MCOTEA uses for OT&E, the Program Office is generally the proponent, supported by the SCP.
Simulation Control Panel	Provides independent technical expertise. Helps the Operational Test (OT) Accreditation Agent (ACA) and Developmental Test (DT) ACA understand model functionality; ensures that the M&S Developer delivers the intended M&S capabilities; and assists in V&V activities.
M&S User	Individual, group, or organization that uses the results or products from a specific application of M&S. For all uses of M&S and associated data in MCOTEA test and evaluation, MCOTEA is the M&S User.
Subject Matter Expert (SME)	An individual who, by virtue of education, training, or experience, has expertise in a particular technical or operational discipline, system, process, or M&S.
Specific Intended Uses (SIU)	The SIUs are a statement of how the test team expects to use M&S results in support of the MCOTEA evaluation; they represent M&S support essential to the MCOTEA OT&E.
Verification	Process of determining that a model or simulation and their data accurately represent the developer's description and specifications.
Validation	Process of determining the degree to which a model or simulation and their data accurately represent the real world based on the model's intended uses.
V&V Agent	Individual, group, or organization that performs the verification, validation, or both. M&S Proponent designates. If the M&S Developer functions as the Agent, an independent entity (such as the SCP) should check the work.
Accreditation	Official certification that a model or simulation and their data are acceptable for a specific purpose.
Accreditation Authority (AA)	Individual or organization responsible for approving use of a model, simulation, or federation of such for a particular application. MCOTEA's AA is the Director or a designee.
Accreditation Agent (ACA)	Individual, group, or organization that conducts an accreditation assessment for an M&S application. At MCOTEA, the OTAD designates the ACA.
Accreditation Criteria	Accreditation criteria are the set of standards that must be met in order for the M&S to be accredited for a particular use.

Deciding to Use M&S

When conducting operational testing and evaluation, it is often useful to consider supplementing the data obtained during live testing with data supported/generated by M&S. From MCOTEA's perspective, M&S may offer an opportunity to examine the system under test in ways that are operationally useful, but not otherwise feasible during an operational test. For example, firing an anti-radiation missile (ARM) at a radar to understand its ability to detect a threat might not be practical during operational testing, but if an M&S could simulate an ARM threat and stimulate the radar satisfactorily, the test team could evaluate the radar's ability to detect the threat as well as the radar operators' reaction to the threat.

MCOTEA uses M&S to augment, but not replace, operational testing. M&S can be used to

- Help design tests
- Predict what happens during tests
- Provide stimulation during tests
- Use test data as input to examine outcomes that cannot be directly tested
- Generate data to supplement data generated during tests

Particular reasons to use M&S include the lack of test asset availability, lack of sufficient time to generate adequate data, test range limitations, cost, and safety considerations.

Public law restricts the use of M&S in OT&E such that the results of the operational evaluation cannot be "… based exclusively on computer modeling; simulation; or an analysis of system requirements, engineering proposals, design specifications, or other information contained in program documents" (Title 10 USC 2399). In addition,"M&S shall not replace the need for OT&E and will not be the primary evaluation methodology" (SNI 5000.2, section 5.4.7.9).

Within these constraints, M&S can provide a powerful way for MCOTEA to supplement the information derived from operational testing.

The decision to use M&S occurs early in the OT&E planning stages, before the SEP is finalized and in conjunction with the Program Office and the T&E WIPT's development of the Test and Evaluation Strategy (TES). The overall T&E approach should define the types of data that can be expected from each potential test venue. Real test data is generally preferable to M&S data; however, any data required by the overall evaluation that cannot reasonably be gathered during a test event is a candidate for M&S support. MCOTEA does not generally develop models or simulations. Sources for obtaining appropriate M&S are listed later in this section.

"Models include any physical, mathematical, or otherwise logical representation of a system, entity, phenomenon, or process. Simulations include a method of implementing a model over time" (SNI 5200.38A).

Expanded Definitions of M&S

Physical models used in operational testing include tank hulls, armor plating, humans, and buildings. MCOTEA once used a Rocket Propelled Grenade (RPG) in an operational test to replicate the appropriate visual signature of an enemy role player. This physical model had all of the visual features of an RPG, but like all models it was not without limitations: the RPG could not be fired and its weight was not representative of a real RPG. Regardless of these limitations, the physical model was very useful in the operational test because the limitations had no real effect on how the model was used in the test.

Mathematical models represent aspects of a system. For example, the Reliability of a component may be modeled using equation $R(t)=e^{-t/MTBF}$, where R(t) is the probability of operating for a time t without a failure and MTBF is Mean Time Between Failures for the component.

A simulation imitates the operation of a real-world process or system over time. The simulation generates an artificial history for a system. The data composing the artificial history is then used to draw inferences about the behavior and characteristics of the real system. Simulations are composed of one or more models and are governed by a set of assumptions, inherent in the models and the way they are applied, concerning the system's operation. For example, if the mathematical model in the equation above were used in a simulation, the simulation would assume that the probability of system failures can be modeled using an exponential distribution. The models composing a simulation are often used through a computer program, thus giving rise to a computer model.

A simulator is a special case of a simulation. A simulator is a device used to artificially duplicate real world conditions so that its operators can practice reacting to those conditions. The simulator represents an actual operational system to varying degrees of fidelity and requires the operators to interact with the simulator much like they would interact with the real system.

A stimulator causes a real-world response to simulated inputs in a system or causes a corresponding real-world reaction by an operator. Stimulators are generally associated with test execution and are used to enable the examination of operationally relevant responses and reactions that otherwise could not be tested. In this role, stimulators are a cost-effective means to test operational aspects of a system the testing of which would otherwise be cost prohibitive, unrealistic, or hazardous.

Although this section defines several types of models, MCOTEA will generally be concerned with computer models. Therefore, this chapter is written assuming the M&S is a computer program/model.

Using M&S in the Test Process

MCOTEA uses M&S to minimize risk by leveraging test venues and scenarios in a way that maximizes the information obtained in test planning, execution, and evaluation.

During Test Planning, M&S can

- Help develop scenario and test setup
- Predict outcomes before testing occurs

During Test Execution, M&S can

- Stress systems under test with large numbers and higher densities than feasible during actual testing
- Present test situations that could not safely or practically be done during actual testing
- Present enemy threats, systems, or counter measures not otherwise available during actual testing

During Evaluation, M&S can

- Examine alternative environments and conditions
- Examine the implications of system deficiencies and test limitations
- Apply test data to conditions, subjects, and scenarios that cannot otherwise be safely or practically tested

Requirement for VV&A

DOD policy states that all models, simulations, and associated data used to support DOD processes, products, and decisions undergo verification and validation throughout their lifecycles (DODI 5000.61). In general MCOTEA does not perform V&V; MCOTEA's responsibility is to accredit the M&S. In order to be used in MCOTEA test and evaluation, all M&S must undergo accreditation. MCOTEA may use M&S in some of the following ways:

- Test assets
- Test stimulators
- Test planning aids
- Pre/posttest analysis tools
- The tools' input data
- The tools' produced data

Once the test team decides that M&S support is needed, the MCOTEA Operations Test Analysis Division (OTAD) appoints an Accreditation Agent (ACA) to the team. The ACA is responsible for organizing MCOTEA's accreditation assessment and obtaining all information necessary to support an accreditation decision.

Developing Specific Intended Uses and Accreditation Criteria

The first task undertaken by the ACA in conjunction with the test team is to determine the Specific Intended Uses (SIU), statements of how the test team expects to use M&S results in support of the MCOTEA evaluation; they represent M&S support essential to the MCOTEA OT&E. After identifying the data that must come from M&S, the test team can define the SIUs.

SIUs generally begin as high-level requirements that become more detailed as the test program develops. The earlier and more precisely the SIUs can be stated, the better. In the end, if the M&S cannot be accredited for certain SIUs, the unachieved SIUs will generally represent limitations to the OT&E.

An effective SIU provides detailed information to the M&S developers. This clarity in expectations helps the developers deliver what is needed; they know the software development goal at an early stage.

This clarity in purpose also helps the ACA establish the accreditation criteria, used to ascertain whether the models are able to deliver the SIU with sufficient fidelity. Accreditation criteria are the set of standards that must be met in order for the M&S to be accredited for a particular use. The criteria should include quantitative standards to the maximum extent possible and should be revisited periodically to ensure that they remain appropriate and sufficient for the application. (Typically, the acceptable Verification of the M&S is also an accreditation criterion, but it is not associated with an SIU.)

Locating the Right M&S

Once the required SIUs are determined, the MCOTEA test team in conjunction with the Program Office select a source for the M&S support. When M&S support is required for both developmental and operational testing, it might be reasonable to use the same M&S to support both. In any case, the required M&S support can be obtained either by using existing models or generating new ones.

Existing Models

The following sources are useful when an existing or generic model can satisfy an SIU:

- DOD M&S Catalog (https://MSCatalog.osd.mil), based on DODI 5000.61.
- SURVIAC—the DOD institution responsible for collecting, analyzing, and disseminating information related to all aspects of survivability and lethality for aircraft, ground vehicles, ships, and spacecraft. SURVIAC maintains several approved models. www.bahdayton.com/surviac/
- Human Effects Center of Excellence (HECOE)— maintains several models designed to address the effects of certain stimuli on humans under various conditions. HECOE is located at Brooks City-Base, San Antonio, TX.

The test team must research their particular area of interest when searching for useful existing models. The team must also consider the following points:

- The accreditation process for an existing model should leverage previous V&V efforts to the greatest possible extent, but the level of vigor with which existing models have been verified and validated varies widely.
- An existing model might save development time and money, but MCOTEA must

An ineffective SIU might read: "The M&S will be used to stimulate the radar by producing performance data under a wide range of scenarios."

An effective SIU might read: "The M&S will be used to stimulate the radar by modeling the radar cross section, delivery profile, and kinematic performance of a threat anti-radiation missile and overlay this data onto the data stream containing the returns of the actual radar during the live play of operational testing."

still accredit it for SIUs using appropriate accreditation criteria.

New/Modified Models

Creating new M&S may best be done by the system developer if the M&S will be used to predict system performance or to stimulate the system, and if the system developer is capable of creating the M&S. The system developer knows the system capabilities and is probably best positioned to convey this knowledge to an M&S. However, the test team must be aware that the system developer will be motivated to make their M&S as well as their system look good. Therefore, the test team must be vigilant in monitoring the M&S's development, verification, validation, and quality assurance.

The test team should revisit the Simulation Interface Units (SIU) after deciding on a model because

- Increased familiarity with test concepts may inspire the need for different/additional SIUs
- The test team might want to take advantage of the supplemental capabilities of a previously unidentified existing model

The VV&A process is rooted in confidence: "Confidence in a particular model or simulation must be justified before its results are used to make decisions involving large sums of money or risk to human life" (SNI 5200.40).

All models using stochastic processes should also provide guidance on the number of iterations necessary considering runtime, confidence levels, and output stability.

M&S Funding and Schedule Requirements

Once the test team understands the required level of M&S support for OT&E, the OTPO notifies the Program Office of MCOTEA's requirements and plans so that sufficient funding for model development and V&V efforts is allotted. This funded support must include the technical expertise required to develop, manage, operate, verify, validate, accredit, and apply the results of the M&S application. In addition, the funded support should include the SCP (explained later in this chapter) and any supplemental, independent SME support required to ensure that final VV&A requirements are met. The OTPO and ACA must assess the adequacy and technical soundness of the PO's approach to satisfying MCOTEA M&S and VV&A requirements.

The program's schedule must include a realistic amount of time for the required level of VV&A, at a minimum several months. The OTPO and the MCOTEA Test Manager must be kept apprised of the M&S schedule for overall test planning purposes. If the M&S will be new, realistic time must also be provided for model development. Constraints on schedule and cost exist for any program, but the MCOTEA SIUs represent data essential for overall system evaluation. To the extent that funding or schedule are not adequately provided, any neglected SIUs become limitations to the OT&E.

Verifying, Validating, and Accrediting M&S

Although MCOTEA's role in the VV&A process is to provide the accreditation, generally not the V&V, being familiar with the entire VV&A process allows MCOTEA to understand the M&S's capabilities, reduce the risk associated with using the M&S, and make informed decisions about using an M&S in support of OT&E.

Furthermore, MCOTEA is closely involved in helping to determine V&V activities and in witnessing the results, reinforcing the idea that understanding the complete process is essential for appropriate MCOTEA participation. This section focuses on MCOTEA's accreditation responsibilities, supported by essential information about the V&V process. Annexes A and B contain detailed information about V&V.

- MCOTEA must be confident that a particular M&S is accurate and suitable for the SIUs
- This confidence must be based on an unbiased assessment of the M&S
- The justification of this confidence must be communicated for future use using established reporting mechanisms as described in this chapter (also see SNI 5200.40).

The term VV&A is often used in the context of a single activity, but the VV&A process

contains three distinct and separable activities (DODI 5000.61):

- Verification: The process of determining that a model or simulation implementation and its associated data accurately represent the developer's conceptual description and specifications.
- Validation: The process of determining the degree to which a model or simulation and its associated data are an accurate representation of the real world from the perspective of the intended uses of the model.
- Accreditation: The official certification that a model or simulation and its associated data are acceptable for use for a specific purpose.

General Uses of M&S Requiring Accreditation

MCOTEA must accredit an M&S when the M&S is used to support operational testing or if the data will be used in the evaluation of OE, OS, OSur. For M&S used in DT (i.e., MCOTEA involvement is only to determine if the threshold has been met), then the TEMP will identify the agencies responsible for accreditation.

MCOTEA accreditation applies to all categories of models, simulations, and data, including:

- Live, virtual, and constructive simulations
- Unitary, federated, or distributed simulations
- Commercial off-the-shelf/Government off-the-shelf/Non-developmental Item software or hardware, emulators, and prototypes
- Simulators
- Stimulators
- Data needed to verify M&S requirements, validate the M&S, perform experiments on/with the M&S, or run combat M&S decision aids

Ideally, verification is an integral part of model development, meaning that the model verification techniques are identified a priori and execution of these techniques is documented as they are performed during model development. However, if verification of legacy M&S was not accomplished, was inadequate, or was not documented, MCOTEA may need to require additional verification.

Note: Many organizations (including for-profit contactors, not-for-profit contractors, universities, and various government organizations and labs) develop models that might supplement a MCOTEA OT&E. The concepts, techniques, procedures, and standards MCOTEA uses to accredit an M&S apply to all M&S developers, regardless of their pedigree.

VV&A Stakeholders and Their Roles

Overall VV&A involves the following stakeholders:

- M&S User
- Accreditation Authority
- Accreditation Agent
- M&S Proponent
- V&V Agent
- M&S Developer
- SME

The MCOTEA stakeholders in the list are the M&S User (for M&S used in OT&E), the Accreditation Authority (AA), and the ACA. It is also possible for MCOTEA to be the M&S Proponent, if MCOTEA funds the M&S, but funding typically occurs through the Program Office.

Accreditation Authority

Generally speaking, the AA is the organization or individual responsible for approving the use of a model, simulation, or federation of models and simulations for a particular application. The AA ensures that resources are available for the VV&A effort. For all uses of M&S and associated data in MCOTEA tests or evaluations, the AA is the Director, MCOTEA or a designated representative. The AA's chief responsibilities are as follows:

- Determine the appropriateness of an M&S

Software applications requiring MCOTEA accreditation do *not* include the software that is an integral part of the system under test, including firmware and other software required to drive the system. MCOTEA expects this type of software to be verified along with the verification of other system specifications during developmental testing.

for the required SIUs

- Ensure that adequate V&V has occurred before using the M&S in OT&E
- Require Independent Verification and Validation of the M&S if needed
- Sign the Accreditation Decision Letter based on the Accreditation Report (documentation is explained in detail later in this chapter)

Accreditation Agent

Designated by the MCOTEA OTAD, the fundamental job of the ACA is to gather the information the AA needs to make an informed accreditation decision. The ACA fulfills an essential role and should function as an adjunct to the test team, without other team responsibilities. (If M&S will be used to support DT, the PMO will assign its own Accreditation Agent to the program.) The OTAD can designate an independent organization or assign a staff member or support contractor as MCOTEA's Accreditation Agent; ideally, the ACA will have experience in V&V to help provide an understanding of the entire VV&A process. The ACA's chief responsibilities are as follows:

- Ensure, early in the VV&A process, that the M&S Developer and M&S Proponent are familiar with MCOTEA V&V requirements.
- Represent the test team at all internal and external discussions of M&S in support of OT&E.
- Lead the test team discussion to determine the desired SIUs. Leading the discussion gives the ACA a deep understanding of the T&E strategy, the role each SIU is intended to perform in the system evaluation, and the corresponding importance of each SIU.
- Lead the MCOTEA effort to determine the appropriate M&S for OT&E
- Conduct the research to support which M&S to use
- Provide the research results to the OTPO for selecting the best M&S options
- Function as the resident MCOTEA expert on the M&S being considered for accreditation
- Become familiar with M&S assumptions, capabilities, limitations, and history
- Monitor the resolution of technical issues to ensure that the M&S will be capable of executing the MCOTEA SIUs
- Participate in the SCP
- Prepare an Accreditation Plan and Accreditation Report

In addition, the ACA coordinates all required MCOTEA participation in inspections, analyses, demonstrations, and tests in support of M&S accuracy and requirements verification; M&S capability validation; supporting data V&V; and validation that the M&S satisfies the accreditation criteria of all MCOTEA SIUs.

Other Important V&V Roles

The following section defines other, non-MCOTEA stakeholders and their roles in the V&V process.

M&S Developer. The individual, group, or organization responsible for developing or modifying an M&S in accordance with a set of design requirements and specifications. The M&S Developer can also be responsible for executing the Configuration Management (CM) Plan and the V&V Plan.

M&S Proponent. The organization with primary responsibility for

- Ensuring that the M&S satisfies the stated requirements
- Developing the V&V Plan and Report
- Performing V&V activities
- Ensuring effective CM
- Ensuring that sufficient information is gathered to support the M&S accreditation

It is not unusual for the M&S Proponent to contract with the M&S Developer

to generate drafts of the CM Plan, the V&V Plan, and the V&V Report for independent review and approval. The M&S Proponent is responsible for delivering this documentation to the DOD M&S Catalog (https://MSCatalog.osd.mil).

M&S User: The individual, group, or organization that uses the results or products from a specific application of an M&S. For all uses of M&S and associated data in MCOTEA tests or evaluations, MCOTEA is the M&S User.

Subject Matter Expert: An individual who, by virtue of education, training, or experience, has expertise in a particular technical or operational discipline, system, process, or M&S. SMEs are on the SCP primarily to help the MCOTEA ACA and the DT ACA gain an understanding of model functionality, ensure the M&S Developer delivers the intended M&S capabilities, and assist in verification and validation activities.

Verification and/or Validation Agent: The individual, group, or organization designated by the M&S Proponent to perform the verification, validation, or both, of a model, simulation, or federation of models and simulations, and their associated data. If this is the M&S Developer, the V&V should be checked by an independent entity such as the SCP.

Further detailed information about stakeholders' individual roles and responsibilities can be found in the M&S VV&A Implementation Handbook.

VV&A Process in Total

The following two flow charts present the basic VV&A (focusing on accreditation) process, beginning with seven steps generic to any accreditation once the test team decides to use M&S. Figure 3-2 (next page) traces the process of accrediting a new or modified model. Figure 3-3 (next page) depicts the accreditation process for an existing model. Although MCOTEA does not generally perform V&V, the ACA supports the determination of acceptable V&V, and thus must be familiar with the V&V activities and processes.

The Accreditation Process

The accreditation process applies to a specific version of an M&S, a predetermined set of SIUs, and everything that supports the M&S to be used in OT&E. The process can be organized into six parts:

1. Assessing the M&S Developer's experience and processes

2. Assessing M&S functionality and assumptions

3. Verifying and validating M&S input data

4. Verifying requirements satisfaction and model accuracy

5. Validating replication of the real world

6. Ensuring that accreditation criteria for each SIU are satisfied

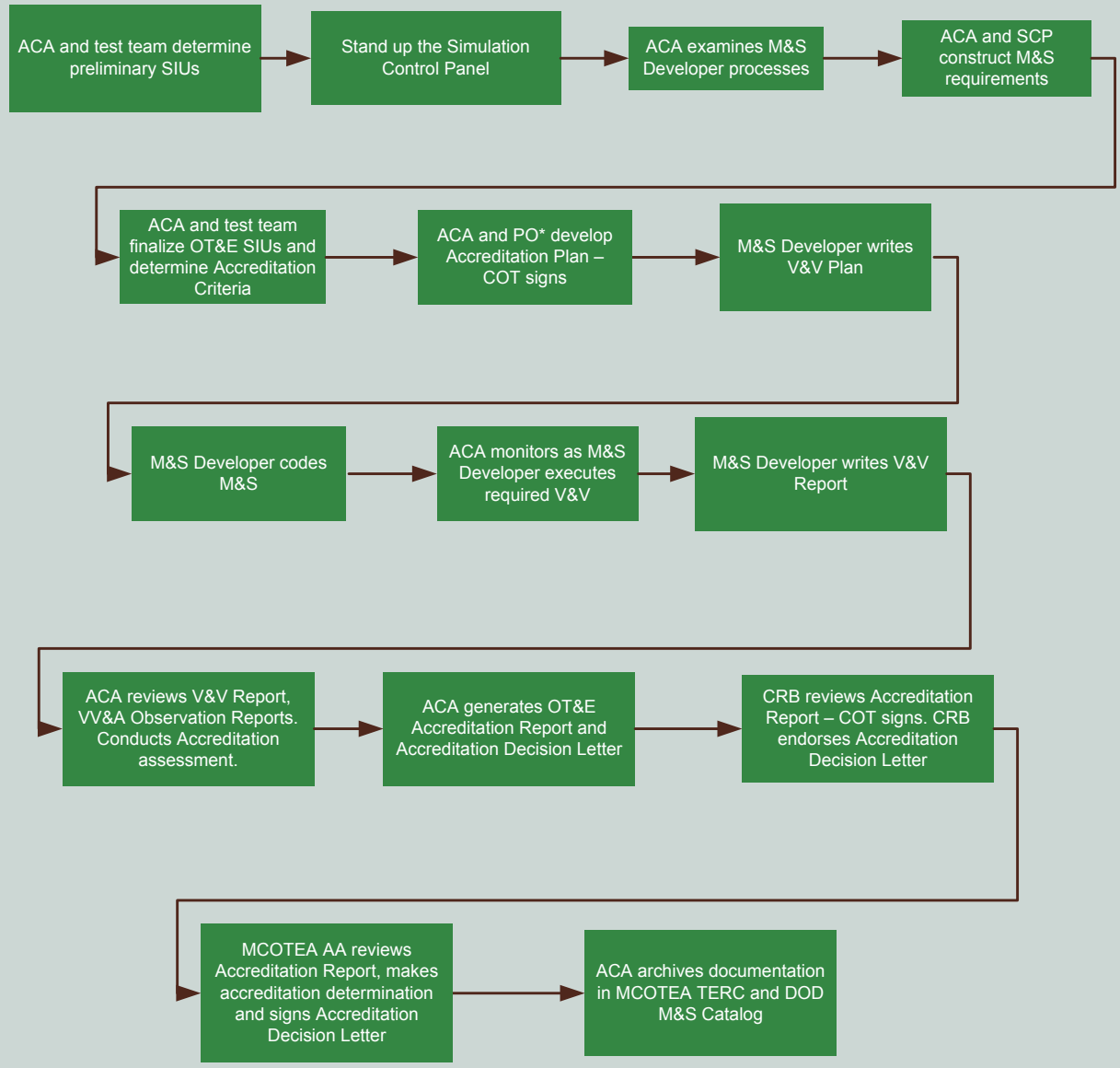
The results of the accreditation effort establishes the level of credibility MCOTEA bestows on the M&S with respect to the SIUs in question. The assessment may or may not result in MCOTEA M&S accreditation for all SIUs.

Accreditation Schedule

In all cases, the version of an M&S intended to support OT&E must be locked down (no further changes), with SIUs fully accredited, 30 days before the OTRB (120 days before OT). If certain SIUs cannot be fully accredited by this time, the test team must determine alternative ways to satisfy the requirements of the unaccredited SIUs or plan to report them as Test Limitations. The test team must brief alternative plans for each unaccredited SIU at the OTRB.

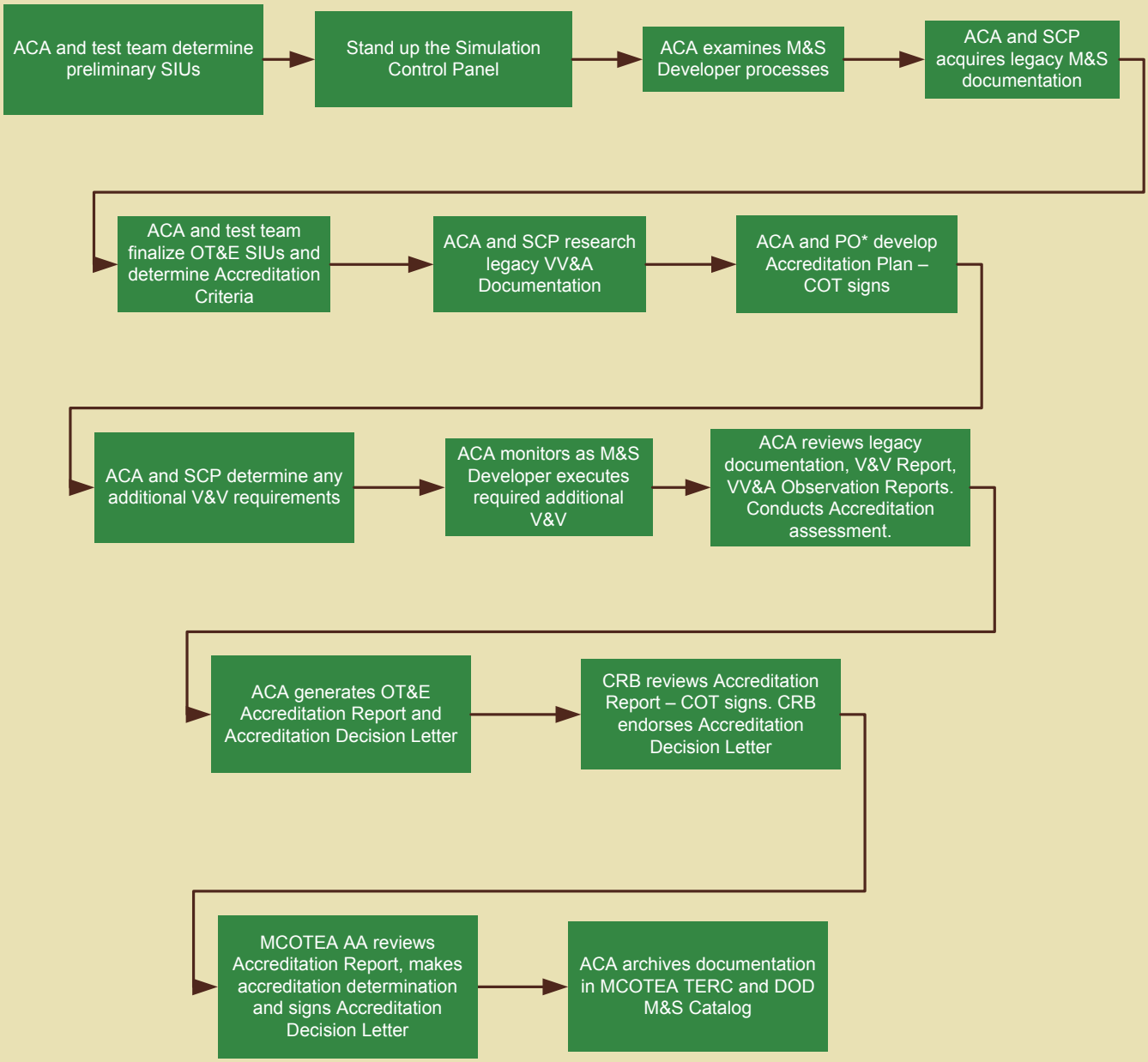
Overview of the Accreditation Assessment

A MCOTEA accreditation assessment requires much more than simply



* Assumes joint MCOTEA/PO Accreditation Plan

Fig. 3-2. Accrediting a New or Modified Model



* Assumes joint MCOTEA/PO Accreditation Plan

Fig. 3-3. Accrediting an Existing Model

Simulation Control Panel

The primary purpose of the SCP is to help the OT ACA and the DT ACA understand model assumptions and functionality, ensure the M&S Developer delivers the intended M&S capabilities, and assist in verification and validation activities.

If the PMO funds the accreditation effort, the PMO charters the SCP in collaboration with the MCOTEA ACA. In this case, the SCP is composed of the MCOTEA ACA, the PO Accreditation Agent representing DT SIUs, the V&V Agent, PO SMEs, the model developer, and independent government and contractor technical experts as determined by the PO and MCOTEA. The PO appoints the SCP chair.

If MCOTEA funds the accreditation effort, MCOTEA charters the SCP with the same membership. In principle, each M&S requires its own SCP. In practice, if several different but related M&Ss are needed for a program, it is often convenient to include the necessary SMEs in a single SCP. The SMEs can be domain professionals (such as experts in what is being modeled by the M&S, e.g., doctors and system experts as required.)

The SCP meets periodically to review and approve model methodologies, use of algorithms, model assumptions, accuracy of approach, adequacy and applicability of input data, model developer processes, and documentation. Based on this activity, the SCP provides periodic reports of model status, plans, and schedule to the DT and OT ACAs. The chairman writes (or assigns) minutes for every meeting and appoints others to write on special topics as needed.

The chief responsibilities of the SCP are as follows:

•Serve as a communication conduit between the ACAs and the M&S developer.

•Provide independent expertise to help address important technical issues and assist the ACAs in gathering relevant technical information.

•Probe the operating details of the model to understand model assumptions and methodologies.

•Provide guidelines for data V&V

§ Why and how the data was generated (the more detail the better)

§ Any assumptions made in generating the data.

•Provide guidelines for the V&V Plan

§ Outline and schedule

§ Any needed clarification of accreditation requirements

•Review and approve final V&V Plan

•Possibly require V&V techniques in addition to those found in the reference (DOD VV&A RPG, www.vva.msco.mil, Reference Documents, V&V Techniques) during V&V Plan execution

•Provide guidelines for the V&V Report

§ Outline and schedule

§ Any needed answers about content and distribution

•Review and approve V&V Report

•Deliver the V&V Report to DT and OT ACAs

Once the accreditation decisions are made, the SCP is dissolved.

reading reports. It requires the active participation of the ACA as described in this section. It is important that the accreditation activities begin early in a program so the M&S Developer is aware of MCOTEA V&V requirements and can react to them (fig. 4-4). The accreditation agent needs first-hand knowledge of the M&S V&V activities and techniques, the assumptions used by the M&S, and an understanding of the strengths and weakness of the M&S relative to the OT&E SIUs. This requires the early and active participation of the ACA in V&V activities. This section describes the types of activities the MCOTEA ACA is expected to accomplish to perform an accreditation assessment.

1. Assessing the M&S Developer's Experience and Processes

Confidence lies at the root of all actions associated with verifying, validating, and accrediting an M&S.

To generate the required confidence, the ACA must first gather general information about the M&S Developer's software development history and processes. This information will indicate the degree to which the M&S Developer follows established software development and software quality assurance procedures, leading to ACA awareness of how closely the M&S Developer's work must be scrutinized. The following issues are examples of general information that can indicate the M&S Developer's competence. These issues are not all inclusive, and the OT&E team/ACA are encouraged to ask additional questions that may apply more closely to their specific circumstances. Positive answers to all of the following issues will create confidence in the developer's ability to construct a quality M&S. If the developer cannot satisfactorily address one or more of the following issues, additional scrutiny by MCOTEA may be warranted.

Historical error detection efficiency. If the M&S Developer has this information, it is probably valid for some amount of time after initial release. Error detection efficiency can be calculated as (number of errors detected before release of software)/(number of errors detected before release + number of errors detected after release). The higher the result the better. The error detection efficiency for some organizations exceeds 99 percent. In general, a rate below 90 percent indicates that the M&S warrants additional scrutiny for errors.

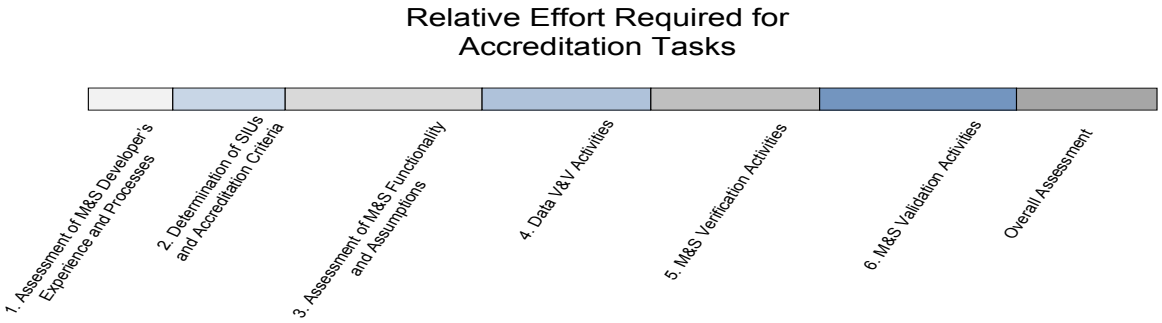


Figure 3-4. The major categories of accreditation tasks where the horizontal width of each layer represents the relative amount of effort required to accomplish the given task. This figure indicates that the "Assessment of M&S functionality and assumptions" can be expected to take roughly as much effort as "M&S validation activities." Both of these tasks can be expected to take roughly four times the effort of the "Assessment of M&S Developer's Experience and Processes" tasks. This figure is intended to be used as a tool for planning the accreditation.

Programming languages used in developing the M&S and number of software lines of code (SLOC) required during programming. Estimate SLOC for codes under construction. The SLOC can be used to indicate the level of the code's complexity. A model comprising 10,000 lines of code can be expected to harbor more errors than a model comprising hundreds of lines.

Software development and software quality assurance processes and best practices, including supporting documentation. Generally, software quality is emphasized in an organization when actual documented processes exist and the developer is conversant in these processes. Even more confidence-inspiring is an industry-recognized rating (such as a mid- to high-level CMMI rating) of the developer's processes. Lack of documented processes or lack of process awareness should be considered a red flag.

Defined cutoffs for code modifications. This refers to modifications in the code's capability, not the correction of errors, and has implications for configuration management. Having clearly defined cutoffs indicates awareness of the basic tenets of configuration management. Vagueness in this area can indicate version creep and schedule slippage.

Verification process execution. The process should specify exactly what will be done (e.g., module testing) and who will do it (e.g., an independent team of software engineers). See Annex A for verification techniques.

M&S error inspection. Ideally, outside experts in the language used to write the software should inspect each module from an independent perspective.

SME availability to answer software engineers' questions. The engineers writing the code will inevitably need to ask operational or technical questions. Having SMEs available will help to minimize erroneous operational assumptions.

The ACA shall assess the responses to all these issues and make an overall statement pertaining to the quality of work that can be expected from the M&S Developer in the Accreditation Report.

2. Assessing M&S Functionality and Assumptions

The ACA is not required to know the language used in programming the M&S under consideration; however, the ACA is required to have a good understanding of what the model is intended to do, the methodology it uses, and the assumptions made in coding and running the model. This knowledge can be obtained by reading the model description, user's manual, Software Requirements Specification, independent model reviews, and any other documentation about the model that is available and relevant.

The SMEs on the SCP can be very helpful in understanding the right questions to ask as well as in interpreting the explanations associated with those questions. Basic understanding of the M&S pays dividends when the ACA witnesses V&V events and can judge the significance of the results.

The ACA shall summarize the documents reviewed and other steps taken to gain an understanding of M&S functionality in the Accreditation Report. In addition, the ACA shall list the major assumptions that are made by the M&S and state the effect, if any, of each assumption on the performance of each OT&E SIU in the Accreditation Report.

3. Verifying and Validating M&S Input Data

If the M&S uses any form of input data or has parameter values hardwired into the code, the ACA addresses how that data was verified and validated. Even if the M&S is functioning perfectly, accurate results from the model cannot be guaranteed if the data required by the M&S has not been acceptably V&V'd, and the M&S has not been accredited for uses that depend on

non-V&V'd data. The ACA is expected to certify that the data used by the M&S is accurate, consistent, and suitable for use by the M&S. The data must be V&V'd in accordance with the "Data Verification and Validation" section of this chapter.

The specific activities addressing data V&V should be documented in the Accreditation Report. Key information on MCOTEA expectations for data V&V and ACA responsibilities in this regard can be found in Annex B.

4. Verifying Requirements Satisfaction and Model Accuracy

Verification is an assessment of how well the M&S satisfies its software requirements and how accurately the M&S performs. The ACA or a suitable government representative witnesses M&S verification activity in accordance with this chapter. The ACA provides an overall assessment of the verification techniques used and whether the verification activities were comprehensive and thorough in locating software errors. The ACA also assesses the verification of software requirements and comments on any that were not adequately verified.

Verification efforts are expected to locate errors in the M&S. Therefore, the model is expected to undergo a great deal of change during verification. Changes to the M&S can also be expected in validation. Once the version of the M&S that will support OT&E is finalized, previous verification tests should be rerun as a best practice.

The ACA shall summarize the verification activities in the Accreditation Report. Key information on MCOTEA expectations of verification and ACA responsibilities in this regard can be found in Annex A.

5. Validating M&S Results

It is expected that different and complementary validation techniques will be performed on the M&S to build confidence that the M&S can realistically and accurately support the SIUs needed to support OT&E. The ACA assesses each validation technique used in the overall assessment of the M&S validation. As with verification, if the M&S is changed as a result of a validation test, the ACA describes the type and adequacy of the regression testing.

The ACA summarizes all validation activity in the Accrediation Report. Key information on MCOTEA expectations for validation and ACA responsibilities in this regard can be found in Annex A.

6. Ensuring that Accreditation Criteria for Each SIU are Satisfied

The ACA must assess the satisfaction of accreditation criteria associated with each SIU. The ACA assesses the adequacy and accuracy of the data collected independent of the M&S to support a comparison with M&S results. The ACA then examines whether the accuracy levels and confidence levels (if stated) in the Accreditation Criteria are met.

Overall Assessment

MCOTEA accredits a specific version of an M&S by individual SIU. Each SIU receives its own accreditation assessment and recommendation, based on the results of the preceding steps. However, going directly to step 6, if the M&S fails to meet an accreditation criterion for any SIU, the ACA cannot recommend accreditation for that SIU. Conversely, meeting accreditation criteria does not guarantee accreditation for that SIU. For example, the data required for input and used to support the SIU may not be sufficiently V&V'd, or the M&S may make an inappropriate assumption regarding the SIU. For each SIU that passes its accreditation criteria, the ACA considers all of the preceding accreditation steps before recommending accreditation for that SIU.

VV&A Documentation Process

The following four flow charts present the basic VV&A (focusing on accreditation) process, beginning with seven steps generic to any accreditation once the test team decides to use M&S (fig.XX) From there the flow charts break into particular situations: VV&A of legacy M&S deemed suitable (fig. XX); VV&A of legacy M&S deemed unsuitable. Although MCOTEA does not generally perform V&V, the ACA supports the determination of acceptable legacy V&V, and thus must be familiar with the determination process.

Conceptual Model

According to SNI 5200.40, enclosure 1, section 2.b(2), "The conceptual model serves as a bridge between the defined requirements and the M&S design, providing the developer's interpretation of the requirements to which the model or simulation will be built." The documented conceptual model is constructed by the M&S Developer before coding begins and should contain the fundamental assumptions used by the M&S, the availability of data required by the M&S, descriptions of the functional modules (e.g., subroutines, objects, etc.) in use by the M&S, as well as the architecture used to relate the functional modules to one another and to other models or simulations. The conceptual model describes what the M&S is expected to do along with any supplemental information and data and their sources. Although MCOTEA need not participate in the development of the conceptual model, the information it contains is important to the overall understanding of the M&S. If this information is not explicitly contained in something called the conceptual model, it will have to be obtained elsewhere. Candidate alternative sources of the type of information typically found in the conceptual model are the User's Manual, M&S Description, M&S development documentation, and any previous VV&A documentation.

The ACA should review the Conceptual Model for completeness and to learn about the M&S. The Conceptual Model is validated and the validation documentation should be available for review. MCOTEA need not be part of the Conceptual Model validation, but the validation report should be reviewed to ensure the M&S Developer's interpretation of the M&S requirements is consistent with MCOTEA's.

If no conceptual model exists for a legacy simulation, it should be constructed and validated if the simulation is modified. If the conceptual model is constructed for a modified M&S, it should cover both the legacy portions and the modified portions of the M&S.

Figure 3-5 on page 3-18 illustrates the core documentation produced for VV&A. MCOTEA is responsible for producing the Accreditation Plan and the Accreditation Report. The V&V Agent produces the V&V Plan and V&V Report. The normal set of MCOTEA templates for plans and reports is not used for VV&A because the DOD-established VV&A process calls for sharing information among those who verify, validate, and accredit. A set of MIL-STD templates (MIL-STD 3022) is available for this purpose, which MCOTEA uses for consistency with other VV&A stakeholders. Table 1 illustrates the core documents that support the VV&A Process.

The Accreditation Plan and the V&V Plan are analogous to a TEMP in that they set forth the expectations of the entire VV&A process. The V&V Report and the Accreditation Report are analogous to final T&E reports in that they aggregate the results of the VV&A process.

In addition to the Accreditation Plan and Accreditation Report, MCOTEA also produces V&V Observation Plans and V&V Observation Reports, explained further in this chapter.

Writing the Accreditation Plan

The Accreditation Plan is drafted early, typically before or coincident with the V&V Plan, but is intended to be a living document that can be adjusted as the M&S and VV&A process progresses. The ACA should plan on producing the first draft of the Accreditation Plan by the time the MCOTEA SEP is completed. The document may be MCOTEA-only or may be co-written

with the PMO if the M&S will be used in DT.

From MCOTEA's perspective the most important element of the Accreditation Plan is to document SIUs and define their accreditation criteria. In addition, the Accreditation Plan defines the methodology for conducting the accreditation assessment; defines the resources needed for the assessment; and identifies issues or concerns associated with performing the assessment.

The MCOTEA OTAD signs the plan when complete and the ACA ensures that the Accreditation Plan is sent to the DOD M&S Catalog, and is entered into the MCOTEA T&E Reference Center.

Suitable Government Representatives

Normally, the ACA or another member of the OT&E team will witness V&V events. However, if attending a V&V event is not practical, MCOTEA can accept V&V results under the following circumstances:

- MCOTEA receives a copy of the event plan before the event
- The event is witnessed by a suitable government representative (can be a contractor representing the government) familiar with the M&S
- The government representative cannot be employed by, or subcontracted to, the M&S Developer or the system development contractor (if the M&S supports the system under test)
- The government representative records detailed observations, all deviations from the plan, and all caveats associated with data elements
- The government representative is available to answer MCOTEA questions after the event
- MCOTEA has access to all recorded event data
- MCOTEA receives a copy of all reports generated by the verification team pertaining to the verification event

The MCOTEA ACA documents the results of any V&V event witnessed by a suitable government representative in accordance with the procedures of V&V

Configuration Management Plan

MCOTEA requires any M&S that will undergo changes to have a configuration management plan. Whenever a change or a group of changes is made to an M&S, either to fix errors or add capabilities, the version of the M&S changes. The CMP is a critical component of the V&V effort because it is essential that the version that undergoes V&V activities is well known and tightly controlled. (Note: a V&V activity is any technique, analysis, inspection, demonstration, or test intended to verify or validate the M&S.) The CMP is normally written by the M&S Developer and reviewed by the SCP. The CMP exercises control of changes to the M&S and supporting documentation by exercising version control and tracking code changes. It secures the code against unauthorized or undocumented changes, and provides an audit trail of all changes to requirements and the M&S all the way back to original software requirements. A good CMP should contain software status accounting procedures, procedures for managing changes to software requirements, control points governing scheduled reviews, as well as requirements and procedures for regression testing when changes are made to the M&S.

As part of good configuration management, the following should be marked with the appropriate M&S version number: source code, executable code, relevant documentation, input data, any special hardware associated with the M&S, and any other applicable materials. The MCOTEA accreditation process applies to everything that supports the specific version of M&S that will support OT&E.

Observation Report.

V&V Observation Plan

MCOTEA requires a V&V Observation Plan (sample p. 6-56) before any MCOTEA representative witnesses an M&S verification or validation event. This plan, written by the ACA or other MCOTEA representative witnessing the V&V event, is similar to the Observation Plan format and process MCOTEA uses for DT Observation. The V&V Observation Plan details exactly what is being verified or validated, how the V&V event is expected to proceed, and describes the anticipated results and what they mean. Typically, several verification and/or validation techniques or activities will be scheduled for a single observation event. The ACA should obtain the plan for the V&V event as soon as it is available. The V&V event Plan is then used as the basis for MCOTEA's Observation Plan. Each observation requires an Observation Plan, but the same plan can be used to observe multiple V&V events close together in time. The ACA submits the V&V Observation Plan to the OTAD for approval.

V&V Observation Report

Even though V&V tests are generally performed by other entities, the MCOTEA ACA, another member of the OT&E team, or a suitable government representative must be present to certify the results of the observed inspections,

Accreditation Plan	V&V Plan	V&V Report	Accreditation Report
Executive Summary	Executive Summary	Executive Summary	Executive Summary
1. Problem Statement	1. Problem Statement	1. Problem Statement	1. Problem Statement
2. M&S Requirements and Acceptability Criteria	2. M&S Requirements and Acceptability Criteria	2. M&S Requirements and Acceptability Criteria	2. M&S Requirements and Acceptability Criteria
3. M&S Assumptions, Capabilities, Limitations & Risks/Impacts	3. M&S Assumptions, Capabilities, Limitations & Risks/Impacts	3. M&S Assumptions, Capabilities, Limitations & Risks/Impacts	3. M&S Assumptions, Capabilities, Limitations & Risks/Impacts
4. Accreditation Methodology	4. V&V Methodology	4. V&V Task Analysis	4. Accreditation Assessment
5. Accreditation Issues	5. V&V Issues	5. V&V Recommendations	5. Accreditation Recommendations
6. Key Participants	6. Key Participants	6. Key Participants	6. Key Participants
7. Planned Accreditation Resources	7. Planned V&V Resources	7. Actual V&V Resources Expended	7. Actual Accreditation Resources Expended
		8. V&V Lessons Learned	8. Accreditation Lessons Learned
Suggested Appendices	*Suggested Appendices*	*Suggested Appendices*	*Suggested Appendices*
A. M&S Description	A. M&S Description	A. M&S Description	A. M&S Description
B. M&S Requirements Traceability Matrix	B. M&S Requirements Traceability Matrix	B. M&S Requirements Traceability Matrix	B. M&S Requirements Traceability Matrix
C. Basis of Comparison	C. Basis of Comparison	C. Basis of Comparison	C. Basis of Comparison
D. References	D. References	D. References	D. References
E. Acronyms	E. Acronyms	E. Acronyms	E. Acronyms
F. Glossary	F. Glossary	F. Glossary	F. Glossary
G. Accreditation Programmatics	G. V&V Programmatics	G. V&V Programmatics	G. Accreditation Programmatics
H. Distribution list	H. Distribution list	H. Distribution List	H. Distribution List
	I. Accreditation Plan	I. V&V Plan	I. Accreditation Plan
		J. Test Information	J. V&V Report

Figure 3-5. Outlines of Four Core VV&A documents

analyses, demonstrations, or tests independently from those conducting the V&V activities and in their own words. After returning from a V&V Observation, the ACA (or the actual attendee) writes the V&V Observation Report (sample p. 6-57), which records the outcome of all activities for the observed V&V event and the extent to which the planned V&V activities were executed.

The report must include any deviations from the plan, any activities not performed, or any activities added to the original event plan. The V&V Observation Report documents the results of all V&V activities observed from the MCOTEA perspective. The report should include all relevant test plans, any relevant data, and the results of the testing, if known. MCOTEA may forward the Observation Report to the M&S Developer and M&S Proponent after OTAD signature if the content is substantial enough that the recipients would benefit from seeing it. Otherwise MCOTEA retains the report internally as part of the official record of VV&A activity.

The Observation Reports are used again towards the end of the VV&A process after MCOTEA receives the official V&V data and compares the record of observed events with the V&V Report.

Typically, the V&V Agent rolls up V&V event results into one aggregated report (the V&V Report), meaning that MCOTEA will most likely have to wait until all V&V is complete before receiving data from any one event; however, the ACA should contact the owner of the V&V results if there are any questions on any particular V&V event. (MCOTEA may also request data along the way if an early look would be beneficial to the accreditation process.)

Once the V&V Report is received, the ACA analyzes the data and results for accuracy, completeness, and for fulfillment of accreditation criteria. This analysis is included in the Accreditation Report.

Accreditation Report

The Accreditation Report is typically written by the ACA and summarizes all data, information, and activity, explicitly or by reference, used in the accreditation assessment. To enable informed accreditation decisions, the Accreditation Report must provide insight into M&S capabilities, limitations, and any uncertainties about M&S capabilities related to the SIUs. The ACA must ensure that the following information is accounted for in the report or its annexes:

- Name and the version number of the M&S being accredited
- Date of report and the name/organization of author (accreditation agent)
- Description of the M&S
- Summary of model assumptions
- Summary of V&V activities/processes performed in support of this accreditation
- Summary of previous VV&A activities that apply to this accreditation and why they apply
- Assessment of each of the six aspects of a MCOTEA accreditation as explained in the Accreditation Process section of this chapter

External references and documentation that support recommendations in the report must be archived in the MCOTEA T&E Reference Center, regardless of who wrote them.

The ACA forwards the Accreditation Report and a draft Accreditation Decision Letter (explained below) to the MCOTEA CRB. The COT signs the approved report and forwards it and the draft Accreditation Decision Letter to the Accreditation Authority.

Accreditation Decision Letter

The ACA drafts the Accreditation Decision Letter as a standard naval letter. The letter must specify M&S name, version number, and version date being accredited. The letter's

content is based on the recommendations resulting from the Accreditation Assessment, including the following:

- The degree to which each SIU is accredited (Fully, Partially, Decision Pending, or Not Accredited)
- Configuration management requirements of the M&S in order to maintain accreditation
- Any requirements for the data used as input to the M&S or restrictions on the data generated by the M&S
- Any additional V&V requirements by SIU
- Any additional questions that must be answered before accreditation by SIU
- Any additional documentation required before accreditation
- A description of any limitations on the accreditation decision

The Accrediation Decision Letter remains in effect for the accredited version of the M&S as long as the intended uses remain unchanged, or until revoked, in writing, by the AA.

Accreditation Decision

The MCOTEA AA has the following options regarding each SIU:

- Full accreditation. Fully accredits the SIUs that merit full accreditation.
- Partial accreditation. SIUs are accredited under certain conditions by placing constraints under which the SIUs may be applied to OT&E.
- Accreditation Decision Pending: Full SIU accreditation is still possible assuming additional information is received, additional testing is accomplished, or modifications are made to the M&S.

Separation of DT and OT&E Accreditations

In many programs, the PMO will have uses for the same models in DT that MCOTEA intends to use in support of OT&E. If the PMO intends to use the models under consideration for DT SIUs, MCOTEA can leverage the PMO's V&V efforts. Under these circumstances it will probably make sense for the DT and OT&E accreditations to use the same Accreditation Plan since most of the plans' required content will be the same.

Even with a shared plan, the SIUs and accreditation criteria for DT and OT&E must be called out separately and DT and OT&E SIU accreditations remain completely separate for four fundamental reasons:

- The Accreditation Authorities for DT and OT&E are different
- SIUs for DT and OT&E are independent of one another and most likely differ from each other
- Different validation information will apply to different SIUs
- DT and OT&E timelines are different, and accrediting OT&E M&S later than DT M&S may allow MCOTEA to take advantage of validation opportunities that might arise during DT and/or OA event execution.

When the same M&S is used to address both DT and OT&E issues, MCOTEA works closely with the PMO and SCP to resolve any issues associated with accreditation to increase the probability that both accreditations can be successfully accomplished. MCOTEA OT&E team members, in particular the ACA, should strive to participate in all of the V&V activities associated with the DT accreditation. All of the verification activities associated with DT accreditation are also required by the MCOTEA process, and the DT validation activities will be useful in building MCOTEA confidence in the M&S. The MCOTEA ACA must ensure that, in addition to the DT V&V activities, MCOTEA requirements for verification and validation are met.

- Not accredited: The SIU cannot be accredited to support OT&E.

The different accreditation options apply by SIU. Therefore, an Accreditation Decision could conceivably fully accredit some SIUs, partially accredit others, conditionally accredit some, and not accredit still others.

Partially accredited SIUs, conditional SIUs that do not undergo remediation, or unaccredited SIUs imply limitations to the OT&E.

The MCOTEA AA signs the letter after making any desired changes. The accreditation is not official until the letter is signed. The letter remains in effect for the accredited version of the model as long as the intended uses remain unchanged, or until revoked by the AA.

The ACA is responsible for filing the signed Accreditation Decision Letter in the MCOTEA T&E Reference Center and the DOD M&S Catalog.

Accounting for Previous Accreditation

MCOTEA strives to leverage any previous VV&A activity for the model under consideration to the maximum extent possible, but MCOTEA determines its V&V requirements independently of what has already been accomplished. This independent examination of V&V requirements may result in the need for additional V&V activities.

MCOTEA may reuse any unaltered M&S version previously accredited by MCOTEA for a given set of SIUs assuming the previous accreditation criteria are acceptable for the new application. However, if the M&S has been modified in some way, the SIUs are different, or the accreditation criteria have changed, a new accreditation is required.

Following are four examples of situations that require new accreditation but can most likely accept previous V&V or portions of it:

Modified M&S Version

Situation 1: MCOTEA has accredited M&S version 1.0 and later would like to modify it and use version 1.1 to support MCOTEA testing or evaluation. Response: MCOTEA must separately accredit version 1.1. In this case, at least some of the original V&V work is likely to be usable in support of version 1.1 V&V.

Same M&S, Different SIUs

Situation 2: MCOTEA has accredited M&S version 1.0 for SIUs for a particular application and would like to reuse this version for different SIUs. For example, MCOTEA may have accredited M&S in support of the OT of a chem-bio protective garment that models chemical penetration of the garment and chemical burns to the wearer. Later use might be to supplement the OT of a non-lethal weapon system by modeling burns from heat sources. Response: Version 1.0 must be reaccredited because the thermal burn SIUs must be accredited separately from the chemical burn SIUs. Presumably, however, most of the original verification efforts and perhaps some of the original validation efforts could be reused in the second accreditation.

Same M&S and SIUs, Different Accreditation Criteria

Situation 3: MCOTEA has accredited M&S version 1.0 SIUs for one test article and would like to reuse this same M&S version and SIUs for a different but related test article (chem-bio garments, for example). Assuming that the accreditation criteria supporting the OT&E of the first garment are different from the criteria for the second OT&E, MCOTEA must accredit the M&S separately to satisfy the new criteria. However, most of the work from the previous VV&A effort presumably

could be reused, leaving only minor validation activities required for the second accreditation.

M&S Accreditation from Another Organization

Situation 4: Another organization accredits an M&S with exactly the same SIUs that MCOTEA needs the model to support. Response: MCOTEA must still independently accredit the M&S to support OT&E, even if the previously accredited version of the M&S is identical to the one MCOTEA wants to use. The first reason for this is directive in nature: only MCOTEA can accredit an M&S for use in a MCOTEA OT&E. In addition (and fundamental to the concept of accreditation), no guarantee exists that the other organization's accreditation process meets MCOTEA's accreditation requirements. In summary, extensive previous use of an M&S or accreditation by another organization does not automatically guarantee accreditation of the M&S for SIUs in support of MCOTEA OT&E. See the section below for details on reaccreditation.

MCOTEA's Reaccreditation Process

"Any subsequent use in a new application domain or modification of the M&S will require a reaccreditation process" (SNI 5200.40). The MCOTEA reaccreditation process is the same as the accreditation process, except that the ACA will leverage as much of the V&V efforts from any previous accreditations as possible.

The degree to which the information from any previous VV&A effort can be reused depends on the quality of the associated documentation. The ACA must be able to discern the following elements of quality in VV&A documentation:

- The exact version of the M&S previously accredited must be evident.
- The M&S must not have changed, or the change and regression testing of the changed M&S must be sufficiently documented.
- Terms such as "accurate," "sufficient," or "adequate" must be supported by documented evidence.
- The documentation must clearly discuss VV&A procedures and data and the results of inspections, analyses, demonstrations, and tests.
- The details of a V&V event should include exactly what was done and under what conditions, who observed and documented the event, the resulting data, how the data was analyzed, and the factual results of the analyses.

Where to start

The ACA begins the reaccreditation process by following the same steps used for initial accreditation. Therefore, the ACA needs to examine the following basic information:

- The M&S Developer's software development and software quality assurance processes
- What the M&S does and how it does it
- The basic assumptions used in the M&S
- Conceptual model
- User's Manual
- Programmer's manual
- Any other available introductory documentation
- Documents that describe past actions
- PRevious Accreditation Plans
- Previous V&V Reports
- Previous Accreditation Reports
- Configuration Management documentation

The Previous Accreditation Plan will show what was intended in the previous accreditation and the Report will show what was actually accomplished. The Previous V&V Report should contain a wealth of information in support of the accreditation. If elements of the Accreditation Plan and V&V Report are not addressed in the Accreditation Report, the ACA needs to understand why this is the case. The Accreditation Report should also include

several references, typically sources of data or documented tests used to support the accreditation.

For MCOTEA to accept previous accreditation results, the M&S must have been under strict configuration control between the previous accreditation and the present. If the M&S version has changed in any way since the previous accreditation but no record exists of the changes or of V&V to support those changes, the ACA must plan appropriate V&V activities to compensate for this shortfall.

Using Previous Verification &Validation Efforts

Previous Verification

MCOTEA's accreditation requirements for verification remain the same for first-time verification and in support of reaccreditation.

If the model of current interest to MCOTEA has changed from the original, verified version, the ACA can still use the previous VV&A information to gain familiarity with the model's capabilities. Although the code itself will have changed from version to version, functional modules within the code may or may not have changed. To the extent that functional modules have not changed from the original version, the verification efforts of those modules may still be applicable. However, those efforts may yet be insufficient to meet MCOTEA's needs. Depending on the thoroughness of the previous verification effort, MCOTEA may require additional verification of codes that have already undergone a set of verification procedures.

In any case, when the previously verified model version has changed, all modified functional modules of that version and the interactions between all modules need to be re-verified. If changes to the M&S were not sufficiently documented, the entire code will require some level of new verification activity. Under these circumstances some of the verification techniques described in this chapter, such as modular string testing, should be considered.

The ACA must document all previous verification activities used in the MCOTEA accreditation in the Accreditation Report, along with any supplemental verification activities required by MCOTEA.

Previous Validation

Past successful validation efforts should give the ACA a degree of confidence in the M&S. However, unlike certain verification techniques that focus on functional modules of the M&S, validation testing tends to examine the validity of the overall M&S. Therefore, if the M&S has changed at all since the last accreditation, previous validation efforts relevant to the current SIUs may need to be repeated and new validation activities may be required. At a minimum this will involve ensuring that the M&S meets the accreditation criteria in the new Accreditation Plan.

Independent Verification and Validation

Independent Verification and Validation (IV&V) is, by definition, independent of the M&S Proponent and the M&S Developer's regular V&V of a model. IV&V is optional unless directed by the M&S Proponent, the AA, or a higher authority. MCOTEA may direct that an IV&V be conducted on an M&S if the MCOTEA AA believes it is necessary to establish the requisite level of confidence in a model for support of OT&E. The requirements for IV&V are identical to those of a regular V&V as described in this manual; the only difference is the entity performing the V&V.

Results of the IV&V are documented in a separate IV&V Report, analogous to a V&V Report, and should contain all of the V&V Report elements. The report should be delivered to the ACA for consideration

Further Information About V&V Documentation

Verification and Validation Plan

MCOTEA requires a V&V Plan, in accordance with SNI 5200.40, for any M&S that requires additional verification, validation, or both. The V&V Plan is developed by the M&S Developer or the M&S Proponent, with MCOTEA input. The ACA and SCP should review the plan for accuracy and completeness. The SCP must specify the due date of the V&V Plan based on realistic estimates for model completion and program schedule. The V&V Plan is a living document, adjusted as the M&S and VV&A processes progress.

The contents of the V&V Plan are seen in figure XX. Ideally the V&V Plan includes the test plans for all V&V activities that require testing; however, this level of detail may be filled in later. MCOTEA must receive a copy of the detailed test plan at least 15 days before any V&V event.

For legacy models (modified or requiring additional V&V activities), the V&V Plan addresses legacy model assumptions, capabilities, and any previous VV&A activities as well as an explanation of all planned M&S enhancements and all planned V&V activities. Although MCOTEA leverages all previous, relevant V&V activities, MCOTEA determines its V&V requirements independent of what has already been accomplished. If the previous V&V efforts were insufficient or undocumented, MCOTEA may require additional V&V. In addition, MCOTEA will still require that the M&S satisfies the accreditation criteria for OT&E SIUs.

Verification and Validation Report

MCOTEA requires a V&V Report, in accordance with SNI 5200.40, to document and describe the details of all V&V events. The V&V Report is developed by the M&S Developer or the M&S Proponent, with MCOTEA input. This report documents evidence supporting the functionality and fidelity of M&S to satisfy OT&E SIUs, M&S requirements, and model accuracy requirements. The V&V Report documents the M&S Developer, the M&S Description, M&S assumptions, and any risks associated with using the M&S or associated data. The V&V Report details all verification and validation activity to include

- a complete description of V&V methodologies, organizations, and individuals involved in V&V and a summary of their findings
- a description of actions taken as a result of V&V
- explicit identification of known M&S capabilities, limitations, and restrictions
- detailed descriptions of all V&V techniques, analyses, inspections, demonstrations, and tests to include scope, limitations, methodology, scenarios, environments, participants, and all supporting data
- a compilation of any V&V reports pertaining to previous relevant V&V activities being leveraged for the current V&V effort
- data V&V activities including the original reason the data was generated, how the data was generated (the more detail the better), and any assumptions made in generating the data

The V&V Report should be designed for use as a reference for follow-on VV&A activities and for future regression testing.

Annex A. V&V Process and Techniques

and analysis in conjunction with other material in support of the accreditation.

Basics of Verification

Before discussing how to verify something, it is useful to repeat the definition of verification. According to reference (DODI 5000.61), verification is, "The process of determining that a model or simulation implementation and its associated data accurately represent the developer's conceptual description and specifications."

The bolded words indicate the two aspects of verification. The first is accuracy. For example, if the model is a computer code, one must acknowledge there are always undetected errors in the code; the larger the code, the more undetected errors. (Anybody doubting this assertion should consult Microsoft about the "Blue Screen of Death".) The goal is to minimize the number of undetected errors, thus ensuring the code is "accurate".

The second aspect of verification is to ensure the code reflects the specifications spelled out for its construction (normally in a Software Requirements Specification). If the code doesn't do what it was supposed to do, it doesn't matter how accurately it does it. Typically, model developers will emphasize this aspect of verification because it is easy to list requirements and show how they will be verified. Checking the M&S for accuracy is arguably more difficult.

At this point it is useful to note that checking an M&S against its requirements is typically a verification function. Occasionally, a requirement will spell out an M&S capability as compared to the corresponding real world capability (resulting in a validation of that requirement), but this is rare. Requirements are typically "verified", while validation is used to confirm the M&S is a realistic representation of the real world and is capable of satisfying the designated specific intended uses.

The remainder of this section describes selected verification techniques. The DOD VV&A Recommended Practices Guide contains several additional verification techniques that should be considered. The ACA should coordinate with the V&V Agent/SCP and the Model Developer to determine the set of verification procedures that makes the most sense for the M&S under consideration.

It is extremely important that all techniques used to verify the M&S be thoroughly documented in the V&V Report, and summarized in the Accreditation Report. This increases the credibility of both reports and allows for reuse of V&V work in future accreditation efforts.

Verifying an M&S for Accuracy

As discussed, the object is to minimize the number of undetected errors in the M&S. When a code is first being written, if there are errors in implementing the computer language, the code will generally not run until these "syntax" errors are corrected. These are the easy errors to track down, and an experienced programmer can frequently construct a section of code without any errors in syntax.

It is the errors in logic that are the most difficult to detect and locate. When a software engineer constructs a code, he/she invariably is required to make certain, seemingly benign assumptions. Since it is a rarity for the software engineer to be a SME in the real world processes being modeled, these assumptions are sometimes erroneous. It is wise to have SMEs available to answer the questions of the software engineers while coding the M&S; however, even if an SME is available to answer questions during coding, erroneous assumptions and other errors in logic can still be implemented in the code. The following describe some techniques

useful in understanding the M&S and locating errors within it. Other verification techniques can be found in reference (DOD VV&A Recommended Practices Guide, 2001).

Basics of Verification

Before discussing how to verify something, it is useful to repeat the definition of verification. According to reference (DODI 5000.61), verification is, "The process of determining that a model or simulation implementation and its associated data accurately represent the developer's conceptual description and specifications."

The bolded words indicate the two aspects of verification. The first is accuracy. For example, if the model is a computer code, one must acknowledge there are always undetected errors in the code; the larger the code, the more undetected errors. (Anybody doubting this assertion should consult Microsoft about the "Blue Screen of Death".) The goal is to minimize the number of undetected errors, thus ensuring the code is "accurate".

The second aspect of verification is to ensure the code reflects the specifications spelled out for its construction (normally in a Software Requirements Specification). If the code doesn't do what it was supposed to do, it doesn't matter how accurately it does it. Typically, model developers will emphasize this aspect of verification because it is easy to list requirements and show how they will be verified. Checking

Basics of Verification

Before discussing how to verify something, it is useful to repeat the definition of verification. According to reference (DODI 5000.61), verification is, "The process of determining that a model or simulation implementation and its associated data accurately represent the developer's conceptual description and specifications."

The bolded words indicate the two aspects of verification. The first is accuracy. For example, if the model is a computer code, one must acknowledge there are always undetected errors in the code; the larger the code, the more undetected errors. (Anybody doubting this assertion should consult Microsoft about the "Blue Screen of Death".) The goal is to minimize the number of undetected errors, thus ensuring the code is "accurate".

The second aspect of verification is to ensure the code reflects the specifications spelled out for its construction (normally in a Software Requirements Specification). If the code doesn't do what it was supposed to do, it doesn't matter how accurately it does it. Typically, model developers will emphasize this aspect of verification because it is easy to list requirements and show how they will be verified. Checking the M&S for accuracy is arguably more difficult.

At this point it is useful to note that checking an M&S against its requirements is typically a verification function. Occasionally, a requirement will spell out an M&S capability as compared to the corresponding real world capability (resulting in a validation of that requirement), but this is rare. Requirements are typically "verified", while validation is used to confirm the M&S is a realistic representation of the real world and is capable of satisfying the designated specific intended uses.

The remainder of this section describes selected verification techniques. The DOD VV&A Recommended Practices Guide contains several additional verification techniques that should be considered. The ACA should coordinate with the V&V Agent/SCP and the Model Developer to determine the set of verification procedures that makes the most sense for the M&S under consideration.

It is extremely important that all techniques used to verify the M&S be thoroughly documented in the V&V Report, and summarized in the Accreditation Report. This increases the credibility of both reports and allows for reuse of V&V work in future accreditation efforts.

Verifying an M&S for Accuracy

As discussed, the object is to minimize the number of undetected errors in the M&S. When a code is first being written, if there are errors in implementing the computer language, the code will generally not run until these "syntax" errors are corrected. These are the easy errors to track down, and an experienced programmer can frequently construct a section of code without any errors in syntax.

It is the errors in logic that are the most difficult to detect and locate. When a software engineer constructs a code, he/

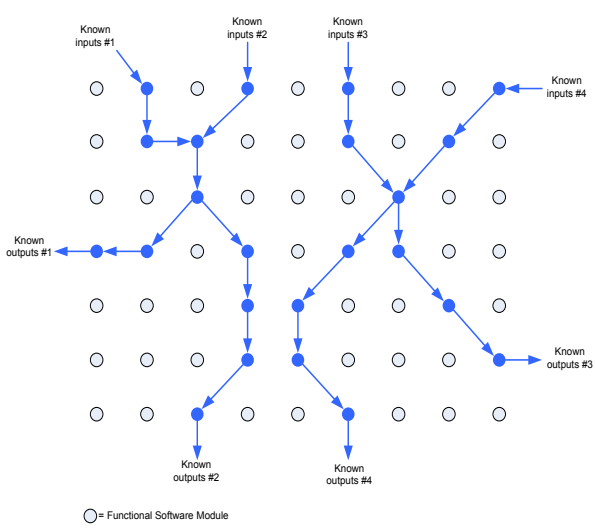


Fig. 4-6.
Testing Strings of Software Modules

she invariably is required to make certain, seemingly benign assumptions. Since it is a rarity for the software engineer to be a SME in the real world processes being modeled, these assumptions are sometimes erroneous. It is wise to have SMEs available to answer the questions of the software engineers while coding the M&S; however, even if an SME is available to answer questions during coding, erroneous assumptions and other errors in logic can still be implemented in the code. The following describe some techniques useful in understanding the M&S and locating errors within it. Other verification techniques can be found in reference (DOD VV&A Recommended Practices Guide, 2001).

techniques can be found in reference (DOD VV&A Recommended Practices Guide, 2001).

Another aspect of verifying a model for accuracy is the examination of how the M&S accommodates unanticipated or out of specification range inputs. The model should be protected from erroneous data entry. Furthermore, models should not allow representations that violate the laws of physics.

Documentation Walkthrough

It is important to identify the assumptions made when the model was coded. This helps in determining whether a model is appropriate or not for a specific use. One way to identify these assumptions is to systematically go through the model documentation. Many of the explicit assumptions made in the construction of the M&S, its internal parameters, or other input data can be determined by a careful review of the M&S User's Manual or any other documentation that describes the logic used in the M&S. Ideally, this review would be accomplished by the appropriate SME.

In some cases, the M&S documentation required for the Documentation Walkthrough will not exist. In those cases, the verification effort is obviously weakened and the explicit assumptions will have to be identified by interviewing the software engineers who wrote the M&S and by inspection of the M&S itself.

M&S Inspection

The source code should also be checked to see exactly how the explicit assumptions are implemented. The inspection of each functional module (subroutine, object, etc) is accomplished by software engineers conversant in the computer language used in constructing the M&S. It can be performed by employees of the M&S Developer, but it is preferable this be performed by software engineers not involved in coding the M&S.

Inspection serves three purposes. First, as the software engineer goes through the module, he/she makes note of how all explicit assumptions were implemented in the model. In addition, this inspection can be used to identify assumptions *implicit* in the way the model itself was coded, including noting any fixed parameters coded into the M&S. These implicit assumptions should be checked with the appropriate SME for accuracy. Lastly, the inspection of each functional module also generally represents the first time independent experts have had an opportunity to locate logic errors in the M&S. Again, the presence of a SME will facilitate finding the logic errors at this stage.

It would be ideal if the SME that examines the documentation and the software engineer that examines the model are the same person, but it is rare to find these diverse capabilities in one person. These operational and software engineering reviewers can be part of the DT team, members of the SCP, independent contractors, or employees of the software developer, but they should not be the same people that developed the M&S in the first place. The fact that the inspection(s)

was performed, how it was performed, and the results of the inspection, to include a review of all identified assumptions and any errors that were discovered, should be documented in the V&V report. Additional information on the formal inspection process can be found in reference (DOD VV&A Recommended Practices Guide).

Modular String Testing

Before checking strings of modules, individual modules (subroutine, object, etc.) should be checked for correct and accurate behavior. When checking the behavior of individual modules, it is often worthwhile to "instrument" the module, that is, insert additional code in the M&S in order to record parameter values at strategic points in the module. This instrumentation of the module allows for tracking parameter values to ensure the module behaves as expected.

After testing functional modules individually, it is useful for the M&S Developer to test strings of modules to ensure interface logic between modules and model outputs are consistent with expectations (fig. 4-6 previous page). First, the software engineer decides on a logical grouping of modules to test. After constructing the necessary input data, the software engineer does a hand calculation on the expected outputs, based on his/her understanding of each functional software module being tested. The inputs, modules being tested and expected outputs are all documented in the V&V plan and V&V report. The results of each test run are used to locate any logical errors in the modules under test. Some notional examples of this technique are shown in figure C. Some modules may be tested more than once, but all should be tested *at least* once. Furthermore, the paths through the M&S that will be frequently used should also be tested in this way. Note that this technique only tests operational logic that the software engineer understands. Here again it is helpful to enlist an operational or system SME in order to capture as many logical errors as possible in the M&S. Instrumentation of the modular string is also useful in locating errors and confirming desired results.

From MCOTEA's perspective, the OT&E team's ACA only needs to see each modular string test performed once – correctly. This allows the M&S Developer to conduct the test as many times as needed to catch all the logic errors the test is capable of catching before MCOTEA verifies the test has been successfully completed. The modular string testing, including the input data, the modules tested and the output attained should be reported in the V&V report and the Accreditation Report.

A big advantage to this technique applies to regression testing of the M&S when the model is changed because errors are corrected in the version of interest, or when verifying a follow-on version of the model in a future verification effort. Since these test cases are thoroughly documented in the V&V plan, V&V report, and Accreditation Report, it should be relatively easy to repeat the testing, as required, to ensure no unwanted changes in M&S behavior have been introduced by changes to the M&S due to error correction or changes in model functionality. If, during the verification and validation process, changes are made to a functional software module that are designed to fix newly discovered errors, at a minimum all verification tests that involve that module must be re-run on the final version of the model (the version to be used during OT&E). As a best practice, all documented verification tests, regardless of whether the included functional software modules have been modified or not, should be re-run on the final version of the model.

Verifying That an M&S Meets Specifications

Typically the expectations for an M&S are spelled out in the Software Requirements Specification (SRS). Each of the

requirements should be called out in the V&V plan. Verification of the requirements is typically done by Inspection, Analysis, Demonstration, or Test. The choice of verification methods to use on a particular requirement is left to the organization doing the verification (probably the M&S Developer) with concurrence of the SCP. The verification methods are described below:

Inspection: The examination and review of descriptive documentation and comparison of appropriate characteristics with a predetermined standard. This method may require access to the source code.

Analysis: Analysis includes quantitative and/or qualitative proof that the code meets specific requirements by technical evaluation using mathematical equations, charts, graphs, and representative data.

Demonstration: This involves the operation or adjustment of the code. The code may be instrumented and its performance monitored, but only as an indirect function in support of the demonstration. Quantitative measurements are generally not taken except in cases where test operators make visual measurements/counts or where simple devices such as a stopwatch are used to estimate time performance. Generally, demonstration results may be noted by a simple YES or NO. Success and failure criteria will be established for each demonstration objective prior to the demonstration.

Test: Exercising the applicable code under appropriate conditions with instrumentation to collect/analyze/evaluate the data to ensure the requirements are met. Acceptability of the code will be determined by pre-established quantitative criteria consistent with the required characteristics stated in the applicable specification. A test plan is generated before each test.

Just as in the case with verifying the M&S for accuracy, when verifying that the M&S meets its requirements, MCOTEA only needs to see results when the M&S Developer is comfortable the verification will be a success. MCOTEA requirements for the four verification methods are:

Inspection: MCOTEA receives a plan describing what is to be inspected and how it will be inspected before the verification event. A member of the MCOTEA OT&E team or a suitable government representative is present during the inspection, can ask questions during the inspection and answer future questions about the inspection, and can independently confirm the inspection results.

Analysis: MCOTEA receives a copy of the full analysis and any associated assumptions and data. MCOTEA must have access to those that did the analysis to answer questions. A member of the MCOTEA OT&E team must independently confirm that the analysis is correct.

Demonstration: MCOTEA receives a plan describing the demonstration, including what is to be demonstrated. A member of the MCOTEA OT&E team or a suitable government representative is present during the demonstration, can answer future questions about the demonstration, and can independently confirm the demonstration results.

Test: MCOTEA receives a copy of the test plan for review and comment. A member of the MCOTEA OT&E team or a suitable government representative is present during the test, can answer questions about the test, and can independently confirm the test results.

Generally speaking, MCOTEA will want the appropriate member of the OT&E team to witness any verification event. The team member witnessing the verification event is responsible for documenting the verification results in accordance with the V&V Observation Report template (notice the similarity to the DT Observation Report) as shown in this chapter. The

ACA will reference this documentation, and may need additional information from the OT&E team member or suitable government representative that witnessed the event, when checking the V&V report for accuracy.

Validation Process

Basics of Validation

Reference (DODI 5000.61) defines validation as, "The process of determining the degree to which a model or simulation and its associated data are an accurate representation of the real world from the perspective of the intended uses of the model." Validation is accomplished by comparing the output of an M&S, with respect to its intended uses, to real world known or expected behavior of the subject it represents. In order to be valid, the M&S output must replicate the real world subject being modeled within the established degree of fidelity. If an M&S does not produce valid representations of the real world system or processes in question, conclusions based on using the M&S will be erroneous resulting poor decisions or actions. Therefore it is essential to establish the validity of the M&S prior to using it to support any decisions or actions.

M&S are used to support OT&E when using the actual systems or processes being modeled to gather sufficient data are impossible, unsafe, or impractical. Since an M&S represents an approximation of the real world, it will always have limitations. A given M&S will never be absolutely valid. For this reason the SIUs in support of OT&E are identified early so the V&V effort remains focused on the right M&S uses. M&S validation activities should be accomplished with an eye toward the SIUs. This is not meant to limit M&S validation to just an examination of the SIUs with respect to their associated accreditation criteria (although this is a key part of validation). However, all validation activity should be related to the SIUs to avoid needlessly expending scarce resources by attempting to make an M&S more capable than it needs to be.

All aspects of the M&S require validation to include the model itself, the data used by the model, any look-up tables used, any extrapolation techniques used, any methodologies used that are external to the M&S, and any required interfaces between the M&S and another M&S, system, or entity. The M&S may be validated in pieces, but it shall also be validated in its final configuration, using the applicable input data, as it will be run during support to OT&E. If an M&S consists of a federation of models, the federated M&S shall be validated as it is intended to be run. Even if all the components of the federation have been independently validated, the federation of models shall be validated while functioning as the intended federation. The OT&E team member (likely the ACA) witnessing the validation event is responsible for documenting the validation results in accordance with the V&V Observation Report template (notice the similarity to the DT Observation Report) as shown in this chapter.

The following section describes some common validation techniques. For more information on these and other techniques, see reference (DOD 2001). The more validation techniques used to successfully validate an M&S functionality, the more confident the MCOTEA accreditation agent and authority will be that the M&S is a credible representation of that functionality.

Common Validation Techniques Using Data From the Modeled System or Environment

If the M&S represents the operation of an existing system, the best means of validation is to compare M&S results to the behavior of the actual system under as close to identical conditions as possible. This can be problematic, in that

the M&S conditions can be precisely specified, while the operating conditions of the actual system, although as tightly controlled as possible, may result in sources of comparison error. The comparison errors introduced by real world operations must be accounted for in the validation criteria. One way to define validity is if the M&S results fall within a specified error interval, say ±10%, at a desired confidence level, say 80%. Note that the error interval defines whether a particular M&S result compared to the real world is valid, while the confidence level defines the number of trials required. The error interval and confidence level together set the validation criteria for each validation check.

However, validation of the M&S shall be accomplished regardless of whether or not the corresponding real world system currently exists or the real world environment is available for comparison. If a system corresponding to the M&S does not currently exist, or the environment is not available for comparison, there are other validation options.

Using Data From Related, Existing Systems or Environments

Lacking an existing system or suitable environment from which to gather validation data, data from a related, existing system or environment can be used to help validate the M&S. The technique would be to construct a preliminary M&S of the existing system or related environment and perform a validation of this preliminary M&S. Once the preliminary M&S is suitably validated, it is modified to create the desired M&S that represents the proposed system or environment. The fewer the modifications needed to the preliminary (validated) M&S, the higher the confidence in the desired M&S. Greater confidence in an M&S constructed in this way can be obtained by employing some of the other techniques in this section.

Using SMEs

Whether or not there is an existing system or suitable environment corresponding to the M&S, SMEs can be helpful in building confidence that an M&S is valid. SMEs view the M&S output under various conditions for reasonableness, based on their experience. The selection of SMEs is important. SMEs should be experts in the warfare area or technical area corresponding to or using the system being modeled by the M&S or experts in systems similar to the system being modeled. The SMEs should also have an understanding of the OT&E SIUs designated for the M&S. If SMEs are used for validation, more than one should be used and they must come to a consensus before the validation is useful. If the SMEs think the M&S results are reasonable, that strengthens the case for validation.

A useful exercise, called a Turing Test, is to show the SMEs data from the real world and corresponding data from the M&S without knowing the sources of the data sets. If the SMEs can accurately discriminate between the two data sets, the reasons they cite can be useful in correcting errors in the M&S. If the SMEs cannot agree on the sources of the data sets, that is another argument in favor of M&S validation for the uses implied by the data sets.

Using Another Model

Once MCOTEA has accredited a model for a specific use, the results of that specific version of the model are trusted by MCOTEA for that specific usage. Therefore, the previously accredited model may be used to validate the results of another model, as long as it is for the previously accredited usage.

This validation technique should be used with caution for the following reasons. Typically a model is considered accurate if its results fall within the desired accuracy of the real-world results. If this accuracy were say ±10 percent of the real-world value, the first model (model A) could be as much as 9.999 percent off and still pass. If this

model were then used to validate another model (model B), and model B was also off of the model A results by 9.9999 percent in the same direction, model B would be close to 20 percent off the real-world value. But since it was less than 10 percent off the model A results, model B would still pass using this validation method.

Another reason this technique should be used with caution is the situation where model B is based on model A. This often happens when there is a desire to improve model A. If there were an undiscovered systematic error in model A, and model B were based on model A, this undiscovered error would probably be conveyed to model B, the daughter of model A. If model A were used to then validate model B, the error would never be discovered, since model B would reproduce the same erroneous results as the original model A. Under these circumstances, model A could only be used to verify that there were no new errors introduced during the coding of model B.

Therefore, using a previously accredited model (model A) to validate second model (model B) can only be done under the following circumstances:

- MCOTEA has previously accredited model A for the SIUs under consideration
- Model B was constructed independently from model A, that is model B is not a daughter of model A
- The usage being validated for model B is identical to that previously accredited for model A
- If a future model (model C) uses this validation technique, only the originally accredited model may be used as the validation tool, that is, model A can be used to validate model C for a previously accredited usage, but model B (validated using model A results) cannot. An exception to this rule is if model B was derived from model A and model C is derived from model B. Model A cannot be used to validate model C, its granddaughter.

Using Predictions

A prediction is obtained by running an M&S under conditions that will be experienced in the future. Predictions are useful since there is no way to consciously or unconsciously "back in" the model results. The M&S is run, predicted values and data are recorded, and the M&S results are then compared to the real world results at some future time when the predicted conditions are experienced. Predictions accurate to the required level of fidelity support M&S validation. Predictions can also be used to discover errors in the M&S or to update parameter values when the M&S results disagree with the real-world results.

Sensitivity Analyses

Sensitivity analysis determines factors having the greatest impact on M&S results and that should be modeled most carefully. Clearly sensitivity analyses can be used to locate coding errors and might be considered part of the verification process. However, unexpected behavior during sensitivity analysis might indicate invalid behavior as well. If small changes in a value correspond to large changes in M&S output, sensitivity analysis will also reveal those values that need to be specified with the most accuracy.

Candidates for sensitivity analysis in the M&S are:

- Parameter values
- Probability distribution selection
- Assumptions

These things should be chosen in a way that most closely represents reality.

Annex B. Data Verification and Validation

Typically, MCOTEA is interested in using the data that is output from an M&S to support some aspect of OT&E. However, many models require certain parameters be set or certain data be input in order for them to produce the needed output. These input data can be fed into the M&S by an operator in order to fill required data fields prior to each run of an M&S, or these data might be "hardwired" as fixed parameters within the model itself. The accuracy of M&S output is just as dependent on the input data as it is on the accuracy of the M&S itself. Therefore, in addition to any M&S to be used by MCOTEA in support of OT&E, any data used as input to the M&S, or as fixed parameters within an M&S must be verified and validated. Ref (MIL-STD-3022) defines data V&V as, "The process of verifying the internal consistency and correctness of data and validating that it represents real-world entities appropriate for its intended purpose or an expected range of purposes." The types of data that require V&V are data used:

- To verify M&S requirements
- To verify M&S accuracy
- To build the conceptual model
- To validate the M&S
- To perform experiments in support of OT&E or M&S V&V
- To run combat support decision aids
- As input to any M&S supporting OT&E

Even if the data are consistent and accurate, the data set may not be suitable for a given application. The data might be incompatible with the application, it might generated based on assumptions that are not compatible with the M&S assumptions, or it might not have been generated at an appropriate level of fidelity. Given this, any data requiring V&V must be accompanied by information concerning the original reason the data was generated, how the data was generated (the more detail the better), and any assumptions made in generating the data. This will give the ACA information pertaining to the quality of the data and if the data is appropriate for the intended use.

The age of a data set is irrelevant. As long as the data can be V&V'd in accordance with this chapter, it may be used to support MCOTEA OT&E.

Data Verification

The verification of data focuses on its accuracy. The idea is to ensure the data has been accurately translated, is complete, is credible, is interpreted correctly when used by the M&S, and supports the input requirements of the M&S. Data can be verified by inspection using a process much like proof-reading; it helps ensure the data isn't inadvertently changed when transcribing it from its point of generation to the M&S input. A SME is useful in data verification, since a SME can often identify data that appear unreasonable under a given set of conditions. A SME can help decide if the data comes from a credible source and that the data has been interpreted correctly when translated into M&S parameters.

Another aspect of data verification is ensuring it comes in the expected form and is properly prepared for use in the M&S. For example, phone numbers in the United States come in the form xxx-xxx-xxxx. A data entry (phone number) not conforming to this form may be erroneous. A more sophisticated verification check on this data might involve ensuring the first 3 digits represent a valid area code within the U.S.

From the MCOTEA perspective, the ACA must ensure that data verification procedures for the M&S are in place, are being executed, and all input data are verified before M&S execution in support of OT&E. All data verification activities and processes shall be documented in the V&V Report and the Accreditation

Report.

Data Validation

Data is validated to ensure it accurately and adequately represents the real world to be simulated. Data is validated by comparing it to a set of acceptance criteria. The acceptance criteria are crafted in a way that ensures the data set will be acceptable for its intended use, therefore, the ACA must approve all data acceptance criteria applicable to MCOTEA SIUs.

One way to validate a data set is to compare it to real world data and establish the degree to which the two data sets must match. In some applications, the data input to an M&S comes directly from the real world. For example, if an M&S models the performance of a given radar system, and the M&S uses the antenna pattern obtained from the actual radar it is intended to model, the antenna pattern already represents validated data because it comes directly from the real world system being modeled.

Data can also be validated by comparing it to an analogous real world system/situation, again within the constraints of the approved acceptance criteria. In the example above, the antenna pattern to be used by the first M&S might be generated by another M&S. The computer-generated antenna pattern can be validated by comparing it to the actual antenna pattern. If the computer-generated pattern compares favorably to the real antenna pattern within previously agreed upon acceptable limits (standards), the computer-generated pattern is considered validated for use by the first M&S.

Data validation can be assisted by SME inspection. SMEs view the data under various conditions for reasonableness, based on their experience. The selection of SMEs is important. SMEs should be experts in the warfare area or technical area corresponding to the system being modeled by the M&S or in systems similar to the system being modeled. The SMEs should also have an understanding of the OT&E SIUs corresponding to the data under examination. If SMEs are used for data validation, more than one should be used and they must come to a consensus before the data validation is useful. If the SMEs think the data being validated are reasonable, that strengthens the case for data validation.

In general, in order to be validated, any data used as input to an M&S intended for use by MCOTEA must either be the relevant real world data itself, or it must compare favorably within pre-defined, acceptable limits to the relevant real world data. The ACA shall ensure that all input data intended to support MCOTEA SIUs is validated against the approved acceptance criteria. The data acceptance criteria shall be defined in the Accreditation Plan, and the validation shall be documented in the V&V Report and the Accreditation Report.

Use of Surrogate Data

It is always preferable to use the data explicitly required by the M&S; however, occasionally the data required as input to an M&S may not exist. In such cases similar data may exist and can be used to approximate the desired M&S input data. For example, controlled data on how human skin reacts to heat (human burn data) might be hard to find or might not exist. However, controlled experiments dealing with how animal skin reacts to heat does exist. In this case, it will be necessary to run the M&S based largely on the animal data, then extrapolate the M&S output to effects on humans.

The extrapolation of the surrogate data is part of the model, so it (the surrogate data and the extrapolation technique) must be V&V'd. Evidence supporting the verification and validation of the surrogate data and the extrapolation technique shall be included in the Accreditation Report.

Marine Corps Operational Test and Evaluation Activity
2032 Barnett Avenue
Quantico, VA 22134-5014

Acronyms

Bibliography

Acronyms

Note: This list accounts for acronyms in both volumes of this manual.

AAP	Abbreviated Acquisition Program
ACAT	Acquisition Category
ACMC	Assistant Commandment of the Marine Corps
ACOR	Assistant Contracting Officer's Representative
ADM	Acquisition Decision Memorandum
ANOVA	Analysis of Variance
A_o	Operational Availability
AoA	Analysis of Alternatives
ARL	Army Research Laboratory
ASN(RDA)	Assistant Secretary of the Navy (Research, Development, and Acquisition)
BDAR	Battle Damage Assessment and Repair
BH&T	Ballistic Hull & Turret
C2	Command and Control
C4ISR	Command, Control, Communications, Computers, Intelligence and Reconnaissance
CAC	Common Access Card
CAE	Component Acquisition Executive
CBRN	Chemical, Biological, Radiological, and Nuclear
CBT	Component Ballistic Testing
CD&I	Combat Development & Integration
CDD	Capability Development Document
CDT	Controlled Damage Testing
COE	Concept of Employment
COI	Critical Operational Issue
COMOPTEVFOR or COTF	Commander Operational Test and Evaluation Force
CONOPS	Concept of Operations
COR	Contracting Officer's Representative
COS	Chief of Staff
CPD	Capability Production Document
CRB	Consolidated Review Board
CRTC	Cold Regions Test Center
CSSTD	Combat Service Support Division
CTEIP	Central Test and Evaluation Investment Program
DC	Data Collector
DC V&V	Data Collection Verification and Validation
DC, CD&I	Deputy Commandant, Combat Development and Integration

DIRLAUTH	Direct Liaison Authorized
DM	Data Manager
DOD	Department of Defense
DON	Department of the Navy
DOT&E	Director, Operational Test and Evaluation
DT	Developmental Test
DT&E	Developmental Test and Evaluation
DTS	Defense Travel System
EDP	Event Design Plan
EOA	Early Operational Assessment
EOAR	Early Operational Assessment Report
ESOH	Environmental, Safety, and Occupational Health
ETD	Expeditionary Division
FD/SC	Failure Definition/Scoring Criteria
FOS	Feasibility of Support (Message)
FOT	Follow-On Operational Test
FOTP	Follow-On Operational Test Plan
FRP	Full-Rate Production
FSR	Fleet Support Request
FUSL	Full-up System Level
GCTD	Ground Combat Division
GOTS	Government off the Shelf
GRS	General Records Schedule
I&I	Inspector and Instructor
IA	Information Assurance
IAR	Intermediate Assessment Report
ICD	Initial Capabilities Document
IOT	Initial Operational Test
IOT&E	Initial Operational Test and Evaluation
IOTP	Initial Operational Test Plan
IPR	In-Process Review
IPT	Integrated Product Team
ITT	Integrated Test Team
JCIDS	Joint Capabilities Integration and Development System
JCTD	Joint Capabilities Technical Demonstration
JITC	Joint Interoperability Test Command
JPO	Joint Project Office
JT&E	Joint Test and Evaluation
KPP	Key Performance Parameter

KSA	Key System Attribute
LCCE	Life Cycle Cost Estimate
LFSEP	Live Fire System Evaluation Plan
LFT&E	Live Fire Test and Evaluation
LRIP	Low-Rate Initial Production
LTI	Limited Technical Inspection
M&S	Modeling and Simulation
MAGTF	Marine Air-Ground Task Force
MAGTF C4ISR TD	MAGTF C4ISR Division
MARFOR	Marine Forces
MaxCMT	Maximum Corrective Maintenance Time
MCCLL	Marine Corps Center for Lessons Learned
MCL	Mission Capability Level
MCMT	Mean Corrective Maintenance Time
MCOTEA	Marine Corps Operational Test and Evaluation Activity
MCSC	Marine Corps Systems Command
MCTL	Marine Corps Task List
MDA	Milestone Decision Authority
MDD	Materiel Development Decision
MEF	Marine Expeditionary Force
mef	Mission Essential Function
MFR	Memorandum for the Record
MOE	Measure of Effectiveness
MOIC	Marine Officer In Charge
MOP	Measure of Performance
MOS	Measure of Suitability
MOSur	Measure of Survivability
MOT	Multi-Service Operational Test
MOT&E	Multi-Service Operational Test and Evaluation
MOTP	Multi-Service Operational Test Plan
MOU	Memorandum of Understanding
MR	Maintenance Ratio
MS	Milestone
MTBF	Mean Time Between Failure
MTBOMF	Mean Time Between Operational Mission Failure
MTBUM	Mean Time Between Unscheduled Maintenance
MTTFL	Mean Time To Fault Locate
NARA	National Archives and Records Administration
NET	New Equipment Training

NMCI	Navy Marine Corps Intranet
O&MMC	Operations & Maintenance, Marine Corps
OA	Operational Assessment
OA	Operations Analyst
OAG	Operations Advisory Group
OAP	Operational Assessment Plan
OAR	Operational Test Agency Assessment Report
OCI	Organizational Conflict of Interest
OE	Operational Effectiveness
OER	Operational Test Agency Evaluation Report
OFER	Operational Test Agency Follow-on Evaluation Report
OIC	Officer in Charge
OMAR	Operational Test Agency Milestone Assessment Report
OMF	Operational Mission Failure
OMS/MP	Operational Mission Summary and Mission Profile
OpT	Operating Time
ORM	Operational Risk Management
ORSA	Operations Research and System Analysis
OS	Operational Suitability
OSD	Office of the Secretary of Defense
OSur	Operational Survivability
OT&E	Operational Test and Evaluation
OTA	Operational Test Agency or Operational Task Analysis
OTICC	OSD Test Investment Coordinating Committee
OTPO	Operational Test Project Officer
OTRB	Operational Test Readiness Board
OTRR	Operational Test Readiness Review
PEO-LS	Program Executive Officer – Land Systems
PGD	Product Group Director
PM	Program Manager
$P_{mission}$	Probability of Mission Success
POA&M	Plan of Action and Milestones
POC	Point of Contact
POM	Program Objective Memorandum
PoPS	Probability of Program Success
PQDR	Product Quality Deficiency Report
PT	Pilot Test

QRA	Quick Reaction Assessment
QRT	Quick Reaction Test
R	Reliability
RAM	Reliability, Availability, and Maintainability
RDC	Rapid Development Capability
RDT&E	Research, Development, Test, and Evaluation
RFP	Request for Proposal
ROM	Rough Order of Magnitude
RSO	Range Safety Officer
RTF	Regional Transportation Facility
RTM	Requirements Traceability Matrix
RTT	Requirements Transition Team
S&T	Science and Technology
SA	Scientific Advisor
SAP	System Assessment Plan
SAR	System Assessment Report
SECNAVINST	Secretary of the Navy Instruction
SEP	System Evaluation Plan
SIT	Systems Integration Test
SITREP	Situation Report
SME	Subject Matter Expert
SoS	System of Systems
SSIC	Standard Subject Identification Code
STA	System Threat Assessment
T&E	Test and Evaluation
T&E BOD	T&E Board of Directors
T&E WIPT	Test and Evaluation Working-level Integrated Product Team
TEIN	Test and Evaluation Identification Number
TEMP	Test and Evaluation Master Plan
TES	Test and Evaluation Strategy
TIR	Test Incident Report
TM	Test Manager
TMC	Test Management Counsel
TMO	Traffic Management Office
TRMC	Test Resources Management Center
VV&A	Verification, Validation, and Accreditation
WIPT	Working-level Integrated Product Team
WNRC	Washington National Records Center
WSERB	Weapon System Explosive Safety Review Board

This page intentionally left blank for two-sided printing.

Bibliography

ATEC et al. 2012. *Memorandum of Agreement on Multi-Service Operational Test and Evaluation and Operational Suitability Terminology and Definitions*. 2012.

Blanchard, Benjamin S. and Wolter J. Fabrycky. 1990. S*ystems Engineering and Analysis*, 2nd edition. New Jersey: Prentice Hall.

Chairman of the Joint Chiefs of Staff. 2012. *Joint Capabilities Integration and Development System*. CJCSI 3170.01H.

Clemen, Robert T. and Reilly, Terence. 2001. *Making Hard Decisions with Decision Tools*. Pacific Grove: Duxbury Press.

Defense Acquisition University. 2012. *Defense Acquisition Guidebook*. Virginia: Defense Acquisition University Press.

———. 2011. *Glossary of Defense Acquisition Acronyms and Terms*, 14th edition. Virginia: Defense Acquisition University Press.

———. 2005. *Test and Evaluation Management Guide*, 5th edition. Virginia: Defense Acquisition University Press.

Department of Defense. 2007. *The Defense Acquisition System*, DODD 5000.01.

———. 2012. *Department of Defense Dictionary of Military and Associated Terms.*

———. 1987. *Distribution Statements on Technical Documents*, DODD 5230.24.

———. 2008. *Documentation of Verification, Validation, and Accreditation for Models and Simulations*, MIL-STD-3022.

———. 2005. *Guide for Achieving RAM*.

———. 2007. *Modeling and Simulation Management*, DODD 5000.59.

———. 2003. *Information Assurance Implementation*, DODI 8500.2.

———. *DOD Modeling and Simulation Verification, Validation, and Accreditation*, DODI 5000.61.

———. *DOD M&S Catalog.*

———. 2008. *Operation of the Defense Acquisition System*, DODI 5000.02.

———. 2004. *Procedures for Interoperability and Supportability of Information Technology and National Security Systems*, DODI 4630.8.

———. 1982. *RAM Primer*.

———. 2009. *Reliability, Availability, Maintainability, and Cost Rationale Report Manual.* Washington, DC: Office of the Secretary of Defense.

———. 2000. *System Safety and Program Requirements*, MIL-STD-882D.

———. 2001. *VV&A Recommended Practices Guide*.

——— *1987. Distribution Statements on Technical Documents, DODD 5230.24.*

———. 1995. *Withholding of Unclassified Technical Data from Public Disclosure*, DODD 5230.25.

———. 2005, Guide for Achieving RAM. DODI 4151

Department of the Navy. 2012. *Department of Defense—Navy Assurance for the Protection of Human Research Subjects.* Assurance Number: DoD N-40074. Expiration Date: 30 June 2015. Institution: Marine Corps Operational Test and Evaluation Activity.

Department of the Navy. 2005. *Department of the Navy Correspondence Manual*, SECNAVINST 5216.5D CH-2.

———. 2004. *Modeling and Simulation Verification, Validation, and Accreditation Implementation Handbook*, Volume 1.

———. 2008. *Naval PoPS Criteria Handbook*, Version 1.0.

———. 1996. *Unit Training Management Guide*. MCRP 3-0A.

Fields, Andy. 2005. *Discovering Statistics Using SPSS*, 2nd edition. Thousand Oaks: Sage Publications, Inc.

Giadrosich, Donald L. 1995. *Operations Research in Test and Evaluation*. Washington, DC: American Institute of Aeronautics and Astronautics, Inc.

Joint Staff. 2011. *Universal Joint Task List*. CJCSM 3500.04F.

Joint Test and Evaluation Methodology. 2009. *Analyst's Handbook for Testing in a Joint Environment.*

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Kirkwood, Craig W. 1997. *Strategic Decision Making: Multiobjective Decision Analysis with Spreadsheets*. Belmont: Duxbury Press.

Klass, Gary. 2008. *Just Plain Data Analysis: Finding, Presenting, and Interpreting Social Science Data*. New York: Rowman and Littlefield Publishers.

Marine Corps Combat Development Command. 2008. *Rapid Engagement Precision Rifle Capabilities Production Document*. Quantico: Marine Corps Combat Development Command.

Marine Corps Operational Test and Evaluation Activity. April 2012. *MCOTEA Policy and Procedures for Research with Human Subjects*. MCOTEA Policy Letter 1-12. Quantico.

Marine Corps Systems Command. 2007. *Project Team Leader's Pocket Guide*, Version 1.3. Quantico.

National Research Council. 1998. *Statistics, Testing, and Defense Acquisition: New Approaches and Methodological Improvements*. Washington, DC: National Academy Press.

NIST/SEMATECH e-Handbook of Statistical Methods. 2012. http://www.itl.nist.gov/div898/handbook, April 2012.

Office of the Secretary of Defense. 2008. *Memorandum*: *Definition of Integrated Testing*.

———. 2010. *Memorandum*: *Guidelines for Operational Test and Evaluation of Information and Business Systems.*

Operational Test and Evaluation of Defense Acquisition Programs, U.S. Code 10 § 2399. 2008.

Parnell, Gregory S. 2007. *Value-Focused Thinking*. In *Methods for Conducting Military Operational Analysis.* Andrew G. Loerch and Larry B. Rainey, 619-655. Military Operations Research Society.

Rossi, Peter H., Mark W. Lipsey, and Howard E. Freeman. 2004. *Evaluation: A Systematic Approach*., 7th edition. Thousand Oaks: Sage Publications, Inc.

Scriven, M. 1991. *Evaluation Thesaurus*, 4th edition. Thousand Oaks: Sage Publications, Inc.

Secretary of the Navy. 2002. *Department of the Navy Modeling and Simulation Management*, SECNAVINST 5200.38A.

Secretary of the Navy. 1999. *Verification, Validation, and Accreditation of Models and Simulations*, SECNAVINST 5200.40.

———. 2007. *Department of the Navy, Navy Records Management Program Records Management Manual*, SECNAVINST 5210.1.

———. 2011. *Implementation and Operation of the Defense Acquisition System and the Joint Capabilities Integration and Development System*, SECNAVINST 5000.2E.

———. 2009. *Navy Enterprise Test and Evaluation Board of Directors*, SECNAVINST 3900.44.

———. 2008. *Standard Subject Identification Code Manual*, SECNAV M-5210.2.

———. 2004. *Verification, Validation, and Accreditation of Models and Simulations*, SECNAVINST 5200.40A.

———. 2002. *Department of the Navy Modeling and Simulation Management, SECNAVINST 5200.38A.*

———. 2008. *Implementation and Operation of the Defense Acquisition System and the Joint Capabilities Integration and Development System, SECNAVINST 5000.2D.*

Surgeon General. 8 June 2012. Department of Defense-Navy Assurance for the Protection of Human Research Subjects.

Surviac, last accessed on May 2011. www.bahdayton.com/surviac/

Tufte, Edward R. 1996. *The Visual Display of Quantitative Information*. Cheshire: Graphics Press.

United States Army Training and Doctrine Command and US Army Material Command. 1987. *RAM Rationale Report Handbook. Alexandria, VA.*

United States Congress. 2008. Major systems and munitions programs: survivability testing and lethality testing required before full-scale production, 10 U.S. Code 10 § 2366.

United States Congress. 2008. *Operational Test and Evaluation of Defense Acquisition Programs, U.S. Code 10 § 2399.*

United States Marine Corps. 1992. *HQMC Supplement to the Department of the Navy (DON) Manual*, MCO 5216.20.

———. 1996. *The Marine Corps Technical Publications System*, MCO P5215.17.

———. 2007. *Universal Naval Task List* Version 3.0, MCO 3500.26A.

———. *2010. USMC Integrated Test and Evaluation Handbook*.

University of Chicago. 2010. *The Chicago Manual of Style*, 16th edition. University of Chicago Press.

U.S. Army Training and Doctrine Command and US Army Material Command. 1987. *RAM Rationale Report Handbook*. Alexandria, VA.

U.S. Congress. 2008. *Major systems and munitions programs: survivability testing and lethality testing required before full-scale production*, 10 U.S. Code 10 § 2366.

Winer, B.J. 1971. *Statistical Principles in Experimental Design Second Edition.* McGraw-Hill Inc, New York, NY.

Marine Corps Operational Test and Evaluation Activity
2032 Barnett Avenue
Quantico, VA 22134-5014