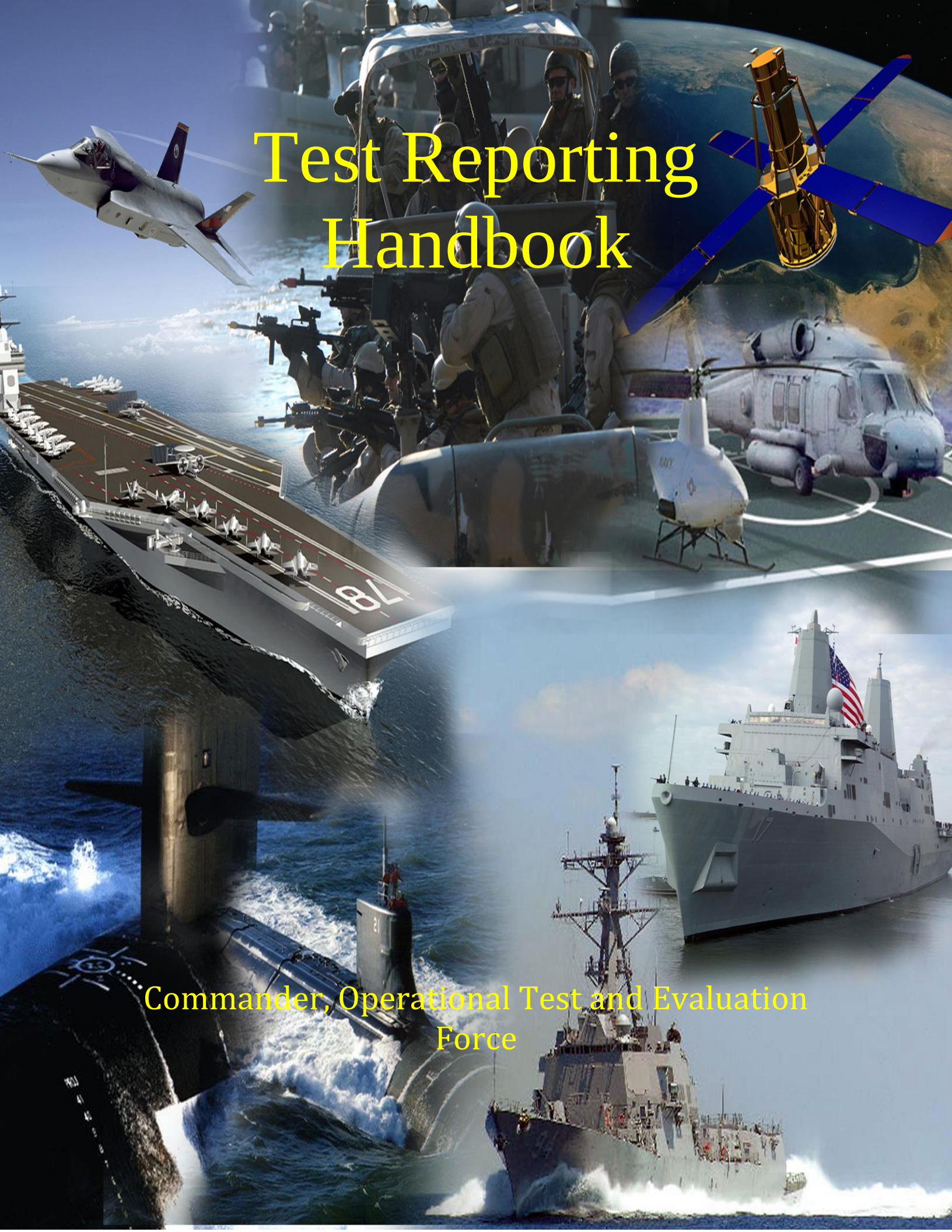
Test Reporting Handbook



Commander, Operational Test and Evaluation Force

Version 2.0

04 February 21

FRAENKEL.RICHARD.JOSE.1136730416

Digitally signed by FRAENKEL.RICHARD.JOSE.1136730416
Date: 2021.02.04 17:06:29 -05'00'

RECORD OF REVISIONS

Number of Change	Summary of Changes	Updated
1	This is the initial Test Reporting Handbook	11 May 20
2	Incorporates Cross Domain Working Group recommendations, updates ACTD terminology to JCTD, replaces Chiclet Chart with PMT View, updates Scoring Board guidance, further addresses Data Analysis Summary requirements, and establishes better linkages to the Cyber Survivability T&E Handbook.	04 February 21

THIS PAGE INTENTIONALLY LEFT BLANK

Forward

To succeed in our mission to "evaluate warfare capabilities in realistic combat environments with Fleet warfighters to ensure Naval Forces can fight and win," Operational Test and Evaluation Force (OPTEVFOR) must translate evaluative results into accurate, concise, understandable, and useful reports to the Fleet, Warfare Centers of Excellence, Office of the Chief of Naval Operations (OPNAV) Resource Sponsors, and Acquisition Authorities.

This handbook was prepared to help the OPTEVFOR military, government civilian, and support contractor team produce evaluation reports that communicate System Under Test (SUT) capabilities and limitations to meet the needs for both the Warfighter and acquisition community. The overarching objective is to impartially provide operational truth to the Fleet based on defendable test results.

This handbook describes Commander, Operational Test and Evaluation Force's (COMOPTEVFOR) process for preparing test reports. Complexity of acquisition programs vary widely. Consequently, test teams are expected to tailor reporting content to the needs of their particular program, working in collaboration with program offices, Warfare Division leadership, and OPTEVFOR Competencies. Unique aspects of the test reporting process dealing with cyber survivability are covered in the Cyber Survivability Test and Evaluation (T&E) Handbook.

THIS PAGE INTENTIONALLY LEFT BLANK

Test Reporting Handbook

TABLE OF CONTENTS

TABLES

FIGURES

Introduction

The test report communicates COMOPTEVFOR's conclusions regarding a system's operational effectiveness, operational suitability, and cyber survivability and recommendations regarding Fleet introduction, further development, additional Operational Test and Evaluation (OT&E), etc. Evaluations of operational effectiveness, operational suitability, and cyber survivability are made based on the contribution of the SUT to the System of Systems (SoS) warfighting effectiveness.

Test reports are COMOPTEVFOR's most important contribution to the acquisition process. Test reports inform acquisition decisions by articulating the effectiveness, suitability, and cyber survivability of new systems and capabilities. COMOPTEVFOR reports are also intended to clearly communicate the capabilities of newly acquired systems, and any operational risk being transferred to the Fleet, to Fleet users and Warfare Centers of Excellence. In addition, test reports provide a historical record of testing.

COMOPTEVFOR policy regarding reporting is contained in the OT&E Manual. The guidance contained in this handbook serves as a foundation for test report preparation. This handbook covers the timeframe from the test team's return-to-base following completion of test events to test report signature and issuance. Additional report guidance may be found on the OPTEVFOR share drive: Y:\OT&E Production Library\Test Reports. Unique aspects of the test reporting process dealing with cyber survivability are addressed in the Cyber T&E Handbook.

Test team members preparing to undertake their first operational test are encouraged to attend the Blue and Gold Sheet Writing Course and the Post-test Iterative Process Course approximately three to nine months prior to test commencement.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 1 - Types of Operational Evaluation and Other Reports

1.1 REPORT TYPES

Several types of reports are listed in table 1-1 and are further described below.

Table 1-1. Report Types
UNCLASSIFIED

Report Type	Test Type	Purpose	Format
OAR	EOA/OA	Early involvement OT reports used in identifying system enhancements and significant areas of risk to the program's successful completion of IOT&E in the form of risk sheets. OARs are assessment reports that support all stakeholders, but do not support specific MS decisions.	Full Report (EOA/OA Report templates)
OMAR	EOA/OA	Early involvement OT reports used to identify system enhancements and significant areas of risk to the program's successful completion of IOT&E in the form of risk sheets. OMARs are assessment reports used to support MS decision meetings.	Full Report (EOA/OA Report templates)
OER	IOT&E	To report a full, complete phase of testing. Consists of a report letter signed by the Commander, a deficiency letter signed by the Warfare Division Director, and a data analysis summary memorandum signed by the Technical Director.	Full Report (IOT&E Report Template)
OFER	FOT&E	To report a full, complete phase of testing. Consists of a report letter signed by the Commander, a deficiency letter signed by the Warfare Division Director, and a data analysis summary memorandum signed by the Technical Director.	Full Report (FOT&E Report Template)
VCD	VCD	To report results for validating correction of specific deficiencies (specific COIs only) from previous testing (end-to-end testing may not be required).	Report (VCD Report template)
QRA	QRA	To report findings for operational considerations/system capabilities when it is necessary to achieve a rapid capability in the Fleet. QRAs do not replace formal OT&E. They are used to support a rapid deployment of a capability to the Fleet.	Report (QRA Report template)
Interim Report	EOA/OA/ IOT&E/ FOT&E	Report provided when evaluation results are required prior to publication of the full OT report. The report provides the status of testing, an assessment of available data, and a recommendation (if appropriate). Use of this report is at the Commander's discretion. The full formal report is still required.	Report (Interim Report template)
DT Assist LOO	DT Assist	Per the memorandum of agreement coordinated between the PM and the Warfare Division Director.	Letter with enclosed Blue/Gold risk sheets (LOO template)
Letter	AOC	Per the memorandum of agreement coordinated between the PM and the Warfare Division Director.	Letter with enclosed Blue/Gold deficiency sheets (AOC letter template)
MUA, LMUA, or OUA	JCTD	Products for the JCTDs that provide an assessment of military utility demonstrated. Not to be used for acquisition programs.	Full JCTD Report

AOC-Assessment of Operational Capability
COI-Critical Operational Issue
DT-Developmental Test
EOA-Early Operational Assessment
FOT&E-Follow-on Test and Evaluation
IOT&E-Initial Operational Test and Evaluation
JCTD-Joint Capabilities Technology Demonstration
LMUA-Limited Military Utility Assessment
LOO-Letter of Observation
MS-Milestone
MUA-Military Utility Assessment
OA-Operational Assessment
OAR-OTA Assessment Report
OER-OTA Evaluation Report
OFER-OTA Follow-on Evaluation Report
OMAR-OTA MS Assessment Report
OT-Operational Test
OTA-Operational Test Agency
OT&E-Operational Test and Evaluation
OUA-Operational Utility Assessment
PM-Program Manager
QRA-Quick Reaction Assessment
VCD-Verification of Correction of Deficiencies

1.2 OPERATIONAL TEST AGENCY EVALUATION REPORT (OER) AND OPERATIONAL TEST AGENCY FOLLOW-ON EVALUATION REPORT (OFER)

For IOT&E and FOT&E, system evaluations of operational effectiveness, operational suitability, and cyber survivability, are made on the contribution of the SUT to the SoS' warfighting effectiveness. A separate operational effectiveness and suitability evaluation may be provided for the SoS capability to perform its mission in the operational environment only when there is sufficient data to conclude the SoS performance differs from the SUT conclusion. A fielding recommendation is provided in the OER or OFER.

1.3 OPERATIONAL TEST AGENCY ASSESSMENT REPORT (OAR) AND OPERATIONAL TEST AGENCY MILESTONE ASSESSMENT REPORT (OMAR)

EOAs and OAs, whether conducted as stand-alone OT, combined Developmental Test (DT)/OT, or an Integrated Test (IT), assess a program's progress towards a successful IOT&E and Fleet introduction. The OAR or OMAR commonly supports Defense Acquisition Boards or MS decision meetings.

1.4 OBSERVING DT

A DT Assist letter of observation (LOO) is used to communicate with the program manager when accomplishing a DT Assist. This feedback is in the form of observations of system performance using the DT Assist LOO format. The format for DT Assist LOOs is a brief letter to the PM with attached Blue/Gold risk sheets for each performance issue identified.

An Assessment of Operational Capability letter is used to document an assessment of a SUT's operational capability observed during DT when there is no future phase of Operational Test (OT) planned. The format for an AOC letter is intended to be a brief letter to the PM and other stakeholders with attached Blue/Gold deficiency sheets for each performance issue identified.

1.5 QUICK REACTION ASSESSMENT (QRA)

If a QRA is required to assess the risk(s) associated with fielding a rapid deployment capability, the program sponsor will initiate a request to Chief of Naval Operations (CNO) N94, copy to COMOPTEVFOR, or the QRA will be identified as a phase of test in the Master Test Strategy (MTS) for Middle-Tier Acquisition (MTA) systems. Once approved, OPTEVFOR will conduct the QRA and the Commander will issue a report as soon as possible. A post-test quick look brief to the Commander may also be warranted for high visibility and urgent QRAs. A QRA will not take the place of a formal OA or IOT&E and will not be used to resolve COIs, make effectiveness, suitability, or cyber survivability calls, or provide a limited Fleet introduction, Fleet introduction, or Fleet release recommendation. A QRA should answer the questions and address the purpose as outlined in the QRA request letter or the Master Test Strategy for Middle-Tier Acquisition systems. As such, the QRA request letter is routed with the test report as the report is staffed for signature. See the Test Planning Handbook for QRA test planning. Information from a QRA may be used by Director, Operational Test and Evaluation (DOT&E) in

support of what has been referred to as a "Section 231" or "Early Fielding" Report to Congress[1] when a system being developed is fielded prior to the completion of IOT&E.

1.6 VERIFICATION OF CORRECTION OF DEFICIENCIES (VCD)

VCDs provide a more rapid response to Department of the Navy (DoN) leadership on the status of the correction of deficiencies than could be provided by waiting for the next scheduled period of OT. A VCD is generally not a preplanned phase in the Test and Evaluation Plan (TEMP), but can be incorporated into the test program at the request of the program manager after a formal phase of OT to validate certain deficiencies have been corrected. No TEMP update is required. A test plan is required to describe the specifics of the given test: what data will be collected, how data will be collected, and how the data will be analyzed/used to determine if the original deficiency has been corrected or mitigated to such an extent as to merit recharacterization. See the Test Planning Handbook for VCD test planning. The VCD report results should indicate whether the deficiencies are corrected; not corrected, but substantially mitigated; or not corrected. For non-DOT&E oversight programs, when COI resolution is discussed in the test plan and if the VCD results enable a change to the resolution of COIs (beyond IOT&E), then those updated COI resolutions will be listed in the VCD report, thereby reducing the scope or eliminating the need for later phases of OT for the specific purpose of verifying the deficiency has been corrected. For programs on DOT&E oversight, the only permitted change in COI resolution during a VCD phase of test is from satisfactory (SAT) to unsatisfactory (UNSAT).

[1] Early fielding reporting requirements pertain to a decision to operationally use or procure a major defense acquisition system beyond low-rate initial production prior to completion of operational testing. These requirements are Public Law, specified in Section 2399(b) Title 10 U.S. Code.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 2 - Evaluation Process

2.1 DISCUSSION

The evaluation process is a standardized, repeatable evaluative process across all warfare domains, for all reports (with the exception of DT Assist LOOs and AOC letters) to classify performance issues, characterize risks/deficiencies, make overall COI assessments or resolutions, and make recommendations for each issue. This process, to include COI Evaluation Working Groups (CEWG), the optional Analysis Working Group (AWG), the optional Blue and Gold Sheet (B&G) Peer Review, System Evaluation Review Board (SERB), and the Executive SERB (E-SERB), is presented below in chronological order. The post-test iterative process (PTIP) checklist found in Appendix A is designed to guide test teams through the process in a step-by-step fashion.

2.2 DURING TEST/INITIAL PERFORMANCE ISSUE IDENTIFICATION

2.2.1 Blue and Gold Sheet Drafting

Upon identification of a potential issue, the Blue and Gold sheet templates are used to document all SUT and SoS performance issues. There is only a single performance issue per sheet and each sheet must stand-alone. Appendix B provides guidance on drafting Blue and Gold Sheets. Cyber survivability Blue and Gold sheets are initially drafted by 01D per the Cyber Survivability T&E Handbook.

2.2.2 Sharing Draft Blue and Gold Sheets

As system performance issues are identified, the raw data and the issue are provided to the PM per the Test Operations Handbook. Draft Blue and Gold sheets are shared with the PM, upon concurrence from the first O-6 in the test team's chain-of-command, and must be clearly marked as preliminary information with the DRAFT watermark on each page and the standard draft disclaimer below on the front page. "Note: This is based on limited initial analysis of the available data. Further data may refine and/or modify the final characterization of the preliminary deficiency [or risk], and will be addressed in the final Deficiency/Risk letter to the Program Manager."

This preliminary feedback to the PM does not include any COI preliminary determinations or recommendations, as the evaluative process is immature. In keeping with the tenets of constructive conflict, inputs from key stakeholders are sought. Feedback to/from the PM is important for several reasons:

- Enables the PM to begin addressing performance issues identified as early as possible.
- Provides insight to the OT team as to causal analysis.
- Identifies additional data that may be available for system evaluation by the OT team.

It is important to note the PM is not being given permission to edit or change Blue and Gold sheets. Instead, the PM is being given an opportunity to provide additional information for consideration by COMOPTEVFOR during the PTIP and the drafting of the Test Report. Warfare Division or Squadron leaders must manage their test teams and ensure each Blue or

Gold sheet is shared with the program office as soon as practicable, on a not-to-interfere basis with test execution. The above does not prevent test teams from sharing issues as they are discovered during test (e.g., daily Situational Reports (SITREPS) and post event summaries).

2.2.3 Post-Test Reviews

- **Scoring Board.** The purpose is to review the data collected and to validate whether the data are suitable for OT use and for addressing the measures delineated in the test plan or Data Collection Plan (DCP). For long duration tests consisting of multiple test periods, scoring boards should occur during test, particularly at the completion of stand-alone test periods on a not-to-interfere basis with test execution. Scoring boards are the first step in the iterative process of data analysis and reporting. Chapter 3 provides details for the conduct of scoring boards.

- **Test Completion.** Once it is clear all necessary data collection is complete and all data have been received, the end of test message or e-mail is issued. Normally, end of test will occur no later than 30 days following completion of the last test event. If more time is required for data reduction and scoring, the Warfare Division Director/Deputy Director is required to brief the Technical Director (00TD) on the planned timeline for test completion.

- **CEWGs.** The purpose of the CEWGs is to provide a systematic process for ensuring measure analyses are conducted within the context of their respective COIs, to review draft risk/deficiency sheets, and to support a balanced COI evaluation. Chapter 4 provides details regarding the conduct of CEWGs.

- **AWG.** At the Warfare Division Director's discretion, convene an AWG to validate data analysis is accurate and complete. Chapter 5 provides details for the AWG process.

- **B&G Peer Review.** At the Warfare Division Director's discretion, convene the B&G Peer Review to examine issue(s) and determine the categorization of risk/deficiency level prior to publishing. A running B&G Peer Review may be conducted during test to classify issues early, and to reduce the scope of the PTIP for large-scale tests. 00TD should be invited to attend B&G Peer Reviews for controversial findings. Chapter 6 provides details for the B&G Peer Review

- **SERB.** A SERB and E-SERB are conducted to thoroughly review the report letter COI resolutions, operational effectiveness, operational suitability, and cyber survivability determinations, and Operational Considerations (OPCON). 00TD should be invited to attend SERBs for controversial findings. Chapter 7 provides details on the test report letter. Chapter 8 provides details of the SERB process.

- **Draft the OT Report Documents.** Drafting the OT report documentation commences before the start of testing with creating initial rough draft report documents by tailoring the report templates for the program concerned. As data are analyzed by the test team during the post-test iterative process, division/squadron analysts and technical writers should review data, calculations, and initial drafts of Blue/Gold sheets associated with measures and issues. Once the CEWG(s), AWG (optional), and B&G Peer Review (optional) are complete, the OTD finalizes the rough draft of the Deficiency/Risk letter, the Data Analysis Summary

memorandum for the record, and the Test Report letter. The Deficiency/Risk letter containing Blue and Gold sheets is published by the Warfare Division Director to the PM and the Data Analysis Summary memorandum is signed by 00TD. Blue and Gold sheets and the Data Analysis Summary are made available upon request. Once the SERB process is complete, the OTD finalizes the smooth draft of the Test Report letter to obtain the Commander's signature at the conclusion of the E-SERB. The OPTEVFOR report templates are located in the Y:\OT&E Production Library. Use of templates are mandatory for COMOPTEVFOR reports. See the Cyber Survivability T&E Handbook for details on how post-test products related to cyber are generated and provided to a warfare division.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 3 - Scoring Board

3.1 DISCUSSION

Scoring boards play a key role in a credible, consistent, and collaborative process that produces results that are relevant to both the acquisition decision maker and the warfighter. To be successful, boards must be conducted in an atmosphere of trust, transparency, and teamwork. Several meetings are usually necessary to address all the data. The number of meetings and their periodicity are initially set with the PTIP POA&M produced during the test planning process. The POA&M should be updated during and after test as required to reflect changes to the timing of the data collection/receipt. As soon as all data for any given COI has been scored, the test team should prepare for and conduct a CEWG for the COI on a not-to-interfere basis with data scoring for other COIs. Cyber survivability scoring boards described in the Cyber Survivability T&E Handbook are conducted separate from effectiveness and suitability data scoring boards. The following guidance is provided to ensure all participants have a clear understanding of roles and responsibilities for scoring effectiveness and suitability data.

3.1.1 Scoring Board Purpose

Primarily, the scoring board serves to qualify data for OT, and in so doing supports the decision to call end-of-test. The scoring board serves the following purposes:

- Review test execution and identify instances where actual execution deviated from the test plan's Detailed Method of Test (DMOT). Test plan deviations must be documented and the impact on the validity of the collected data determined by the scoring board. A test plan deviation is any modification to the approved test execution, data collection, and/or data analysis described in the test plan or subsequent test plan change letter. Deviations such as adjusted data collector assignments or an amended sequence in which events are executed are examples of deviations that do not affect COI resolution and are therefore not limitations to test. These deviations will be documented as deviations in the test report's Data Analysis Summary. Deviations which affect the adequacy of test, such as resource constraints or changes to controlled conditions resulting in the loss of runs associated with a response variable, or changes that impact the expected confidence level associated with critical tasks, result in limitations to test. These deviations shall be documented as limitations to test in the test report's Data Analysis Summary and possibly the Test Report Letter.

- Validate data elements were collected and recorded in accordance with the Test Plan:
  - Design of Experiment (DOE) runs were executed per the controlled conditions and within established tolerances identified in the DOE run matrix.
  - Modeling and simulation involved with data generation or collection was accredited.
  - SUT was in the proper configuration.
  - SUT, SoS, and threat operators were operationally representative, qualified, and trained.
  - SUT was employed/stressed in an operationally representative manner.

- Identify missing, incomplete, or unusable data. Each "data gap" should be assessed for its impact to the assessment/evaluation of the affected COI(s) or cyber survivability. If the data

gap impacts the assessment/evaluation it is a limitation to test. The Warfare Division will need to determine if the data gap can be documented as a test limitation without mitigation to allow the post-test process to be completed without additional data collection. If not, the data gap test limitation will need to be mitigated through additional data collection. The additional data collection does not necessarily require additional operational testing. Depending on the circumstances, it may be possible to qualify Developmental Testing or Fleet data for operational test use.

- Score the result of the observed task, if needed (hit/miss, pass/fail, Operational Mission Failure (OMF)/not, abort/not, etc.). Identify response variable data requiring coordination with OPTEVFOR Test Design Competency (01B) Statistical Analysis Cell per the Analysis Handbook.

3.1.2 Scoring Board Membership

Scoring boards provide a venue for the division and/or squadron conducting the test to review the data collected with the headquarters test competency division staff. To support this effort, a variety of outside organizations may be invited to participate; however, responsibility for execution of a scoring board rests with the cognizant Warfare Division Deputy Director or Squadron Chief Operational Test Director (COTD) conducting the particular test. Appendix A contains a notional listing of Scoring Board participants. Enclosure (2) of the test plan for a given phase of test should be referred to as it may tailor this participant list based on project specifics.

3.1.2.1 Scoring Board Chair Person

The scoring board will be chaired by either the Division Section Head (SH) for tests being conducted by the Warfare Division or the Squadron Branch Head/Department Head for tests being conducted by a squadron. The chair is responsible for conducting the board per the OT&E Manual and this handbook. The chair will adjourn the proceedings and notify the Division Director/Squadron Commanding Officer (CO) anytime he/she feels COMOPTEVFOR processes are being compromised.

3.1.2.2 External Organization Participation

The presence of senior individuals from other organizations in no way relieves the designated chair from responsibility for conduct of the event. Senior individuals from outside OPTEVFOR are to be treated with the courtesy due their positions and are expected to respect COMOPTEVFOR processes.

3.1.3 Scoring Board Rules of Engagement

Materials used to prepare for, presented at, or produced by an OPTEVFOR scoring board are to be clearly marked "Predecisional - Not for External Release." Personally Identifiable Information collected in conjunction with surveys and interviews will only be provided to and reviewed by Operational Test Agency (OTA) and DOT&E scoring board participants. Participants must understand the results of the scoring board are not final until the Data Analysis Summary memorandum is signed by 00TD.

Military personnel and government civilians are expected to conduct themselves professionally at all times. In the rare case where a non-government participant may fail to comply with the procedures governing the scoring board, the chair should raise the issue with the appropriate Contracting Officer Representative (COR) so the matter can be properly handled.

Outside participants are encouraged to express their view with the same candor as members of OPTEVFOR. Support contractors and particularly professional staff members from Federally Funded Research and Development Centers such as the Center for Naval Analyses and the Institute for Defense Analyses bring unique experience and technical knowledge to the proceedings; however, it is important for all participants to understand such individuals are not permitted to speak on behalf of the Government. In a similar vein, care must be taken that nongovernment personnel are not implicitly tasked to perform services. Any requests for additional support/analyses from these individuals must come from their respective COR, not from the scoring board chair or other participants.

In keeping with the purpose of the board, frank, open dialogue is encouraged. While one would expect that in most cases participants will quickly reach consensus based upon a common understanding of the facts, there will be times when individuals will examine the same set of facts and draw different conclusions as to the validity and/or utility of a set of data.

Individual members of the test team and the headquarters staff are encouraged to speak freely and openly. Intellectual disagreements among members of the OPTEVFOR team should be anticipated. Given that free and open discussion, all participants must take particular care not to represent the positions expressed by individual members of OPTEVFOR as the views of the Commander. All deliberations of the scoring board are considered internal OPTEVFOR staff discussions.

There is no desire for a forced consensus. The tendency for "Groupthink" must be studiously avoided. Mature organizations cannot only deal with differences in technical/operational judgment; they will in fact become better from the self-examination that is part of the professional discourse.

In the event that, after an appropriate period of discussion it is evident there is an irreconcilable difference of professional opinion between the test team and the test competency staff, the chair will close the discussion and propose a synopsis of the differences to be placed in the record. Once both parties agree their views have been captured, the chair will continue with the next topic of discussion.

Following completion of the scoring board, the chair will provide the cognizant Division Director/Squadron CO and the Director for Test Planning and Evaluation with the record of the proceedings taking special note of any open areas of disagreement. The Division Director/Squadron CO and the Director for Test Planning and Evaluation will then resolve any open items.

3.1.4 Scoring Board Conduct

Procedures to prepare for and conduct the Scoring Board are contained in Appendix A.

3.1.5 Data Sharing

COMOPTEVFOR is obligated to share all test data with DOT&E (for oversight projects) and all factual data with the applicable program office as it is collected and/or becomes available on a not-to-interfere basis with test execution. At the conclusion of the scoring board, confirm the appropriate data have been shared with the DOT&E Action Officer and the SUT's Program Office.

In the event additional data is collected or obtained to augment in-hand scored data at any point after the last Scoring Board, the test team shall:

- Consider documenting the additional data collection as a test plan deviation.
- If the program is on the DOT&E Oversight List, provide all additional data to the DOT&E Action Officer.
- Score the additional data.
- Expeditiously share the additional raw factual data with the SUT's Program Office.

CHAPTER 4 - CEWG

4.1 DISCUSSION

The COI Evaluation Working Group (CEWG) is intended to be a comprehensive review of each effectiveness and suitability COI, including data scoring (if not yet completed), data analysis, Blue and Gold sheets, COI results paragraphs, and operational considerations (OPCONs). The purpose is to produce balanced COI resolutions and assessments, made within the context of the test design and test scope, using scored data. The review takes place within the Warfare Division/Squadron, supported by the OPTEVFOR Test Design (01B) and Test Planning and Evaluation (01C) Divisions. It can be structured to fit the needs of the specific program. Several meetings are usually necessary to address all the COIs. The number of meetings and their periodicity are initially set with the PTIP POA&M produced during the test planning process. The POA&M should be updated during and after test as required to reflect the actual amount of data collected and the number of potential risks or deficiencies identified in test. As soon as all data for any given COI has been scored, the test team should prepare for and conduct a CEWG for the COI on a not-to-interfere basis with data scoring for other COIs.

The processes used to develop cyber survivability report products are similar to the CEWG processes applied to effectiveness and suitability data. The Cyber Survivability T&E Handbook details the differences. 01D will analyze the cyber survivability data and provide the OTD a cyber survivability data analysis summary input. Unlike effectiveness and suitability Blue and Gold sheets, cyber survivability Blue and Gold sheets, drafted by 01D and provided to the OTD, may document multiple vulnerabilities as part of a single system deficiency impacting one or more mission areas. In addition, 01D may provide the OTD with Blue and Gold sheets that document issues not traced to a system mission impact. Finally, 01D will provide the OTD a cyber survivability results paragraph vice a COI results paragraph.

4.2 CEWG PREPARATIONS

Actions required to prepare for a CEWG are outlined in Appendix A. The following procedures pertain to documenting the evaluation of SUT and SoS operational effectiveness and operational suitability. The associated documents will be reviewed during the applicable CEWG.

4.2.1 Data Analysis

Data shall be analyzed in accordance with the Analysis Handbook and the Suitability Handbook and documented in the Data Analysis Summary. The Post-Test Iterative Process (PTIP) Checklist in Appendix A details the steps required to complete the analysis. The Data Analysis Summary template posted in the Y:\OT&E Production Library provides guidance on how to document the analysis. The Data Analysis Summary should:

- Provide the authoritative details to reflect testing has been completed in accordance with the approved test plan;
- Establish credibility of reported results by "showing the homework" for calculations used in the SUT/SOS evaluations;
- Provide the basis to rationalize the OT&E conclusions documented in the report and/or deficiencies in the event those findings should be challenged by external agency;

- Serve as a source of 'applied learning.' This pertains to gaining equity from each test and design. Each test design and test plan should make use of knowledge learned from its predecessor. For example, if you are designing a test for the next generation of a given SUT, use the analyzed results from the predecessor phase during factor selection to avoid unnecessary resources where possible. If you are designing for a type of SUT that could benefit from prior test results, regardless of Test and Evaluation Identification Number (TEIN), the data analysis summaries for those other SUTs should be used as a resource.
- Include recommendations for future improvements. This could be anything from analysis tools, to workforce knowledge needs, to alternative analysis methods. Be sure to provide supporting rationale for your recommendations and make them as actionable as possible.
- Be understandable to a reasonably qualified external reader.

The format for presenting test results and supporting data may be modified by test teams to effectively and efficiently present their analyzed test results and supporting data. Teams should embed Excel spreadsheets, Power Point presentations, screen shots, or video files, if doing so meets the intent of the Test Results and Supporting Data section, particularly if it builds upon the products used to collect, score, and analyze data. Regardless of your solution, a reader unfamiliar with your test must be able to find and interpret the data, and if necessary, recreate your results. Therefore, describing the analytical method(s) used may be required, to include defining any mathematical formulas or statistical methods used outside of Excel. Additionally, if using Excel, be consistent in how each tab is formatted.

Measures are addressed, COI-by-COI, in the data section of the Data Analysis Summary in the following order:
- Critical measures
- Measures linked to risks/deficiencies
- Remaining non-critical measures

For each measure state/address:
- Measure number, name, and threshold or criterion
- Test date/location/test conditions applicable to the measure
- Results with supporting data (the data necessary to calculate the result)
- Response variable regression analysis. The RV analysis outbrief shall be embedded in the Data Analysis Summary. In addition, the following RV analysis outbrief information shall be addressed in the RV's measure paragraph:
  - List the factors that had an effect on the RV.
  - For each factor or factor interaction that had an effect, quantitatively describe that effect so that an operator can use this information.
  - Note: Discussion of the operational impact of factor effects included in the RV analysis outbrief is intended to inform the associated COI Results paragraph and should not be included in the RV's measure paragraph.
- Confidence Interval or Limit (refer to the Analysis Handbook for guidance on calculating confidence intervals)
- Analytical methodology. Describe the calculations and applied analytical techniques. Detailed discussion of analytical method is not required for calculations with a standard methodology, such as the mean, median, or standard deviation.

- If there is a threshold or criterion, state whether the threshold or criterion was met, or not met.

The cyber data summary provided to the OTD by 01D will include all data requirements collected and vulnerabilities identified per the Cyber Survivability T&E Handbook.

Properly annotate all data, including as a minimum:
- Units of measure for all data,
- Charts, graphs, and figures are titled,
- Chart axes are labeled with variable definition and units of measure,
- Keys are provided to interpret the use of colors or shapes,
- Notes and amplifying remarks as required to properly characterize the data.

4.2.2 Risks/Deficiencies

Performance issues relating to the SUT are documented in "Blue Sheets." Performance issues relating to the SoS are documented in "Gold Sheets." SUT performance issues are those issues directly linkable to what the sponsor has funded the PM to develop and field. SoS issues are those issues that, while not traceable back to the required SUT capability to be delivered, impact the mission accomplishment of the SUT when operating in the SoS environment. Performance issues will be identified as Severe, Major 1, Major 2, Major 3, or Minor risks or deficiencies, depending on the operational impact/mission relation and the phase of test.

Blue and Gold sheet templates and the report formats using the Blue and Gold Sheets are posted in the Y:\OT&E Production Library. The format of these sheets are based on the Six Part Paragraph (6PP) style of writing as described in Appendix B. The intent is for the reader to gain a comprehensive understanding of the issue by reading this single sheet. There will be a single Blue or Gold Sheet for each performance issue identified during testing. A unique number is assigned to each performance issue. The performance issue will be updated as new OT-qualified data are acquired using the same Blue or Gold Sheet, and the unique number will take on a modifier. These stand-alone risk/deficiency sheets are used for initial performance issue identification and continuously updated through verification of correction (cradle to grave). These sheets are used in the CEWG and B&G Peer Review for risk/deficiency level categorization, and inform COI Assessment/Resolution decisions at the SERB and E-SERB. The Blue and Gold sheets are approved by the Warfare Division Director. Issues characterized as Severe are significant enough to drive an UNSAT COI resolution or Red COI assessment and shall be briefed to the Commander prior to approval.

4.2.3 COI Results

The COI results paragraph is the first paragraph under each COI results section and is constructed using the following flow. The Test Report templates in the Y:\OT&E Production Library provide detailed guidance on formatting and content.

4.2.4 Opening Sentence (past tense)

Establish the test conditions. Repeat, verbatim, the COI question in the affirmative or answer format with a verb form "was evaluated" or "was assessed." State what was being evaluated and

within the predicate of the sentence, provide a high-level summary of how the COI was evaluated.

4.2.5 Middle Section (past tense)

This part of the results paragraph presents overall test results and clearly explains and supports the satisfactory or unsatisfactory conclusion stated at the end of the paragraph. Focus should be on results and not test method. The goal is to provide a Fleet focused, clear understanding of the results associated with the COI's critical tasks defined within the test plan. Explain what was successful and what was not. Go further and explain what those results mean to the completion of the mission/COI. Operational impact of factor effects described in the applicable RV analysis outbrief is intended to inform this discussion. Severe and Major risks/deficiencies should be discussed in some detail, describing their operational impact/mission relation to the COI. A comparison to the legacy system(s) may be included, if helpful. Specifically address positive outcomes as enhancing characteristics when the observed performance improved SUT capability as compared to legacy systems or added new capability to the Fleet. The discussion should be centered on the performance of critical tasks and not overly focused on specific quantitative or qualitative critical measure results (which are listed in the Test Report Letter's Major Test Results table(s) preceding the COI Results Paragraphs). The end product should clearly communicate to the reader what capabilities were successfully demonstrated and what were not and their impact to mission accomplishment. The discussion should be a subjective assessment of COI risk (EOA/OA) or resolution (IOT&E/FOT&E) by comparing adverse results against the full scope of the COI. In the end, the reader should come away with a clear view of the positive test outcomes versus the negative outcomes and understand why the scales tipped to either the positive (satisfactory) or negative (unsatisfactory). Past tense transition sentences may be used as needed. If there are additional minor deficiencies for the COI, provide a transition sentence (e.g., "X additional minor deficiencies listed in table 1-4 were noted and are available upon request"). The transition sentence can be included where appropriate within the paragraph or following the conclusion.

4.2.6 Closing Sentence (present tense)

Conclusion. Again, state verbatim the COI question in answer form with a conclusion. The conclusion is along the lines of: "[The COI question in statement form] is evaluated as satisfactory (or unsatisfactory)." When writing the report for an OAR or OMAR, the results paragraph for the COI risk assessment will follow the same general format as above; however, the presentation will be modified to address program or system risk as presented in the following section.

4.2.7 COI Risk to IOT&E/FOT&E Assessments for EOA and OA

4.2.7.1 General Risk Discussion

The EOA and OA reports include risk assessments as part of the Test Report letter and the associated Risk letter containing the Blue or Gold Risk Sheets. Additionally, the EOA and OA reports provide an overall risk assessment of the capability of the SUT and SoS to perform required missions in the intended operational environment. The risk analysis methodology described in the paragraphs below provides the OTD with a transparent, repeatable, and defendable COI risk assessment process that will identify for the PM areas of risk that should be addressed to ensure successful completion of IOT&E or FOT&E. The result of the risk

assessment is a prioritized list of risks for each COI. For significant risks tipping the COI scales to moderate or high risk, a separate assessment of the adequacy of the program office's mitigation plan may be conducted. The risk assessment and the risk mitigation assessment are then subjectively combined to determine the overall risk to successful resolution of the COI at IOT&E/FOT&E. The subjective COI assessment is made by comparing all known adverse risk projected to IOT&E against the full scope of the COI. Consideration is given to the fact that during the EOA/OA, the full scope of each COI may not be assessed and, therefore, unknown at the time of the EOA/OA. The team then weighs projected risk along with unknown performance against the full scope of the COI. Although the ultimate COI assessment is subjective, it is based on objective risks projected to IOT&E and objective COI scope identified in the Mission Based Test Design (MBTD). The rationale for the subjective assessment of the COI is included in the COI results paragraph. Risk assessment methodology is described in the paragraphs below.

4.2.7.2 Overall COI Risk Assessment

The results paragraph for each COI will reflect a roll up of all known risks with a best understanding of potential mitigation of those risks at IOT&E compared to the full scope of the COI as identified in the MBTD. The following examples are provided to reflect the extreme possibilities available to the CEWG and SERB:

4.2.7.2.1 Red COI Risk Assessment

A COI will be assessed RED if a single Severe or Major 1 risk projected to IOT&E is so negative that it dominates the full scope of the COI or,

4.2.7.2.2 Yellow or Green Risk Assessment

A COI may be assessed YELLOW or GREEN if several Severe or Major risks projected to IOT&E are either:

- Isolated to a small subset of the full scope of the COI and will not, in the aggregate, dominate COI performance or;
- The team assesses the risks will not be realized as the risks will be corrected or mitigated prior to IOT&E.

The judgment of the SERB SME assessment of overall COI risk will be presented at the E-SERB for final determination or approval by the Commander.

Overall COI risk assessment, as approved during the E-SERB, will be presented as one of the following:

- High – red
- Moderate – yellow
- Low – green
- Not assessed – white (For COIs that cannot be assessed as a result of system immaturity or lack of information.)

4.2.8 COI Resolution at IOT&E and FOT&E

OPTEVFOR addresses the resolution of COIs by satisfying the questions posed by the COIs. Derived from the MBTD process, the test plan will provide an audit trail from the COI questions through the critical mission tasks to the critical measures. This trail provides a logical flow path so the disposition of COIs is directly related to the evaluation of each task. Thus, when a test

parameter is quantitative, the COI resolution is based on actual results relative to the operational threshold. For non-quantifiable parameters, the COI resolution must be based on two factors: (1) observed results and (2) operational experience and judgment. Additionally, the number and severity of the deficiencies and their cumulative/aggregate impact on mission performance associated with the COI must be considered for COI resolution. The resolution of COIs should be a subjective assessment of COI results by comparing adverse results against the full scope of the COI. In the end, the case should be clearly made to support weighing the positive outcomes versus the negative outcomes for the critical mission tasks and subtasks. The audience should come away with a firm understanding as to why the scales tipped to either the positive (satisfactory) or negative (unsatisfactory). See figure 4-1. These conclusions will be presented to the SERB for validation and final approval by the Commander at the E-SERB.

4.2.8.1
COIs are resolved as follows:

4.2.8.1.1 Resolved
The COI was tested and resolved either SAT or UNSAT. When a COI has been resolved UNSAT, the severe and/or major (1, 2, or 3) deficiencies that caused the UNSAT resolution must be discussed in the Test Report letter. A severe or major (1, 2, or 3) deficiency can impact other COIs and the deficiency can be used to resolve additional COIs UNSAT. The analysis and evaluation will determine the most appropriate primary COI. The deficiency will be reported under the primary COI where it has the main operational impact. The impact to other COIs affected by the deficiency will be discussed in those COI results paragraphs.

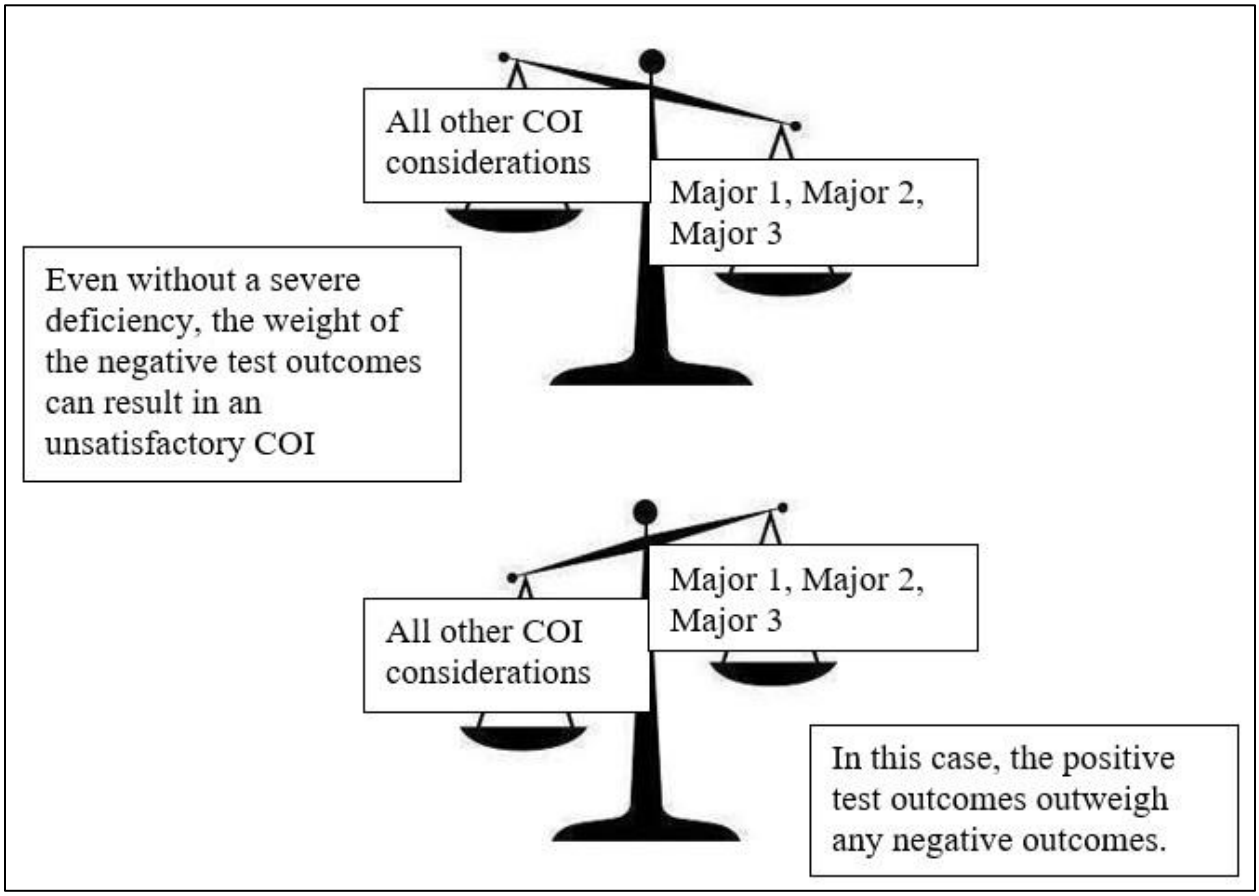
4.2.8.1.2 Unresolved
Used when a COI requires further testing for final resolution due to a major or severe limitation(s). This is used when the COI has been tested, but cannot be resolved**.**

4.2.8.1.3 Split Resolution
Split resolution is used when the COI was tested and resolution is not a singular determination. In these instances, split resolution will be used to clearly communicate the differing conditions impacting the COI resolution. A COI may be split to resolve SUT performance in one condition as SAT and in another condition as UNSAT. Examples include:

- SAT in an Electronic Attack (EA)-clear environment, UNSAT in a EA environment (threat)
- SAT in sea state 4 or below, UNSAT above sea state 5 (environment)
- SAT in APG-XX configuration, UNSAT in APG-YY configuration (SUT configuration)

Additionally, a split COI resolution may be used to communicate when part of the COI is either SAT or UNSAT, and part is unresolved due to a test limitation.

Figure 4-1. Examples of Possible COI Resolution
UNCLASSIFIED



4.2.8.1.4 Not Tested

Used only when the COI was not tested during the particular phase of testing in which it was an issue for resolution. This may be due to the absence of a key test resource that poses a limitation to the test for the COI or it may be due to a decision by OPNAV N94 to defer testing of one or more aspects of the SUT until a future test period.

4.2.9 OPCONs

Tactical employment considerations are by-products of the IOT&E. OPCONs document tactical considerations that inform operational commanders of significant aspects (pro and con) of system employment, or make clear what special measures would be required to make the system more efficient in operational use. Operational impact of factor effects described in the applicable RV analysis outbrief may inform this discussion. An OPCON is a recommendation for the user to consider in the operational employment or management of the SUT and/or SoS. It is not a deficiency by another name. The OPCON paragraph is structured for making recommendations to the operational commander and includes supporting data or test observation. When used this way, OPCONs serve as the starting point for the OPTEVFOR Tactics Guides (for air warfare projects), and tactics inputs for Commander, Naval Surface Force (for surface programs) and the Undersea Warfare Development Center (for undersea programs).

4.3 CEWG MEMBERSHIP

The CEWG membership is as follows:

- OTD – required
- SH/OTC – required
- Lead Test Engineer (LTE) – required
- Warfare Division/Squadron Analyst – required
- 01C Representative – required
- 01B Core Team Facilitator (CTF) – required for COIs with Response Variables (RV); as available otherwise
- Center for Naval Analysis (CNA) Analyst (if assigned) – required
- OTD's Support Contractor(s) – optional.

4.4 CONVENING A CEWG

The OTD or SH/OTC coordinates with the 01C Representative, and 01B CTF to schedule CEWGs when the Division/Squadron is ready to support the review. The OTD must provide read-ahead materials electronically to all CEWG members no later than 2 working days prior to the scheduled CEWG, to include the applicable portion of the Data Analysis Summary, any Blue and Gold Sheets associated with the COI under review, the COI results paragraph, and applicable OPCONs.

4.5 CONDUCT OF THE CEWG

The procedure to conduct the CEWG is contained in Appendix A. When convened, the CEWG is led by the OTD, with the review facilitated by the LTE (or the OTD if no LTE is available). In the event there are differences of professional opinion between the test team and competency division staff, the Warfare Division Deputy Director/Squadron COTD will adjudicate the issue. The CEWG will:

- Review Scoring Board results (if available) or score data per Scoring Board procedures.
- Document and review limitations to test.
- Identify and note the rationale for invalid data which was not used.
- Review the relevant portion of the draft Data Analysis Summary.
  - Review data analysis for all critical measures, per the Test Plan Data Analysis Plan. For RV analysis, use the procedures from the Analysis Handbook.
  - Review data analysis for all non-critical measures.
- Review Blue and Gold sheets.
  - Brief Resource Sponsor concurrence/non-concurrence on derived requirements.
- Review Platform Mission Task (PMT) View (see Appendix C)
- Review COI Results paragraphs.
- Review OPCONS.
- Discuss lessons learned during test execution as they relate to the evaluation/assessment (e.g., test strategy, design, planned scope of testing, etc.) for potential inclusion in the appropriate handbook(s) or a best practice. Highlight any cross-divisional lessons.

4.5.1 CEWG Roles and Responsibilities

4.5.1.1 OTD

With the SH/OTC and LTE, make initial:

- Security classification determination for performance issues identified during test planning, execution, and data analysis,
- Performance issue identification and assessment/evaluation (risk/deficiency level),
- COI resolution and associated rationale.

With the SH/OTC and LTE:

- Document deviations from the approved test plan. Deviations are categorized in one of the following categories:
  - Minor Deviations – Minimal impact to COI resolution. Should be documented as a Minor limitation to test.
  - Major Deviations – Affects and possibly precludes COI resolution. Should be documented as either a Major limitation to test (will affect COI resolution) or a Severe limitation to test (precludes COI resolution).
  - Other Deviations – No impact to the ability to resolve COIs. Should be documented as a deviation.
- Document limitations to test,
- Analyze all measures from the approved test plan,
- Prepare the COI's Data Analysis Summary data section,
- Document performance issues in Blue and Gold sheets,
- Obtain Resource Sponsor concurrence/non-concurrence on derived requirements associated with risks/deficiencies,
- Fill in the PMT View per Appendix C,
- Write COI results paragraph,
- Write OPCONs.
- Document lessons learned during test execution.
- Assess whether or not there is a need to brief external/Fleet stakeholders on test results if the performance of the SUT, as tested, differs significantly from expected/advertised.

4.5.1.2 SH/OTC and LTE

- Assist the OTD in above responsibilities.
- Conduct working-level coordination with the Resource Sponsor.

4.5.1.3 Warfare Division Deputy Director

- Review the products from each CEWG (Data Analysis Summary, Blue and Gold sheets, COI results paragraphs, OPCONs, lessons learned, need for external stakeholder briefings).
- Provide comments and guidance to the test team.
- Adjudicate any disagreements between test team, 01B, and 01C.

4.5.1.4 01C Representative

- Participate in all CEWGs.
- Review the Data Analysis Summary, Blue and Gold sheets, COI results paragraphs, and OPCONs.

4.5.1.5 01B CTF

- Participate in all CEWGs for COIs with response variables (RV).
- Review test execution and analysis related to RVs.

4.5.1.6 Warfare Division/Squadron Analyst and CNA Analyst (if assigned)

- Assist the OTD with data analysis and preparation of the Data Analysis Summary data section
- Ensure analytical rigor supports system evaluations.

4.6 POST-CEWG ACTIONS

4.6.1 Warfare Division Deputy Director

The CEWG is considered complete when the OTD briefs the Warfare Division Deputy Director/Squadron COTD on the results of the CEWG and incorporates any received guidance.

4.6.2 Resource Sponsor Review

OTD or Sections Head confirm the concurrence/non-concurrence on derived requirements have been received from the Resource Sponsor.

4.6.3 Program Manager Review

The Warfare Division is responsible for sharing the draft Blue and Gold sheets with the Program Manager to obtain their comments on the issues prior to final approval by the Warfare Division Director.

4.6.4 RV Analysis Outbrief

Following the completion of a CEWG including a response variable, the OTD provides a copy of the response variable analysis outbrief, and a copy of the draft Data Analysis Summary addressing the response variable to the 01B Director for inclusion in the 01B Statistical Analysis Cell historical database supporting future test designs.

4.6.5 Finalize the Data Analysis Summary

The OTD compiles the complete draft Data Analysis Summary after the last CEWG. The draft Data Analysis Summary is reviewed at an AWG if the Warfare Division Director and/or the Director for Test Planning and Evaluation determine an AWG is warranted. Otherwise, the Warfare Division reviews the draft Data Analysis Summary and routes it to 00TD for approval. The executive summary accompanying a draft Data Analysis Summary containing response variables in the Electronic Document Router should specify the response variable analysis is included in the route folder and that it was reviewed by 01B.

Before the draft Data Analysis Summary reaches 00TD in the Electronic Document Router the OTD, LTE, and test team Lead Analyst (at a minimum) should schedule a 30 minute meeting to brief 00TD on the Data Analysis Summary. The 01B lead analyst should also be invited when response variable analysis is part of the document. A formal briefing with slides is not required nor desired. Instead the OTD should provide an overview of the significant risks/deficiencies

noted during the test and simply talk through the data analysis while highlighting any consequential or potentially questionable measures along the way. Test limitations and test plan deviations should also be discussed. This 30 minute time investment will reduce 00TD's document review time when the Data Analysis Summary reaches him in the Electronic Document Router.

4.6.6 Finalize the Deficiency/Risk Letter

The OTD compiles the draft Deficiency/Risk letter after the last CEWG. The draft Deficiency/Risk letter is reviewed at a B&G Peer Review if the Warfare Division Director determines a B&G Peer Review is warranted. Otherwise, the Warfare Division reviews the draft Deficiency/Risk letter and routes it to the Warfare Division Director for approval.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 5 - Analysis Working Group (AWG)

5.1 DISCUSSION

The AWG is the mechanism for Warfare Division/Squadron leadership and the 01C Division Director to review the draft Data Analysis Summary before the Warfare Division submits it for 00TD approval. The AWG is optional and is convened at the direction of Warfare Division/Squadron leadership. The AWG will validate OT data analysis to ensure accuracy and consistency of data supporting the SERB evaluative process. Appendix A describes the required actions to prepare for and conduct an AWG.

5.2 AWG MEMBERSHIP

The AWG membership is as follows:

- 00TD – as available
- Warfare Division Director or designated representative – required
- VX CO or designated representative – required for VX SUT
- 01C Director or designated representative – required
- Warfare Division Deputy Director – required
- VX COTD or designated representative - required for VX SUT
- SH/OTC – required
- LTE – required
- OTD – required
- Division Analyst – required if assigned
- 01C Representative – required
- 01B CTF – required
- CNA analyst – as available
- OTD's Support Contractor(s) – optional.

5.3 CONVENING AN AWG

When directed, the responsible OTD or SH/OTC coordinates with the 01B CTF and 01C Representative and schedules an AWG prior to the SERB. The AWG should be conducted no less than 5 working days prior to the SERB to allow formal reporting and resolution of data issues discovered during the AWG. The OTD must provide read-ahead materials electronically to all AWG members no later than 2 working days prior to the scheduled AWG, to include the draft Data Analysis Summary, requirements document, approved test plan, and any additional supporting data.

5.3.1 Conduct of the AWG

The AWG focus is on data identified as critical to COI resolution/assessment per the associated test plan and data used to support identified risks/deficiencies. AWG data analysis focuses on:

- All data identified by the test plan as critical for subtask/COI resolution/assessment. This includes the data associated with critical measures. A critical measure is a measure or measures associated with a critical task/subtask. These critical measures include quantitative and qualitative data. They may also include Key Performance Parameters (KPP)/Measures

of Effectiveness (MOE)/Measures of Suitability (MOS) that will be included in the Major Test Results tables of the Test Report letter.
- Data substantiating results discussed in the COI results paragraphs of the report.
- Data used to build the risks/deficiencies.
- Fleet data and any other data the division deems pertinent to risk/deficiency characterization or COI resolution/assessment.

Every AWG starts with a brief overview/description of the SUT and pertinent SoS to aid reviewing analysts to understand system operation and relevance of the test data to be reviewed. Use of the draft SERB brief is encouraged. Following the brief, the OTD, supported by the LTE and test team, will lead the AWG data review using the test plan's traceability matrix, the PMT View, and the draft Data Analysis Summary to include test plan deviations and limitations to test. The review will ensure all data associated with the critical measures are reviewed to include a comparison of the planned number of runs versus the actual number of runs. Any data supporting noncritical measures that are used in Blue and/or Gold sheets should also be examined. Again, the focus of the AWG is data-driven issue identification and evaluation of data-driven measures binned by subtasks associated with COIs.

Following adjudication of AWG comments, the final draft Data Analysis Summary is routed to 00TD for approval.

5.3.2 AWG Roles and Responsibilities

5.3.2.1 01C Representative

- Coordinate with OTD, SH/OTC and LTEs as appropriate and schedule AWGs.
- Participate in all CEWGs prior to the AWG (except cyber survivability CEWGs).
- Collect data review lessons learned and disseminate as best practices.

5.3.2.2 OTD

- With the SH/OTC and LTE:
  - Review of reduced test data supporting data-driven deficiencies and positive results.
  - Review test plan to validate required test data were collected and identify any test data issues (e.g., missing data or data not qualified for OT).
- Coordinate with 01C Representative and schedule an AWG (use telephone conference calls for tests conducted by Air Test and Evaluation Squadrons and to include supporting organizations).
- Prepare the Data Analysis Summary for review by the AWG and a SUT/SoS overview brief. Send both as read-aheads to the AWG membership no later than 2 working days prior to the scheduled meeting.
- Lead the AWG through the data review discussion.
- Adjudicate comments received during the AWG and route the final draft Data Analysis Summary to 00TD for approval.

5.3.2.3 Division Analyst

- If assigned, assist the OTD, SH/OTC and LTE in above responsibilities. As part of the post-test iterative process, the division analyst(s) should already be assisting the OTD and reviewing all data and calculations with the test team. The AWG is the culmination of this effort.
- Prepare supporting OT data for presentation as directed by the OTD, SH/OTC and LTE.
- Participate in AWGs,

5.3.2.4 SH/OTC and LTE

- Assist the OTD in above responsibilities.
- Review test data for completeness, accuracy, deficiencies, or data anomalies that should be identified for AWG discussion and resolution.

5.3.2.5 01B CTF

- Review read-ahead materials prior to the AWG.
- Review associated approved IEF prior to the AWG.

5.3.2.6 Warfare Division Deputy Director

Oversee execution of divisional analytical reviews.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 6 - B&G Peer Review

6.1 DISCUSSION

An optional B&G Peer Review provides the Warfare Division Director a venue for a collaborative review of any CEWG Blue and Gold sheets following Warfare Division Deputy Director review that warrant additional review due to their severity or potentially contentious nature. The collaborative environment is particularly well suited to review contentious or challenging risks/deficiencies. The B&G Peer Review validates the severity characterization of issues identified during OT prior to Warfare Division Director signature and the SERB and E-SERB. Appendix A describes the required actions to prepare for and conduct a B&G Peer Review.

6.2 B&G PEER REVIEW MEMBERSHIP

The B&G Peer Review membership is tailorable depending upon the scope of the review:

- 00TD – invited for controversial findings,
- Warfare Division Director and/or designated representative – required,
- VX CO or designated representative – required for VX SUT,
- Warfare Division Deputy Director – optional
- VX COTD – optional for VX SUT,
- 01C Director or designated representative – required,
- 01D Director or designated representative – required for cyber Blue and Gold Sheets,
- SH/OTC – optional,
- LTE – optional,
- OTD – required,
- Division/Squadron Analyst – optional,
- 01C Representative – optional,
- 01B CTF – optional,
- CNA analyst – optional,
- OTD's Support Contractor(s) – optional

6.3 CONVENING A B&G PEER REVIEW

When directed by the Warfare Division Director, the OTD or SH/OTC coordinates with the 01C Representative and schedules the B&G Peer Review prior to the SERB. The B&G Peer Review should be conducted no less than 5 working days prior to the SERB to allow formal reporting and resolution of issues discovered during the review. The OTD must provide read-ahead materials electronically to all B&G Peer Review members no later than 2 working days prior to the scheduled review, to include those Blue and Gold sheets requiring peer review, the approved test plan, Data Analysis Summary, and any other supporting data or documents, as needed.

6.4 CONDUCT OF THE B&G PEER REVIEW

The B&G Peer Review focus is on those Blue and Gold sheets identified by the Warfare Division that warrant peer review. 00TD should be invited to attend B&G Peer Reviews for controversial findings. While all aspects of the risks/deficiencies will be reviewed, particular attention will be given to the description of the issue, the supporting data, the mission relation,

and the severity characterization. The objective of the B&G Peer Review is concurrence by the membership on the Blue and Gold sheets, supporting the approval and signature by the Warfare Division Director.

6.5 FINALIZING THE DEFICIENCY/RISK LETTER

Following adjudication of B&G Peer Review comments, the final draft Deficiency/Risk letter is routed to the Warfare Division Director for approval.

6.6 ROLES AND RESPONSIBILITIES

6.6.1 01C Representative

- Coordinate with OTD, SH/OTC and LTEs as appropriate and schedule B&G Peer Reviews.
- Participate in all CEWGs prior to the B&G Peer Review (except cyber survivability CEWGs).

6.6.2 OTD

- With the SH/OTC/LTE, update all the Blue and Gold sheets post-CEWG, ensuring all comments have been adjudicated.
- Obtain Warfare Division Deputy Director concurrence and comments on draft Blue and Gold sheets, including post-CEWG comment adjudications.
- Coordinate with 01C Representative and schedule the B&G Peer Review (use telephone conference calls or Video Teleconferences (VTC) for tests conducted by Air Test and Evaluation Squadrons).
- Obtain resource sponsor position for all derived requirements related to Blue sheets.
- Prepare Blue and Gold sheets for review by the B&G Peer Review. Send all Blue and Gold sheets, the approved Test Plan, and the Data Analysis Summary as read-aheads to the B&G Peer Review membership no later than 2 working days prior to the scheduled meeting.
- Step through each draft Blue and Gold sheet during the B&G Peer Review discussion.
- Document all action items from the B&G Peer Review.
- Ensure all B&G Peer Review action items are complete prior to routing Blue and Gold sheets with cover letter for Warfare Division Director approval and signature.

6.6.3 Division Analyst

- Assist the OTD, SH/OTC and LTE in above responsibilities.
- Prepare supporting OT data for presentation as directed by OTD, SH/OTC, or LTE.
- Participate in B&G Peer Reviews.

6.6.4 SH/OTC

- Assist the OTD in above responsibilities.
- For VX SUTs, obtain resource sponsor position for all derived requirements related to Blue sheets.

6.6.5 Warfare Division Deputy/Squadron COTD

Review all CEWG products, provide concurrence, comments and guidance. Prior to the B&G Peer Review, review the adjudication of all comments on Blue and Gold sheets. Coordinate with OTD/SH/OTC/LTEs as appropriate and schedule B&G Peer Review.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 7 - Test Report Letter

7.1 DISCUSSION

OPTEVFOR's evaluative process segregates performance issues discovered during OT into SUT issues and SoS issues. System evaluations of operational effectiveness, operational suitability, and cyber survivability are made on the contribution of the SUT to the SoS' warfighting effectiveness. A separate operational effectiveness, operational suitability, and cyber survivability determination may be provided for the overall SoS capability to perform its mission in the operational environment if warranted. The intent of this guidance is to implement a standardized, repeatable process for OT reporting of all findings, while recognizing every program is unique with subjective judgments based on operational experience being required.

Additionally, an effective evaluation report is balanced. Balance is added to the report, specifically in the Test Report letter, by including discussion of positives and negatives for the SUT. Relate the positives and negatives discussed to the results/outcomes of the critical tasks and measures associated with resolving any given COI. For the Test Report letter, the positive and negatives are associated with the roll-up of missions (COIs) and associated capabilities of the SUT/SoS. Balance should be an outcome of properly explaining why the COI was resolved as satisfactory or unsatisfactory, or in the case of effectiveness, why the SUT is operationally effective or not operationally effective. The same logic applies to the suitability call. Avoid the tendency to focus solely on the deficiencies (or negatives).

7.2 DEFINITIONS

The following definitions for SUT and SoS performance issues apply.

7.2.1 SUT Performance Issues

A SUT is defined by either specified or derived requirements the Navy sponsor has funded the PM to deliver. The SUT evaluation is based on the contribution of the SUT, as defined by specified and derived requirements, to the SoS warfighting capability. SUT performance issues identified during test are characterized as risks (EOA and OA) or deficiencies (IOT&E or FOT&E). SUT performance issues will be used in the risk assessment/resolution of appropriate COIs; SUT operational effectiveness, operational suitability, and cyber survivability determinations; and fielding recommendations.

7.2.2 SoS

A SoS performance issue is any effectiveness, suitability, or cyber survivability issue that is not within the purview of the SUT, yet is necessary for mission accomplishment of the SUT when operating in the overall SoS environment. These include those capabilities:

- Identified as MOE and MOS performance objectives that adversely impact SUT mission accomplishment.
- Required for the full employment of the SUT in its intended overall SoS operating environment (including Joint).

SoS risks/deficiencies will inform operational commanders of significant performance issues that need addressing to achieve full mission capability of the SUT. SoS performance issues will be characterized as SoS risks (EOAs and OAs) or deficiencies (IOT&E or FOT&E). SoS

performance issues will be used in the risk assessment/resolution of appropriate COIs and in determining whether the SUT is operationally effective, operationally suitable, and cyber survivable within the intended SoS. Although sufficient data may be available to determine the SoS is not operationally effective, not operationally suitability, and/or not cyber survivable, OTs are not designed to determine the opposite of this (i.e., that the SoS is operationally effective, operationally suitable, and/or cyber survivable).

7.3 CONCLUSIONS AND RECOMMENDATIONS IN EVALUATION AND REPORTING

At the completion of each phase of OT, COMOPTEVFOR provides conclusions and recommendations to the CNO regarding the system tested via an assessment or evaluation report. There are a number of factors that must be considered before a decision is made to enter into production of a system; OT&E is only one of these many factors. Since COMOPTEVFOR is normally not aware of the status of many of the other issues affecting a production decision, it is inappropriate to comment on production issues based on OT&E alone. Accordingly, no conclusion or recommendation pertaining to production should appear in the evaluation report. The guidelines for determining the key elements of the conclusions and recommendations, based on the results of testing, are:

7.3.1 EOA/OA

7.3.1.1 Conclusions

Prior to IOT&E, conclusions are presented as identification of system enhancements and of risks toward effectiveness and suitability COI resolution at IOT&E.

7.3.1.2 Recommendations

COMOPTEVFOR recommendations for EOA/OA phases of testing will be “is” or “is not” recommended for continued program development.

7.3.2 IOT&E

7.3.2.1 Conclusions

Conclusions in IOT&E *must* be definitive (i.e., effective or not effective, suitable or not suitable, cyber survivable or not cyber survivable). Conclusions of operational effectiveness, operational suitability, and cyber survivability are made for the SUT when it is operating within its intended SoS as it affects the SUT’s capability to deliver the required warfighting capability. Sufficient data should be collected and an evaluation conducted to resolve all COIs. If that is not feasible, the program’s test strategy plan should be reviewed and modified as required before commencing the test.

7.3.2.1.1 Operational Effectiveness

The evaluation of operational effectiveness is always anchored on whether the system supports mission accomplishment. Whether a system meets requirement thresholds is informative, but on its own, is not persuasive. A system may meet no requirements but be useful or it may meet all the requirements and be useless. A good rule of thumb is: will the system make the warfighter more effective than he/she was before. Conclusions normally address overall system

effectiveness. However, in those cases where the system tested had effectiveness issues in discrete warfare area(s) (air, submarine, surface, etc.), mission(s), or environment(s) (e.g., jamming), or in several threat regions, the system should be evaluated in each area or threat region and conclusions provided that address effectiveness in each area. Characterize the system's performance regarding where, or under what conditions, the system was or was not effective (e.g., effective in a non-EA environment, effective against specific threat class, or undetermined against other threat class, etc.). The following are basic definitions to be used during the evaluative process when determining operational effectiveness.

7.3.2.1.1.1 Operationally Effective

Ideally, all effectiveness COIs were satisfactorily resolved and there were no severe or major (1, 2, or 3) deficiencies. However, through the evaluative process, it is possible for the system to be determined operationally effective with one or more major (1, 2, or 3) deficiencies and/or one or more unsatisfactory COI resolutions. If as a result of deferrals or limitations to test, there are COIs or portions of COIs that remain unresolved/not tested, characterize the system effectiveness as accurately as possible and recommend additional OT&E to resolve these areas.

7.3.2.1.1.2 Not Operationally Effective

If the E-SERB concludes on balance that sufficient effectiveness COIs were not resolved as satisfactory due to severe or major (1, 2, or 3) deficiencies, then the system is not effective. Regardless of the SUT performance when compared to the KPPs and the Key System Attributes (KSA), if the operator is unable to successfully employ the system to accomplish the mission, or cannot accomplish the mission as well as with the legacy system, it will be deemed not operationally effective.

7.3.2.1.2 Operational Suitability

7.3.2.1.2.1 Operationally Suitable

The Availability COI is central to the SUT suitability assessment and evaluation, and is supported by the Reliability, Maintainability, and Logistic Supportability COIs. In order to be operationally suitable, the SUT must be available, and to be available, the SUT must be reliable, maintainable, and logistically supportable. If as a result of deferrals or limitations to test, there are COIs or portions of COIs that remain unresolved/not tested, characterize the system suitability as accurately as possible and recommend additional OT&E to resolve these areas.

7.3.2.1.2.2 Not Operationally Suitable

If the E-SERB concludes, on balance, that the operational availability of the SUT did not support mission accomplishment, the system is not operationally suitable. Regardless of the SUT performance when compared to the KPPs and the KSAs, if the operator is unable to successfully maintain and sustain the system to deliver the required warfighting capability, it will be deemed not operationally suitable.

7.3.2.1.3 Cyber Survivability

A system's cyber survivability decision is based on a determination of whether or not, in an operational context, a system survives and operates after exposure to cyber threats. This determination is based on weighing the technical feasibility of system exploitation and the

associated mission impacts of the SUT's capability to prevent, mitigate, and recover from cyber threats. Cyber survivability conclusions are discussed further in the Cyber Survivability T&E Handbook.

7.3.2.2 Recommendations

A recommendation regarding Fleet introduction is obligatory if the system is intended for Fleet use, or to support the Full Rate Production Decision Review, or if the TEMP requires it. COMOPTEVFOR addresses Fleet introduction as follows:

7.3.2.2.1 Fleet Introduction

If the system is concluded to be operationally effective, operationally suitable, and cyber survivable, Fleet introduction will normally be recommended. This recommendation may be made contingent on completing specified actions to correct major (1, 2, or 3) deficiencies observed in IOT&E including, if appropriate, verification of correction in FOT&E.

7.3.2.2.2 Continued Fleet Introduction

In those cases where the system under test has already been fielded/released to the Fleet the recommendation should be for "continued Fleet introduction" or not.

7.3.2.2.3 Limited Fleet Introduction

Limited Fleet introduction can sometimes be recommended if IOT&E results are not generally satisfactory, and it has been concluded the system is not operationally effective, not operationally suitable, and/or not cyber survivable, but there is some benefit to the Fleet by introducing the system in limited quantities to specified units. This recommendation will almost always be made contingent on completion of corrective actions, and may be made contingent on demonstrating those corrective actions in a subsequent phase of IOT&E or FOT&E. When recommending limited Fleet introduction, the conditions that must be satisfied before Fleet introduction should be specified, and will ordinarily include FOT&E whenever system design changes are necessary. The effectiveness, suitability, and cyber features to be demonstrated in FOT&E must be specified. Whenever possible, a recommendation for limited Fleet introduction should specify to what level of units the introduction should be made (e.g., units required for next phase of OT&E, air squadrons operating in specific scenarios, etc.).

7.3.2.2.4 No Fleet Introduction

A recommendation against Fleet introduction may be made if it has been concluded the system is not operationally effective, not operationally suitable, and/or not cyber survivable. A not recommended for Fleet introduction decision is validated during the SERB and approved at the E-SERB.

7.3.3 FOT&E

7.3.3.1 Conclusions

The conclusions drawn in FOT&E will address the system's operational effectiveness, operational suitability, cyber survivability, and Fleet introduction if Fleet introduction was not recommended at IOT&E, or no IOT&E was conducted. When the FOT&E is being conducted to examine the integration of a system into other platforms or aircraft, the conclusion will address

the system's operational effectiveness, operational suitability, and cyber survivability in the platform or aircraft tested and, if applicable, Fleet introduction of the system in the platform or aircraft. In those cases where the FOT&E is conducted to examine an upgrade to a system already in production or release of an improved software revision, the conclusion will address the operational effectiveness, operational suitability, and cyber survivability of the system with the upgrade or new software, and Fleet introduction of the upgraded system or Fleet release of the new software version.

7.3.3.2 Recommendations

- A recommendation regarding Fleet introduction should be made if a recommendation for Fleet introduction has not been made in previous OT&E.

- In those cases where the purpose of FOT&E is to verify correction of deficiencies and/or to complete deferred testing and Fleet introduction was already recommended in the IOT&E final test report, a recommendation regarding continued Fleet introduction is appropriate.

- In those cases where FOT&E is to examine the integration of a system into other platforms or aircraft, or to examine an upgrade to a system already in production, a recommendation regarding Fleet introduction is obligatory.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 8 - SERB

8.1 DISCUSSION

The SERB is a review of PTIP products and provides a peer and senior OPTEVFOR leadership review of all system performance issues identified during test execution and data analysis. The SERB provides a repeatable process for evaluation of COIs along with their respective SUT and SoS performance issues to ensure OT reporting fairly evaluates the SUT and identifies SoS issues impacting the full realization of the SUT capabilities. The SERB results will be briefed to the E-SERB for approval or guidance within 5 working days of SERB completion (based on the Commander's availability). Appendix A describes the required actions to prepare for and conduct SERBs.

8.2 SERB MEMBERSHIP

The SERB membership is as follows:

- 00TD invited for controversial findings,
- Warfare Division Director or designated representative*,
- VX CO or designated representative* (if VX SUT),
- 01C Director or designated representative*,
- Warfare Division Deputy Director,
- 01D Director or designated representative (for test phases with cyber survivability),
- SH/OTC,
- LTE,
- OTD*,
- Warfare Division/Squadron Analyst,
- 01B CTF,
- 01C Representative,
- CNA Analyst (if available),

* Minimum requirement for SERB to be convened.

8.3 CONVENING A SERB

The Warfare Division Deputy Director schedules a SERB as soon as possible, but no later than 5 working days following the last CEWG (or AWG or B&G Peer Review, whichever is later). The OTD must provide read-ahead materials to all SERB members no later than 2 working days prior to the scheduled SERB, to include the approved Blue and Gold sheets, approved Data Analysis Summary, the draft Test Report letter containing draft COI Results paragraphs and draft OPCONs, and the SERB/E-SERB Brief slides; see Y:\OT&E Production Library for the current SERB brief template.

8.4 CONDUCT OF THE SERB

The purpose of the SERB is to conduct a review of the quantitative and qualitative results, the rationale for COI assessment or resolution, any OPCONs, the overall conclusions regarding

operational effectiveness, operational suitability, and cyber survivability, and any recommendations.

- The SERB/E-SERB Brief slides should be presented to the SERB using the template from Y:\OT&E Production Library.

- Following the review of the test results, the OTD will lead the evaluative discussion starting with the COI resolution methodology and the logic behind the resolution call using the PMT View, followed by a high-level review of performance issues for each respective COI identified during test using the summary table of Blue and Gold sheets, and finish the COI discussion with a review of the COI Results paragraph and any OPCONs. After all COI discussions are complete, the draft Test Report letter, to include the operational effectiveness, operational suitability, and cyber survivability conclusions and Fleet introduction recommendation, or continued program development recommendation, are reviewed. The OTD will ensure all directed adjustments to the COI Results paragraphs and OPCONs are documented, as well as the consensus or lack of consensus between the Division Director, the VX Commander (when appropriate), and the 01C Division Director.

- Discuss Fleet stakeholder engagement plans and briefing requirement(s). If testing results differ significantly from expected, this may need to occur before the report is released. Discuss who these stakeholders are, what needs to be briefed, and when briefs need to occur

8.5 POST-SERB DATA SHARING

SERB approved draft COI results paragraphs and the effectiveness, suitability, and cyber conclusions will be shared, by e-mail, with the O-6 level PM by the division or squadron O-6 leadership (A-Code or squadron CO/COTD). The e-mail should include the E-SERB scheduled date (normally 5 working days) and a request for comments from the PM. The division or squadron O-6 leadership shall inform the Commander of the results from the discussions with the cognizant PM(s) and afford the Commander the opportunity to engage the PEO.

8.6 EXECUTIVE SERB (E-SERB) BRIEF TO THE COMMANDER

The purpose of the E-SERB is to inform the Commander concerning SUT and SoS issues, COI assessment/resolution, overall conclusions, and associated recommendations.

8.6.1 Outcome of the E-SERB

Successful completion of an E-SERB yields:

- The Commander's concurrence on:
  - COI resolution/assessment,
  - Operational Considerations (if required),
  - Operational Effectiveness, Operational Suitability, and Cyber Survivability determinations,
  - Fleet introduction or program development recommendation.

Successful completion of an E-SERB means the Commander concurs with the COI Results, the overall test conclusions, and any recommendations. The objective of the E-SERB is for the

Commander to sign the Test Report letter at the end of the brief. If the Commander opts to sign the letter at a later time, E-SERB action items will be documented in formal minutes published by the Division and routed with the test report.

8.7 SERB AND E-SERB ROLES AND RESPONSIBILITIES

8.7.1 OTD

With the SH/OTC and LTE:

- Draft the test report.
- Lead the SERB discussion.
- Document the SERB results and brief to the Commander.
- Document E-SERB action items in formal minutes.

8.7.2 SH/OTC and LTE

- Assist the OTD in above responsibilities.
- Review and assist in drafting the test report.

8.7.3 Warfare Division Deputy Director

- Schedule own warfare division SERBs, including the external participants and an appropriate conference room.
- Participate in own warfare division SERBs.
- Review the test report.
- Participate in SERBs for tests outside own warfare division, as requested.
- Prepare the test report email.

8.7.4 Warfare Division Director

- Conduct coordination with the PM and the Resource Sponsor.
- Chair the SERB.
- Following the SERB, share draft COI results paragraphs and the effectiveness, suitability, and cyber conclusions with the PM and solicit feedback.
- Participate in E-SERB brief to COMOPTEVFOR.
- Approve and promulgate the E-SERB action item minutes.
- Review the test report.

8.7.5 VX CO (as appropriate)

- Participate in SERBs.
- Participate in E-SERB brief to COMOPTEVFOR.
- Review the test report.

8.7.6 01C Director

- Participate in SERBs.

- Participate in E-SERB briefs to COMOPTEVFOR.
- Review the test reports.
- Ensure technical analytical rigor supports system evaluations.

8.7.7 01D Director

- Participate in SERBs.
- Participate in E-SERB briefs to COMOPTEVFOR.
- Review the test reports.
- Ensure technical analytical rigor supports cyber survivability determinations.

CHAPTER 9 - OTHER TEST REPORTS

9.1 INTERIM REPORT

An Interim Report (IR) is a tailored product designed to provide the information needed to accomplish the purpose defined by the stakeholder(s). An IR, while not replacing the requirement for a complete and conclusive report, accelerates delivery of vital information to stakeholders to enable informed decisions. It provides "what we know now."

The OTD should develop a recommended scope for the IR and an outline of the proposed letter for division leadership review and approval. The report's purpose should be clearly stated in the IR's opening paragraph. With Division Director approval, proceed with IR production.

In its briefest form, which is normally used to support a decision on a short timeline, it takes the form of a Navy Correspondence Manual Standard Letter responding to a specific question, with supporting data and Blue and Gold Sheets as enclosures, as needed. This form of IR avoids "acquisition speak" because it will likely be written for a stakeholder outside the acquisition community. It may include factual test results "as they were observed" by the test team. For example, it may be arranged by mission area and/or kill chain with statements such as: "the destroyer was able to detect, track, engage, and destroy four boats using the new gun round" or "the man-pack radio was able to transmit clear voice but was not able to transmit using encryption." Suitability results may be included, for example: "during 200 hours of underway operations, the system had three hardware failures that delayed mission accomplishment." OPCONs may also be included. The PTIP may need to be tailored to meet timeline requirements while still ensuring supporting data, Blue and Gold Sheets, and test results receive appropriate vetting.

A more comprehensive IR can be used when a portion (one or more test periods) of a test phase has been completed and there will be a significant delay in finishing the whole test phase. This delay could be pre-planned, such as after the completion of live-firing events while the associated simulation suite is being validated for later OT use. The delay could be unplanned, such as when a critical test asset requires unscheduled maintenance that pushes out the remainder of testing. In either case, the IR can adopt the more comprehensive format of an OER generated through the process described in Appendix A. This adaptation would include production through the entire PTIP. Individual COIs may be resolved (if all critical data have been collected, scored, analyzed, and evaluated) or COIs may be characterized as "trending satisfactory" or "trending unsatisfactory" given sufficient data on hand to support these trending calls. COI trends may be split based on conditions. For example, a COI can trend unsatisfactory for one threat and the same COI can trend satisfactory for another threat. If insufficient data have been collected to make a COI "trending" characterization, then a statement to that effect should be made.

One or more determinations may be made or the System Under Test (SUT) can be characterized, for example, as "trending operationally effective" or "trending not operationally suitable." The use of "trending'' is what differentiates an IR from a final OER. A trend can be changed later. For example, a system can be reported as "trending not operationally effective" in an IR and be

determined to be "operationally effective" in the final report due to the analysis of more data or due to correction of deficiencies while still in test.

The executive summary accompanying the draft IR in the Electronic Document Router should address the risk to the Fleet if the IR is intended to inform an early fielding or deployment decision. The executive summary should also address any DOT&E concerns with the IR, or the decision it is intended to inform, if the IR concerns a project on the DOT&E oversight list.

Later, when OT is complete and all data have been collected, the Warfare Division will conduct the formal post-test process with the complete data set with the goal of producing a final report. The "trending" characterizations in the IR do not bind the test team in creating the final report, although it is helpful for clarity to the Commander and the customer to highlight any differences between the IR and that report as the document proceeds through the post-test process.

9.2 VCD

The test report for a stand-alone VCD test should indicate whether the in-scope deficiencies are corrected; not corrected, but substantially mitigated; or not corrected. For non-DOT&E oversight programs, when COI resolution is discussed in the test plan and if the VCD results enable a change to the resolution of COIs (beyond IOT&E), then those updated COI resolutions will be listed in the VCD report, thereby reducing the scope or eliminating the need for later phases of OT for the specific purpose of verifying the deficiency that has been corrected. For programs on DOT&E oversight, the only permitted change in COI resolution during a VCD phase of test is from SAT to UNSAT. The OPTEVFOR VCD report templates are located in the Y:\OT&E Production Library.

9.3 QRA

A QRA report takes the form of an operational risk assessment to address the purpose and answer the questions as outlined in the QRA tasking letter or a Rapid Fielding program's Master Test Strategy. A QRA will not resolve COIs; make operational effectiveness, operational suitability, and cyber survivability conclusions; and/or provide a limited Fleet introduction, Fleet introduction, or Fleet release recommendation. The OPTEVFOR QRA report template is located in the Y:\OT&E Production Library.

9.4 OBSERVED DT

A Memorandum of Agreement (MOA) between the PM and the OPTEVFOR Warfare Division A-Code guides reporting content when OPTEVFOR observes DT with an intent to provide a report. Depending on the purpose for observing DT described in the MOA, the OPTEVFOR report will either be a DT Assist LOO or an AOC letter:

- The purpose of a DT Assist is to assess the risk to resolving COIs as satisfactory at a future Initial Operational Test and Evaluation (IOT&E) or Follow-on OT&E (FOT&E). The DT Assist LOO takes the form of an OPTEVFOR O-6/GS-15 Warfare Division Director to an O-6/GS-15 Program Manager letter. The OPTEVFOR DT Assist LOO template is located in the Y:\OT&E Production Library.
- The purpose of an AOC is a special purpose capability assessment specified in the MOA per program office/resource sponsor/Fleet request (typically for introducing/releasing a system for Fleet/operational use) and the system has no future phase of Operational Test (OT) planned. The

AOC letter takes the form of an OPTEVFOR O-6/GS-15 Warfare Division Director to all associated/impacted stakeholders. The OPTEVFOR AOC letter template is located in the Y:\OT&E Production Library.

9.4.1 Post-Test Recommended Actions

Following the completion of testing, the OTD should conduct a DT data analysis review or CEWG-like working group to prepare a Post-Test Brief and a draft report letter for the A-Code. The draft report letter should discuss the observed operational tasks and the SUT's performance as it contributed to the related mission area(s). The purpose of this brief and draft letter is to frame the DT Assist LOO or AOC in the A-Code's mind, which should make it easier to route and will give the A-Code the opportunity to inject his or her perspective into the draft letter.

In preparation for this meeting, the OTD and test team should:

- Use the MOA as a starting point. Be prepared to show the A-Code the difference between what was planned and what was executed, and explain the reason(s) for differences.
- Review the results derived from any DT data for accuracy and to support risks or deficiencies.
- Using the draft outline that should have been prepared during planning and updated during execution, continue to update and prepare the draft letter outline based on what we know from DT observations to cover the PM's requested observations per the MOA. Update the letter to include the differences mentioned above such as asset or schedule limitations.
- The test team should prepare to discuss both what went right as well as what issues were observed and the operational impacts derived from these observations.
- Issues identified while observing DT are documented as Blue and Gold Sheets. LOOs document issues as risks, while AOCs document issues as deficiencies.

Conduct the Post-Test Brief with A-Code.

- Review the draft letter's content and flow to ensure expectations of the MOA are met. Have a copy of the MOA for ready reference.
- Ensure the letter presents the operational perspective. The letter should discuss the observed operational tasks and the SUT's performance as it contributed to the related mission area(s). Remember, the PM already has the test results.
- What makes a DT Assist LOO important is the assessment of risk to future OT and the operational impact of the DT test results. The AOC should address the specific needs of the requesting organization and other stakeholders per the MOA (e.g. capability being delivered or risks for fielding with no further testing planned). The goal for either letter is to clearly articulate observed SUT capability strengths and weaknesses from the Fleet user perspective.
- Ensure the letter represents a balanced discussion of both positive and negative test results. The letter should not focus solely on what the SUT did poorly. If the system will add new capability or capability enhancements to Fleet operations, the letter should so state. The letter should be clear about what capability is being delivered and what may be missing.
- Review draft Blue and Gold sheets for A-Code concurrence on content and level of risk or deficiency. Discuss if the A-Code desires to convene an optional B&G Peer Review.
- Review letter distribution per the MOA or per A-Code direction.
- Note and review directed actions by the A-Code.

- Make changes per division leadership and route for signature.

9.5 JCTD

A Joint Capabilities Technology Demonstration (JCTD), formerly an Advanced Concept of Technology Demonstration (ACTD), is a demonstration of the military utility of a significant new technology and an assessment to clearly establish operational utility and system integrity. On conclusion of a JCTD, an Operational Utility Assessment (OUA) or Military Utility Assessment (MUA) will be produced, signed by the Commander, and forwarded to the JCTD's Operational Manager (OM). The OUA report describes how a JCTD's products affect the resolution of an Operational Problem (OP) and fulfill operational Desired Capabilities (DC). It declares the level of operational utility according to the Concept of Operations (CONOPs) and Tactics, Techniques, and Procedures (TTP) and provides post-JCTD transition, CONOPs, TTP and Doctrine, Organization, Training, Material, Leadership, Personnel and Facilities recommendations.

The OUA/MUA is similar to an EOA/OA. A Demonstration Execution Document (DED), a document similar to a test plan, provides sufficient detail to measure MOPs, MOEs, and MOSs and analyze each COI. OPTEVFOR observations will state the planned and observed outcomes of the demonstration. COIs will be assessed using EOA/OA color codes, assessing military or operational utility. OPTEVFOR will not attempt to resolve JCTD COIs as SAT or UNSAT. Objective, analytical rigor applies to observations reported and the assessments made. Of special importance is the summary paragraph, which details the conditions and limitations under which the data were obtained.

APPENDIX A - PTIP Checklist

DESCRIPTION/PURPOSE: The Post-Test Iterative Process (PTIP) is a collaborative, disciplined evaluation of the System Under Test (SUT) within the context of each Critical Operational Issue (COI). The process begins with a series of scoring board(s) and COI Evaluation Working Groups (CEWG) preceding an optional Analysis Working Group (AWG), an optional Blue and Gold Sheet (B&G) Peer Review, the System Evaluation Review Board (SERB), and the Flag-level Executive SERB (E-SERB). Scoring boards validate operational test data was collected in accordance with the approved test plan and under operationally representative circumstances. The CEWG is a COI-by-COI Division/Squadron led working meeting reviewing the full breadth of COI data, risks or deficiencies, and results. The optional AWG conducts a data and analysis review of all effectiveness and suitability measures, including both quantitative and qualitative data with a focus upon critical measures and data supporting risk/deficiency identification. Normally, the AWG will not be required nor conducted, but may be in the team's best interest when clarity or accuracy are concerns. The cognizant Warfare Division finalizes the Data Analysis Summary following the last CEWG, or AWG if conducted. Similarly, the cognizant Warfare Division finalizes the Blue and Gold Sheets following the last CEWG, or an optional B&G Sheet Peer Review held at the discretion of the Warfare Division Director. The SERB and E-SERB then examine the Test Report letter containing the Executive Summary, the COI Results paragraphs, and any Operational Considerations (OPCON). The Warfare Division is required to brief test findings, conclusions, and recommendations to the decision authority for the decision the test report is intended to inform. In addition, the Warfare Division is responsible for briefing appropriate Fleet commands on COMOPTEVFOR's evaluation of the SUT's operational capabilities and limitations.

This checklist is intended to guide test teams through the PTIP for a single operational test phase of a typical program. As such, the PTIP Plan of Actions and Milestones (POA&M), initially developed during test planning and subsequently updated as required, should be based on this checklist. That said, the Warfare Division, with 01C assistance, can and should tailor the PTIP POA&M when circumstances warrant modification. For example, the PTIP POA&M can be developed or revised to accommodate an interim report during a lengthy test, or to account for an interim report to support an acquisition decision that has to be made before the final test report is completed.

NLT 30 Days Prior to Test

- ☐ Review data analysis plan and the Post-Test Plan of Action and Milestones (POA&M)
  For long duration tests (multiple test periods), the POA&M should include plans for scoring data and starting data analysis during test on a not-to-interfere basis with test execution
- ☐ Create shells of the Data Analysis Summary and the Test Report using the appropriate report templates on the Y drive (Y:\OT&E Production Library\Test Reports)
  - Determine proper enclaves (e.g., NIPRNET or SIPRNET) based on the applicable Security Classification Guide(s)

☐ Format the data section of the Data Analysis Summary (data results and data tables) consistent with measures, to include:
- Measure title and threshold
- Define the measure and type of measure (i.e., specified, derived, or OTA created)
- Introduce the planned analysis methodology
- Format supporting data tables and results tables

Conduct Test Operations

☐ Validate data was collected in accordance with the test plan at post-event debriefs
☐ Document test plan deviations and test limitations as they are encountered
☐ Commence drafting Blue and Gold Sheets as issues are noted
☐ Share raw factual data with the Program Office on a not-to-interfere basis with test execution
☐ Share all collected data for oversight programs with the DOT&E Action Officer (AO) on a not-to-interfere basis with test execution

Initial Post-Test Brief

OTD brief Warfare Division Director/Deputy Director/COTD/ACOTD, normally within 5 working days, but no later than 10 working days, after the last test event

☐ Include the following:
- Approved Test Plan
- Draft Blue and Gold Sheets or, at a minimum, a list of potential risks/deficiencies
- Draft results paragraphs (if available)
- POA&M for data reduction and analysis and follow-on briefs (Scoring Board, CEWGs, AWG, B&G Peer Review, SERB, and E-SERB)

☐ Discuss planned vs. completed test events/runs
- Focus on DOE run matrix and critical tasks/measures
- Highlight any shortfalls in data collection; discuss mitigation options
- Discuss limitations to test that were realized during execution, those that were anticipated and documented in the test plan as well as unanticipated limitations encountered during test execution

☐ Review and discuss SCG guidance for security classification of performance issues identified during test execution and post-test analysis/evaluation
☐ Provide Warfare Division Director/Deputy Director/COTD/ACOTD with a projected end-of-test date based on the estimated time required to receive and complete scoring of all test data
☐ Update the status/schedule of the decision(s) the final report is intended to inform
☐ Identify any deserved "BZ" recognition for other organizations and/or individuals who provided exceptional test support

Scoring the Data

The Scoring Board has three functions: inventory the data, qualify the data for OT, and score the run results. Before calling out-of-test, collected data must be reviewed to ensure all data required by the test plan was collected and is qualified for OT. The conduct of test must be reviewed, comparing actual test execution to the test plan's scope, Detailed Method of Test (DMOT), and the Data Collection Plan (DCP), with particular focus given to the validity of DOE runs. If sufficient data for COI assessment/resolution has not been collected, a decision will need to be made to either stay in test to collect additional data or document the missing data as a limitation to test. Scoring boards may be conducted in various ways at the discretion of the Division/Squadron, depending on specific test requirements, schedule, and objectives. One or more dedicated Scoring Boards (e.g., one for all Effectiveness COIs and one for all Suitability COIs) may be held, individual comprehensive COI-by-COI CEWGs including the scoring function may be held, or a hybrid approach including both dedicated Scoring Board(s) followed by comprehensive CEWGs may be conducted. The following scoring board checklist applies to effectiveness and suitability data; Warfare Divisions and 01D coordinate separately to score cyber survivability data.

Scoring Board

Participants:

- ☐ Section Head/OTC – required
- ☐ Squadron Branch/Department Head – required for VX SUT
- ☐ 01C Representative – required
- ☐ OTD – required
- ☐ Warfare Division/Squadron Analyst – required
- ☐ CNA Analyst (if assigned) - required
- ☐ LTE – required
- ☐ 01B CTF – required
- ☐ OTD's Support Contractor(s) – optional
- ☐ DOT&E AO (for oversight programs) – invited
- ☐ Program Office representative(s) – invited

Scoring Board Preparations

- ☐ Summarize test execution:
  - Planned vs. completed test events/runs
  - Incomplete testing or other deviations from the approved test plan
  - Rationale and mitigations for deviations.
- ☐ Compile data to be scored:
  - Planned run matrix
  - Number of runs completed
  - Qualification criteria for Operational Test (OT) data
  - Conditional criteria and tolerances
  - Scoring criteria (how data will be scored to determine hit/miss, Operational Mission Failure (OMF)/non-OMF, pass/fail, etc.)

☐ OTD provide read-ahead materials to all participants no later than 2 working days prior to the scheduled Scoring Board

Scoring Board Conduct

☐ OTD brief test execution, discussing planned vs. completed test events/runs, identifying any incomplete testing or other deviations from the approved test plan; deviations, to include rationale and mitigations, must be discussed
☐ Qualify the data as OT data by verifying:
- The SUT was operated and maintained by operationally representative operators and maintainers
- The SUT was employed in an operationally representative manner
- The SUT was tested against operationally realistic threats
- For IOT&E/FOT&E, the test article was production representative

☐ Verify data requirements for each measure were collected per the approved test plan
- For each data point, verify conditions were met for a valid opportunity/challenge
- Compare planned run matrix to number of runs completed

☐ Identify and note the rationale for test plan deviations
☐ Discuss impact of missing/incomplete data (coordinate with supporting statistician) including:
- Impact to planned confidence intervals, power calculations, and factor analysis
- Ability to conduct planned analyses
- Impact to planned COI assessment/resolution (including mitigations or workarounds, if they exist)
- Is there a limitation to test?
- If possible, augment with data from DT, alternate logs; otherwise stay in test (continue testing) and/or plan for follow-on testing, as appropriate

☐ For Response Variables (RV) and critical measures, compare planned to actual controlled conditions (including tolerances)
☐ Score the result of the observed task, if needed
☐ For reliability data, ensure operating time data supports determination of Total System Operating Time (TSOT) and ensure each failure is discussed and determination of OMF or abort is made, documenting the rationale for the determination
☐ For availability data, ensure neutral time, if applicable, is accounted for
☐ OTD provide scored RV data to the 01B Senior Analyst and Core Team Facilitator (CTF) for statistical analysis (see the Analysis Handbook) – this does not preclude using other analytical support if/when available (CNA, Corona, JHU APL, etc.)
☐ Commence CEWGs with scored data on a not-to-interfere basis with scoring other in-hand raw data

End of Test Message (normally issued NLT 30 days following completion of the last test event)

- ☐ All data in-hand and scoring board(s) completed (CEWGs do not have to be completed)
- ☐ Section Head, in coordination with the Squadron Branch/Department Head if VX SUT, makes recommendation to the Warfare Division Director for either the End of Test message to be sent, or for the SUT to remain in test to continue collecting data
- ☐ Warfare Division Director/Deputy Director brief 00TD if the end-of-test call cannot be made within 30 days of the last test event
- ☐ OTD prepare end-of-test message/email and provide to Warfare Division Director for release
- ☐ OTD ensure all raw data has been provided to the DOT&E AO
- ☐ OTD ensure all raw factual data (no raw subjective/survey/interview data) has been provided to the Program Office

CEWG Reviews

DESCRIPTION: A systematic data authentication process which ensures measure analysis is conducted within the context of the applicable COI. This review takes place within the Warfare Division/Squadron, supported by 01B and 01C for effectiveness and suitability COIs, and is structured to fit the needs of the Division/Squadron for the specific program. Several meetings/discussions may be necessary to cover all COIs. These meetings include data scoring (if not yet completed), data analysis, blue/gold sheets, results paragraph, and OPCON review and critique. The CEWGs ensure balanced COI resolutions and assessments are made within the context of the test design, using scored data of SUT performance, including observed capabilities and shortfalls. The number of meetings and their periodicity should be determined based on the length of the test phase and the amount of data being collected. The OTD, in collaboration with 01C and the Warfare Division Deputy Director/ACOTD, Section Head/Branch Head, and LTE should use the POA&M initially developed during the test planning process to guide the post-test iterative process. As soon as all data for any given COI has been scored, the test team may conduct a CEWG for the COI. When assigned, the LTE should facilitate and guide the conduct of the CEWG. When an LTE is not assigned, the OTD should facilitate the CEWG. The following CEWG checklist applies to effectiveness and suitability COIs; Warfare Divisions and 01D coordinate separately to accomplish the cyber survivability CEWG.

Participants:

- ☐ OTD – required
- ☐ Section Head/OTC – required
- ☐ OTD's Support Contractors – as available
- ☐ Warfare Division/Squadron Analyst – required
- ☐ LTE – required
- ☐ 01C Representative – required
- ☐ 01B CTF – invited
- ☐ CNA Analyst (if assigned) – required

CEWG Preparations

- ☐ Review Scoring Board results
- ☐ Discuss and document any incomplete testing or other deviations from the approved test plan. Deviations are categorized in one of the following categories:
  - Minor Deviations – Minimal impact to COI resolution. Should be documented as a Minor limitation to test.
  - Major Deviations – Affects and possibly precludes COI resolution. Should be documented as either a Major limitation to test (will affect COI resolution) or a Severe limitation to test (precludes COI resolution).
  - Other Deviations – No impact to the ability to resolve COIs. Should be documented as a deviation.
- ☐ For each COI, list the associated critical tasks (see the Traceability matrix in Appendix B of the test plan and the Data Analysis Plan in Enclosure (2) of the test plan)
- ☐ For each critical task, list the KPPs and critical measures per the approved test plan
- ☐ Review data requirements for each measure
- ☐ Document limitations to test
  - If data requirements are missing or incomplete for a critical task, or a critical task was not accomplished, this is likely a major limitation to test
  - If data requirements are missing or incomplete for a non-critical task, or a non-critical task was not accomplished, this is likely a minor limitation to test
  - Refer to the OT&E Manual for policy guidance on defining limitations
- ☐ Analyze all critical measures per the Data Analysis Plan in Enclosure (2) of the test plan
  - For RV analysis, see the Analysis Handbook
  - Ensure RV Analysis Outbrief is complete
- ☐ Analyze all non-critical measures listed in the Measures-to-Data Requirements table in Appendix B of the test plan
- ☐ Calculate confidence intervals where appropriate (see the Analysis Handbook)

NOTE: Sample size, power, and confidence calculations (all part of DOE before the test) depend on the exact functions to be used in the analysis (after the test); it is critical that post-test analysis is consistent with planned data analysis methodology.

- ☐ In collaboration with 01C, determine the scope of non-critical measures to be included in the data section of the Data Analysis Summary; otherwise, the expectation is that all measures will be included
- ☐ Prepare the COI's data section of the Data Analysis Summary, filling in/building on the shell of the Data Analysis Summary created prior to test execution
  - Clearly define the measure and explain the analysis methodology as executed
    - Conditions under which data was collected
    - State the formula or analysis technique
    - Discuss any tools or software used for calculations
    - Any external agencies involved in any portion of the analysis and their analysis responsibilities
    - For measure with criterion, state whether or not the measure was met
  - Present data in a clearly annotated format (table, figure, plot, chart, etc.)

- List the overall measure result(s) followed by the supporting data used to calculate the measure(s)
- Display results not supporting task accomplishment in red font
- Incorporate RV analysis inputs
- Present results in the units of measure specified in the requirements document (specified) or approved test plan (derived)
- Include the threshold/criterion in the measure result table if defined; otherwise, specify the measure's criterion is "No Threshold"
- Ensure calculation results are rounded correctly and presented with the correct number of significant digits per the Data Analysis Summary template (one digit beyond the threshold value)

☐ Use appropriate report template on the Y drive for Blue and Gold Sheet preparation (Y:\OT&E Production Library\Test Reports)

- Determine Blue or Gold Sheet using the SUT/SoS definitions in the test plan
- Write a concise description of the problem in paragraph 1
- Open paragraph 2 with a description of the test conditions – what/where/when – focusing on the task
- Paragraph 2: Explain the impact on the operator, mission, or task. Include any work-arounds.
- Use the mission relation boilerplate: "During [mission, task, or subtask as appropriate], the [problem] will cause [impact to system, operator, or task completion] resulting in [impact to mission]."
- Ensure Conclusions and Recommendations are consistent across the SUT
- Request Resource Sponsor concurrence/non-concurrence on the applicability of any derived requirements the Blue Sheets are based upon (for VX SUT, the OTC communicates with the Resource Sponsor)

☐ Compile the draft PMT View (see Appendix C)

☐ Write COI results paragraphs per the COI Resolution Methodology described in enclosure (2) of the Test Plan and in accordance with Test Report template guidance, ensuring both positives and negatives are discussed, with a focus on SUT capabilities, functionalities and accomplishment of tasks

- Clearly state the capability of the SUT to accomplish or support 1st level critical tasks (as depicted in the IEF and Concept of Test Brief)
- Comply with the Suitability Handbook for Suitability COIs
- Account for each applicable Blue and Gold Sheet

☐ Write OPCONs (or Design Considerations for an EOA), as necessary

☐ Provide the following to all participants (or ensure all participants have access to) no later than 2 working days prior to the scheduled CEWG:

- Approved Test Plan
- Rough Draft data section of the Data Analysis Summary for the selected COI
- Rough Draft Blue and Gold Sheets for the selected COI
- Rough Draft COI Results paragraph
- Rough Draft PMT View
- Rough Draft OPCONs for the selected COI

Conduct a CEWG for each COI

- ☐ Score data per Scoring Board procedures (above), or review results from the Scoring Board
- ☐ Discuss any incomplete testing or test deviations
- ☐ Discuss statistical and operational impact of missing data
  - Changes to confidence and power calculations
  - Ability to execute the Data Analysis Plan
  - COI resolution (including mitigations or workarounds, if they exist)
  - If possible, augment with data from DT, alternate logs; otherwise stay in test (continue testing) and/or plan follow-on testing, as appropriate
- ☐ Identify and note the rationale for excluding any data from analysis; excluded data must be discussed during the CEWG and briefed to the Division Deputy Director (see post-CEWG actions below)
- ☐ Review limitations to test
- ☐ Review the analysis of all critical measures
  - For RVs, verify Analysis Handbook compliance
  - Review RV Analysis Outbrief
  - Review all calculations, including confidence interval calculations
- ☐ Review all non-critical measures
- ☐ Review the draft data section of the Data Analysis Summary
- ☐ Review draft Blue and Gold Sheets
- ☐ Review the draft PMT View
- ☐ Review the COI results paragraph
- ☐ Identify significant points to be considered for inclusion in the OE, OS, or CS paragraphs in the Test Report letter
- ☐ Review OPCONs

Post-CEWG Actions

- ☐ Following the incorporation/adjudication of all comments, OTD brief the Division Deputy Director or Squadron COTD on all CEWG products
- ☐ Incorporate all comments from the Division Deputy Director or Squadron COTD
- ☐ OTD provide RV Analysis Outbrief and completed draft data section of the Data Analysis Summary to the 01B Senior Analyst and 01B CTF per the Analysis Handbook
- ☐ Repeat for ALL COIs
- ☐ When the last COI's CEWG is completed:
  - Finalize the data section of the Data Analysis Summary per below
  - Finalize the Blue and Gold Sheets per below
  - Compile all sections of the Test Report per the appropriate template on the Y drive (Y:\OT&E Production Library\Test Reports) into a complete draft final Test Report (or documents, if a portion of the report will be distributed separately for classification reasons) and provide to the Division Deputy Director and Squadron COTD for review
  - Brief the Division Deputy Director and Squadron COTD on significant lessons learned discovered through the test planning, execution, and the PTIP process to date. Highlight key cross-divisional lessons and discuss if this is a candidate for a command-wide

Internal Post-Test Brief. Identify key stakeholders who may require a briefing on the test results either before or after the report is signed.

- Assemble one-page Fact Sheet providing summary of capabilities tested, deficiencies, and high level results. The Fact Sheet should be a short, concise, easy-to-read summary of test results and is intended to accompany the SERB/E-SERB and Fleet stakeholder briefs.
- Obtain Division Deputy Director and Squadron COTD concurrence prior to sending out the SERB read-ahead

Finalize the Data Analysis Summary

- ☐ Use the appropriate Data Analysis Summary template on the Y drive to consolidate CEWG data analysis

  (Y:\OT&E Production Library\Test Reports)
- ☐ Measures supporting multiple COIs are only reported once (first use); refer the reader to the first measure entry as needed in subsequent COI sub-sections of the data sections
- ☐ OTD complete the remaining sections of the draft Data Analysis Summary (scope of test, limitations, deviations, and resources)
- ☐ OTD provide the draft Data Analysis Summary to the Section Head, LTE, 01C Representative, and 01D Representative
- ☐ Section Head, LTE, 01C Representative, and 01D Representative review the draft Data Analysis Summary and provide comments to OTD
- ☐ OTD adjudicate Section Head, LTE, 01C Representative, and 01D Representative comments
- ☐ Proceed with the AWG checklist below if the Warfare Division and/or 01C determine an AWG is warranted; otherwise:
  - OTD obtain Warfare Division leadership concurrence the draft Data Analysis Summary is ready for 00TD review and approval
  - OTD schedule a 30-minute meeting with 00TD to brief him on the Data Analysis Summary. As a minimum, the LTE (if assigned) and the Lead Analyst should accompany the OTD to this meeting.
  - OTD provide a copy of the draft Data Analysis Summary to the 01B Director and concurrently route the completed draft to 00TD for approval and cover memorandum signature. Also forward a copy of the RV Analysis Outbrief to both 00TD and the 01B Director if RVs are included in the Data Analysis Summary.
  - Warfare Division obtain date and serialization per SOP 14-1 and provide the entire document to the Editors or Vault for conversion to pdf file. Editors or Vault prepare the pdf file and forward to Code 01A and the Warfare Division.
  - 01A archive the Data Analysis Summary in the appropriate Y-drive signed test report folder

AWG (Optional)

Participants:

- ☐ 00TD – as available

- ☐ Warfare Division Director or designated representative – required
- ☐ VX CO or designated representative – required for VX SUT
- ☐ 01C Director and/or designated representative – required
- ☐ Warfare Division Deputy Director/Squadron COTD and/or designated representative – required
- ☐ Section Head/OTC – required
- ☐ OTD – required
- ☐ Warfare Division/Squadron Analyst – required
- ☐ CNA Analyst (if assigned) – required
- ☐ LTE – required
- ☐ 01C Representative – required
- ☐ 01B CTF – required
- ☐ OTD's Support Contractor(s) – optional

AWG Preparation

- ☐ If directed by Division Director/Squadron CO or 01C Director, OTD schedules AWG
- ☐ OTD provide read-ahead materials to all participants electronically NLT 2 full working days prior to the scheduled AWG:
  - Approved Test Plan
  - Adjudicated draft Data Analysis Summary
  - Other supporting data or documents (as needed)

AWG Conduct

- ☐ OTD step through the data section of the Data Analysis Summary for the Division/Squadron Leadership and 01C. Trace critical measures and measures associated with deficiencies/risks to the Test Plan and ensure each is adequately discussed, properly formatted, and correctly calculated.

Post-AWG Actions

- ☐ OTD adjudicate comments received at the AWG
- ☐ OTD obtain Warfare Division leadership concurrence the draft Data Analysis Summary is ready for 00TD review and approval
- ☐ OTD schedule a 30-minute meeting with 00TD to brief him on the Data Analysis Summary. As a minimum, the LTE (if assigned) and the Lead Analyst should accompany the OTD to this meeting.
- ☐ OTD provide a copy of the draft Data Analysis Summary to the 01B Director and concurrently route the completed draft to 00TD for approval and cover memorandum signature. Also forward a copy of the RV Analysis Outbrief to both 00TD and the 01B Director if RVs are included in the Data Analysis Summary.
- ☐ Warfare Division obtain date and serialization per SOP 14-1 and provide the entire document to the Editors or Vault for conversion to pdf file. Editors or Vault prepare the pdf file and forward to Code 01A and the Warfare Division.
- ☐ 01A archive the Data Analysis Summary in the appropriate Y-drive signed test report folder

Finalize the Blue & Gold Sheets

- ☐ OTD obtain Resource Sponsor concurrence/non-concurrence for all derived requirements related to a Blue Sheet risk/deficiency (may have been documented in IEF development) (for VX SUT, the OTC communicates with the Resource Sponsor)
- ☐ OTD obtain Division Deputy Director/Squadron COTD concurrence on the draft Blue and Gold Sheets; ensure satisfactory adjudication from the post-CEWG, Division Deputy Director/Squadron COTD review/briefing
- ☐ Warfare Division share the draft Blue and Gold Sheets with the Program Manager (PM) to obtain PM comments on the sheets prior to approval
- ☐ Warfare Division Directors shall brief the Commander on all Severe Blue and Gold Sheets prior to approval.
- ☐ Proceed with the B&G Peer Review checklist below if the Warfare Division determines a B&G Peer Review is warranted; otherwise:
  - OTD route the Blue and Gold Sheets with cover letter for Warfare Division Director approval and signature
  - Warfare Division obtain date and serialization per SOP 14-1 and provide the entire document to the Editors or Vault for conversion to pdf file. Editors or Vault prepare the pdf file and forward to Code 01A and the Warfare Division.
  - Warfare Division distribute the Risk/Deficiency Letter
  - 01A archive the Risk/Deficiency Letter in the appropriate Y-drive signed test report folder

B&G Peer Review (Optional – following Division Deputy Director Review of the CEWG B&G Sheets)

Participants:

- ☐ 00TD – invited for controversial findings
- ☐ Warfare Division Director or designated representative – required
- ☐ VX CO or designated representative – required if VX SUT
- ☐ Warfare Division Deputy Director/Squadron COTD – optional
- ☐ 01C Director or designated representative – required
- ☐ 01D Director or designated representative – required if cyber survivability is in-scope
- ☐ Section Head/OTC – optional
- ☐ OTD – required
- ☐ Warfare Division/Squadron Analyst – optional
- ☐ LTE – optional
- ☐ 01C Representative – optional
- ☐ 01B CTF – optional
- ☐ OTD's Support Contractor(s) – optional

- ☐ OTD provide read-ahead materials to all participants electronically NLT 2 full working days prior to the scheduled B&G Peer Review:
  - Approved Test Plan
  - Data Analysis Summary

- Post-CEWG, Division Deputy Director/Squadron COTD reviewed draft Blue and Gold Sheets
- Other supporting data or documents (as needed)

☐ OTD step through the draft Blue and Gold Sheets for the Division/Squadron Leadership and 01C
☐ OTD adjudicate comments received at the B&G Peer Review
☐ OTD obtain Section Head, LTE, 01C Representative, 01D Representative, and Division Deputy Director/Squadron COTD concurrence that B&G Sheet Peer Review comments are satisfactorily adjudicated
☐ OTD route the Blue and Gold Sheets with cover letter for Warfare Division Director approval and signature.
☐ Warfare Division obtain date and serialization per SOP 14-1 and provide the entire document to the Editors or Vault for conversion to pdf file. Editors or Vault prepare the pdf file and forward to Code 01A and the Warfare Division.
☐ Warfare Division distribute the Risk/Deficiency Letter
☐ 01A archive the Risk/Deficiency Letter in the appropriate Y-drive signed test report folder

Command SERB

Participants:
☐ 00TD – invited for controversial findings
☐ Warfare Division Director and/or designated representative – required
☐ VX CO or designated representative – required if VX SUT
☐ Warfare Division Deputy Director/Squadron COTD – required
☐ 01C Director or designated representative – required
☐ 01D Director or designated representative – required if cyber survivability is in-scope
☐ Section Head/OTC – required
☐ OTD – required
☐ Warfare Division/Squadron Analyst – required
☐ LTE – required
☐ 01C Representative – required
☐ 01B CTF – invited
☐ OTD's Support Contractor(s) – optional

Command SERB Preparation
☐ Division Deputy Director and Squadron COTD complete post-final CEWG review and concur with the draft Test Report
☐ OTD schedule Command SERB once Warfare Division and Squadron comments have been adjudicated
☐ OTD provide read-ahead materials to all participants electronically NLT 2 full working days prior to the scheduled SERB:
- Approved Test Plan
- Approved Data Analysis Summary
- Approved Risk/Deficiency Letter

- SERB Brief with PMT View (use SERB Brief Template (see Y:\OT&E Production Library\Test Reports\SERB and ESERB))
- Draft Test Report
- Other supporting data or documents (as needed)

☐ OTD build 5 briefing notebooks using 3 tab dividers for SERB participants to include the following:
- Tab 1: SERB Brief
- Tab 2: Draft Test Report
- Tab 3: Approved Risk/Deficiency Letter

Command SERB Conduct

☐ OTD designate test team member as "recorder" to document SERB actions/directions
☐ OTD present the SERB brief
☐ OTD lead the review of the Test Report's enclosure, COI-by-COI
- Review the PMT View and major quantitative and qualitative test results by describing how measured SUT performance contributes to or detracts from the accomplishment of the task(s), then
- Review the COI results paragraph, then
- Review any OPCONs associated with the COI, then
- Record SERB concurrence / non-concurrence
- Repeat for each remaining COI

☐ OTD present the proposed findings, conclusions, and recommendation(s) in the Test Report
☐ OTD brief Fleet stakeholder engagement and briefing plans that may be warranted before and/or after report is released. Discuss who these stakeholders are, what needs to be briefed, and when briefs should occur.
☐ Document final Command SERB comments/recommendations
☐ OTD read back action items to SERB participants for concurrence
☐ OTD collect notebooks and retain notes and markups from each reviewer for later E-SERB use

Post-Command SERB Actions

☐ OTD ensure all substantive comments are adjudicated NLT 2 working days prior to the E-SERB for changes to the:
- SERB/E-SERB brief
- Test Report

E-SERB

Participants:
Same as for the Command SERB, with the addition of:
☐ COMOPTEVFOR – required
☐ Deputy, COMOPTEVFOR – required
☐ Technical Director – required
☐ 01A Director – required

E-SERB Preparations

- ☐ Warfare Division share draft COI results paragraphs with the PM to obtain comments prior to the E-SERB
- ☐ Warfare Division prepare a smooth final draft of the Test Report for Commander's signature
- ☐ Warfare Division Deputy Director prepares the Test Report email
- ☐ OTD prepare briefing notebooks (following SERB preparation procedures above) and provide read-ahead NLT 2 working days prior to the scheduled meeting

E-SERB Conduct

- ☐ OTD designate a test team member as "recorder" to document E-SERB actions/directions
- ☐ OTD present E-SERB brief
- ☐ OTD lead the review of the Test Report's enclosure, COI-by-COI
  - Review the PMT View and major quantitative and qualitative test results by describing how measured SUT performance contributes to or detracts from the accomplishment of the task(s), then
  - Review the COI results paragraph
  - Review any OPCONs associated with the COI
  - Repeat for each remaining COI
- ☐ OTD present the proposed findings, conclusions and recommendation(s) in the Test Report
- ☐ OTD brief Fleet engagement and briefing plans. Identify specific stakeholders, brief content, and timeline. Include a summary of engagements already conducted
- ☐ Document E-SERB decisions
- ☐ OTD read back action items to E-SERB participants for concurrence
- ☐ Warfare Division Director recommend the Commander sign the Test Report or propose a way-ahead to obtain the Commander's signature on the Test Report

Post-E-SERB Actions

- ☐ OTD draft formal E-SERB action item/minutes
- ☐ Warfare Division Deputy Director forward the Test Report email to the Flag Writer
- ☐ When the Commander signs the Test Report:
  - Flag Writer obtain date and serialization per SOP 17-1 and provide the document to the Editors or Vault for conversion to pdf file. Editors or Vault prepare the pdf file and forward it to Code 01A and the Warfare Division.
- ☐ Warfare Division distribute the Test Report after the Commander sends the Test Report email
- ☐ 01A archive the Test Report in the appropriate Y-drive signed test report folder
- ☐ Warfare Division brief test findings, conclusions, and recommendations to the decision authority for the decision the test report is intended to inform
- ☐ Warfare Division brief appropriate Fleet commands on COMOPTEVFOR's evaluation the SUT's operational capabilities and limitations

APPENDIX B - Risk and Deficiency Sheets

1. Initial Effectiveness/Suitability Performance Issue Identification

Each sheet has a unique issue number to track the issue from identification to correction. The numbering scheme uses the program TEIN with a three-digit modifier (i.e., 1420-001, 1420-002, or 3000-371-001). In addition, as the same issue is updated, the number includes a modifier for the revision (i.e., 1420-001, 1420-001 Rev 1; 1420-001 Rev 2).

2. Categorization

SUT and SoS performance issues are categorized as "draft", "risk", "deficiency", or "closed". The following are the categorization descriptions and the reporting product each supports.

a. Draft

A "Draft" categorization is used for performance issue discovery. Draft is used during data collection to document the current information, when there is not enough data to know if there is an effectiveness/suitability performance issue. It is also used to update a risk/deficiency with new data. Draft categorizations do not support any product.

b. Risk

A "Risk" (Severe, Major 1/2/3, Minor) categorization is used for performance issues identified during DT Assists, EOAs, OAs, and QRAs. CEWGs and B/G Peer Reviews are not required for DT Assists.

c. Deficiency

A "Deficiency" (Severe, Major 1/2/3, Minor) categorization is used during IOT&E, FOT&E, VCDs, and AOCs.

d. Closed

A "Closed" categorization is used to document a deficiency that is no longer an issue. Closed is used to record data and analysis of SUT or SoS performance issues that have been corrected or are no longer relevant to the SUT evaluation. Once closed, the issue number is retired.

(1) When there are consecutive DT Assists, EOA, or OA test periods scheduled, and the Program Office takes action to correct performance issues associated with a Risk, then the Risk should be verified as corrected, and may be reported as such using the methodology and format provided for VCD Corrected and Closed (or Not Corrected) Blue and Gold sheets. Coordinate with 01C for program-specific questions.

(2) For performance issues identified during test and subsequently corrected by the Program Office while still in test, whether or not to publish a Blue or Gold sheet is dependent on the status of testing. If the test completion message has not been transmitted and the issue has been verified corrected by OT, the issue need not be documented in the test report. However, if the test completion message has been transmitted, and the issue has

been verified corrected by OT, the issue should be documented in the test report as a Corrected and Closed Blue or Gold sheet.

(3) Blue and Gold sheets may be administratively closed when the original performance issue is no longer relevant to the SUT evaluation. If a "closed" Blue or Gold sheet is desired for an issue being administratively closed, the term "corrected" should not be used. Examples of reasons for administratively closing Blue or Gold sheets include:

- CONOPS change
- Threat changes
- Requirements change
- Relevant component no longer integrated with the SUT
- Issue integrated into a new Blue sheet.

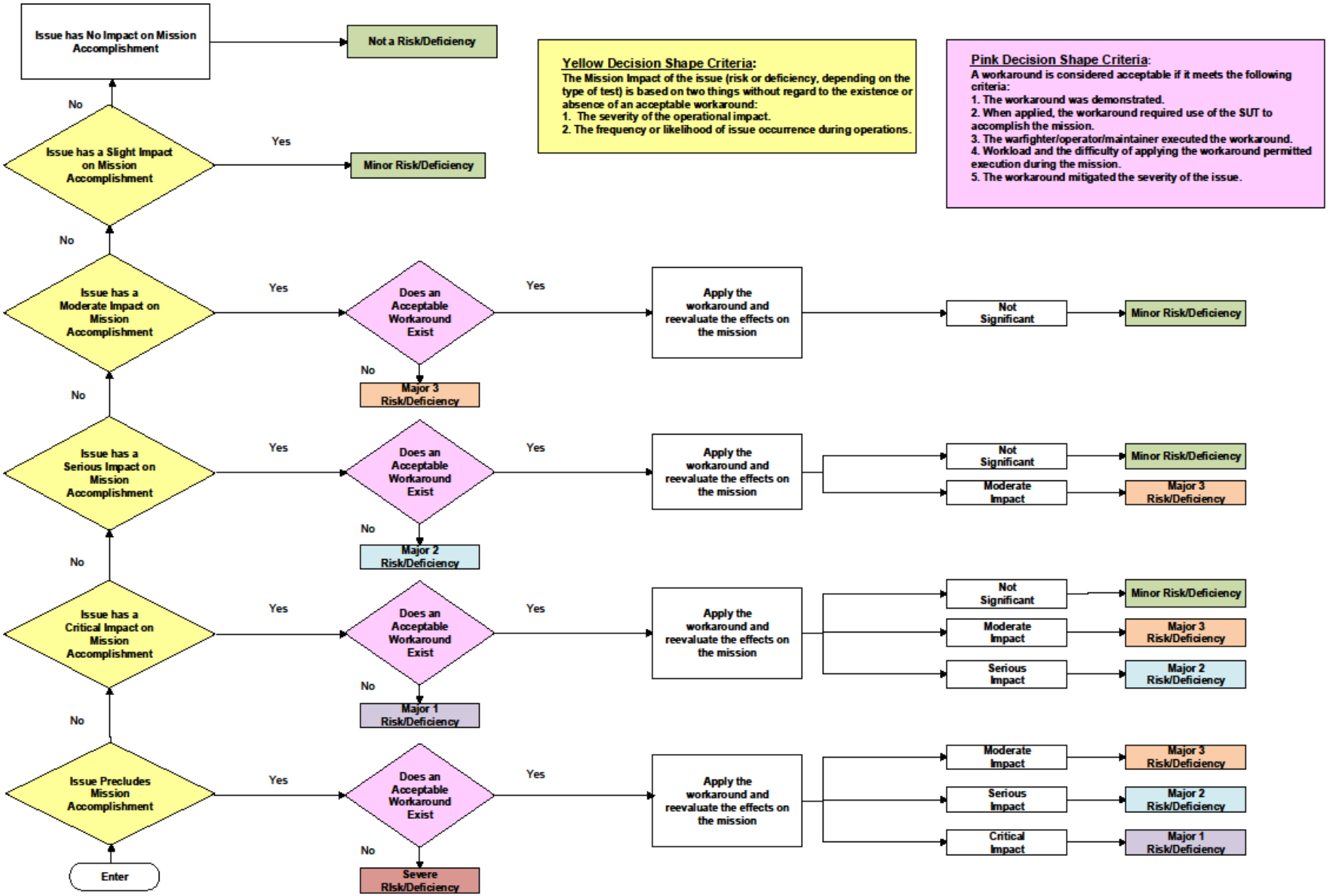
3. Risk and Deficiency Severity Levels

A deficiency is defined as lacking in some necessary quality, capability, or element or not up to a normal standard or complement. Operational capability is defined as a capability or means that is directly traceable to an approved requirement (i.e., CDD, CPD, etc.). Mission-essential capability is defined as a capability that is inherently necessary to complete an assigned mission (e.g., a targeting mechanism is required to properly aim a weapon system, but the targeting mechanism/system may not be part of the weapon SUT). Table B-1 provides the baseline risk and deficiency definitions that shall be used throughout the evaluative process to make a final conclusion as to the risk or deficiency level. See figure B-1 for the baseline risk and deficiency flow diagram.

Table B-1. Baseline Risk and Deficiency Definitions
UNCLASSIFIED

Severe	Precludes mission accomplishment
Major 1	Critical impact on mission accomplishment
Major 2	Serious impact on mission accomplishment
Major 3	Moderate impact on mission accomplishment
Minor	No significant impact on mission accomplishment

Figure B-1. Risk/Deficiency Flow Diagram
UNCLASSIFIED



THIS PAGE INTENTIONALLY LEFT BLANK

4. Blue and Gold Sheet Reviews

Blue and Gold sheets are vetted during the CEWGs and the optional B&G Peer Review. Post-CEWG, draft Blue and Gold sheets approved by the Warfare Division Deputy Director will be shared, by e-mail, with the O-6 level PM by the division or squadron O-6 leadership (A-Code or squadron CO/COTD). The e-mail should include a request for comments on the draft Blue and Gold sheets from the PM.

5. Continuous Risk/Deficiency Revisions

As the SUT progresses through its development and additional OT-qualified data are obtained, the original Blue/Gold sheet should be updated to represent the current status of the risk/deficiency. Updates to a risk/deficiency are annotated by the use of a DRAFT watermark across the sheet, as well as the addition of a "Rev" modifier to the issue number (i.e., 1420-001 Rev 1). Once the revision has been approved by the Warfare Division Director, the draft labeling will be deleted and the date is updated to the date of approval. In summary, when documenting the update using the Blue or Gold sheet, include the following:

- New data,
- Add the DRAFT watermark and update the revision number with the "Rev #" modifier and the date,

For risks or deficiencies, conduct the CEWG, and optional B&G Peer Review.

6. Life Cycle of an Issue Number

A unique number (usually a three-digit number) is assigned to each potential performance issue as it is identified. Once a number is assigned to a performance issue and initially documented in a draft Risk/Deficiency Sheet, that number is tied to that issue only. The number may not be re-assigned or used for any other issue within the same TEIN (regardless of SUT or SoS) even if the draft Risk/Deficiency Sheet is never issued as an approved Risk/Deficiency Sheet. The issue category (Risk or Deficiency) links the performance issue to a type of test, and the Revision number uniquely identifies the issue for subsequent versions within the same categorization. Once an issue becomes a Deficiency, it remains a Deficiency, regardless of the type of test, until it is Closed. Purely for illustration purposes, the following assumes the fourth performance issue identified during a QRA for a program (TEIN 1420) is observed throughout the life-cycle of the program until eventually being corrected during a VCD event for Increment 2.

Increment 1
QRA #1, Risk No. 1420-004
DT Assist #1, Risk No. 1420-004 Rev 1
OT-B1, OA, Risk No. 1420-004 Rev 2
OT-C1, IOT&E, Deficiency No. 1420-004
OT-C1, VCD, Deficiency No. 1420-004 Rev 1

Increment 2
DT Assist #2, Deficiency No. 1420-004 Rev 2
OT-D1, OA, Deficiency No. 1420-004 Rev 3
OT-D2, FOT&E, Deficiency No. 1420-004 Rev 4
OT-D2, VCD, Deficiency No. 1420-004 Rev 4 (Closed)

7. Blue/Gold Sheet Writing

OPTEVFOR communicates the results of OT to stakeholders and customers formally in test reports. The nectar within the various test reports are the risks and deficiencies identified by the testers. All risks and deficiencies are described within Blue or Gold sheets. There are two basic types of Blue or Gold sheets; risk sheets associated with a DT Assist LOO, EOA/OA, or QRA, and deficiency sheets associated with IOT&E/FOT&E reports, VCD reports and AOC letters. As noted above, Blue sheets are associated with and apply to the SUT, while Gold sheets apply to or are associated with the greater SoS. Blue/Gold sheets are formatted descriptions of the performance issue and are intended to stand alone without reference to other documents. The structure of a Blue or Gold sheet is loosely based on the U. S. Naval Test Pilot School 6-part paragraph. It is imperative to convey clearly the intent and logic inherent in the thought process when communicating test results. A properly constructed Blue/Gold sheet will have the following characteristics:

- The result is **goal-directed**. The writer should identify the purpose of the result and its importance to reader.
- The result is **clear, concise, and organized**. The writer should "cut to the chase" with logically formulated, direct, simple language.
- The result is **easily understood by non-experts in the subject matter**. The writer should not assume every reader has his or her skills and experience.
- The result is **defendable**. This characteristic refers to, relies on, and reinforces OPTEVFOR's credibility.

8. The Five Blue/Gold Sheet Paragraphs

a. Paragraph 1, Problem or Issue Description

State the issue. Describe the problem/issue concisely in one sentence using past tense. Only address one issue/problem per Blue/Gold Sheet. Do not address/include the operational impact in paragraph 1; the operational impact is discussed in paragraph 2 of the Blue/Gold Sheet.

b. Paragraph 2, Part 1, Establish the Test Conditions (past tense)

Establish the test conditions. The information here is the start of paragraph 2 of the Blue/Gold sheet. Establishing the test conditions should be done in one or two sentences and should focus on what testing was being conducted when the problem was discovered. Specifically, the writer should describe what vignette, mission, or task/subtask was being performed. The OTD should use the test plan to help frame what was being accomplished. The specific test conditions, which affected the result, apply. These conditions bound the problem and support repeatability. Again, state what was being evaluated and how the evaluation was performed, as well as any pertinent conditions for the test. The "what" is very closely related to the problem or deficiency. The problem or deficiency is called out specifically in paragraph 1 of the Blue or Gold sheet.

c. Paragraph 2, Part 2, Present Data (past tense)

Present data. The data and results presented here follow the test condition sentence(s) at the beginning of paragraph 2. Only present data related to the issue. The Blue/Gold sheet should

stand alone and reference to the data analysis summary memorandum should not be made. Pull the specific data/results needed to make the case to the reader from the data analysis summary. Photos, screen shots, figures, and tables with detailed annotation are encouraged. When annotating figures, photos, or screenshots, use the same wording used in paragraph 2 text to allow the reader to easily follow your discussion. Again, state what data were collected. Only name or call out data related to your analysis and conclusions. Data can be quantitative or qualitative. Include a description of any work-around (using the SUT) if used by the operators to make the system overcome an issue or deficiency. Present the data from a third-person, objective point of view (e.g., do not use wording like "the OTD observed"). Focus should be on results and not test method.

d. Paragraph 2, Part 3, Analyze/Evaluate the Data (past tense)

Analyze/evaluate the data. The information/analysis of the results presented here forms the basis of the sheet and should logically flow into the mission relation. Use the data and results presented to explain the impact upon the operator, mission, or task/subtask. What does the data indicate? The evaluation could include a comparison to legacy systems. Include an evaluation of how any work-around used by the operator mitigates the issue or deficiency and the resulting impact on mission accomplishment both before and after application of the work-around.

e. Paragraph 3, Mission Relation (future tense)

Mission relation. The "so what." The mission relation appears in paragraph 3 and must build upon the data, results, analysis, and evaluation presented in paragraph 2 of the Blue/Gold sheet. This part describes the impact to the Fleet, operator, or mission of the problem described in paragraph 1 of the sheet. Ensure the problem is included in the mission relation and is stated in the same context/meaning as in paragraph 1. Present the mission relation paragraph in future tense. The mission relation should normally be no more than one sentence, and rarely over two sentences. A boilerplate for mission relation is provided in the report templates and its use is strongly encouraged.

e. Paragraph 4, Conclusion (present tense)

Conclusion. The conclusion is paragraph 4 and is one simple sentence where the subject is the problem and the predicate is the level of the risk or deficiency (Severe, Major 1, 2 or 3, or Minor). When reviewing the conclusion statement, ensure the problem is restated in the same context as was stated in paragraph 1 of the Blue/Gold sheet. For conclusions in revised sheets where the level of deficiency has been mitigated or elevated, the new (updated) deficiency level is stated as the conclusion. For issues verified corrected, the conclusion states the issue was corrected and the risk/deficiency is closed. For issues administratively being closed, the conclusion states the issue is administratively closed.

f. Paragraph 5, Recommendation (future tense)

Recommendation. The recommendation is paragraph 5 and normally a standardized sentence, selected from the options presented in the template. The recommendation provides general timing for the desired correction of the risk or deficiency. Refer to the templates and use the recommended choices unless a case-specific recommendation is needed. For example, if a submarine or submarine system had a problem that only mattered when operating under the ice

cap, then the recommendation may be "Correct prior to operating the submarine/system under ice."

9. Constructing the Blue/Gold Sheet

A guide to aid in constructing Blue/Gold sheets, in the form of questions to be answered, is presented below:

- What is your issue (problem)?
  - Clearly articulating the problem is critical to developing a Blue/Gold sheet. The problem appears in paragraph 1 of a Blue/Gold sheet, as well as in the mission relation and conclusion paragraphs. Take special care to be consistent in these paragraphs when describing the problem.
- How bad is it (Severe, Major 1, 2, 3, or Minor)?
  - This should be your initial judgment and may change as the test team critiques the issue or more is learned during the analysis process.
- What is the impact on the mission, if not fixed?
  - The mission relation appears in paragraph 3 and must build upon the data, results, analysis, and evaluation presented in paragraph 2 of the Blue/Gold sheet. Present this paragraph in future tense. The mission relation should normally be one or no more than two sentences.
- Was there a work-around using the SUT and what was the mission impact with the work-around?
- What test conditions were relevant to collecting the data?
  - The information here is the first part of paragraph 2 (Test Conditions, Results, and Analysis). Establishing the test conditions should be done in one or two sentences and should focus on what testing was being conducted when the problem was discovered. Specifically, the writer should describe the mission or task that was being performed.
- What are the data/test results that support the conclusion (qualitative and quantitative test results)?
  - The data and results presented here follow the test condition sentence(s) and is the second part of paragraph 2. Only present data related to the issue. The Blue/Gold sheet should stand alone and no reference to the data analysis summary memorandum should be made. Pull the specific data/results needed to make the case to the reader from the data analysis summary. Photos, screen shots, figures, and tables with detailed annotation are encouraged. When annotating figures, photos, or screenshots, use the same wording used in paragraph 2 to allow the reader to easily follow your discussion.
- Can I help the reader understand the cause(s) of the problem (analysis)?
  - The information/analysis of the results presented here forms the third part of paragraph 2 and should logically flow into the mission relation. Use the data and results presented to explain the impact upon the operator, mission, or task/subtask.
- Recommendation (timeline for correction).
  - Now, the sheet should be assembled and checked for logic and a final proofread:
- Are all parts of the sheet present?

- Does the argument make sense?
- Does the sheet present the message you really want to convey? Does the discussion lead logically to the conclusion?
- Proofread the draft sheet. Look for and correct typographical errors, improper verb tenses, and other grammatical errors.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C - PMT View

1. Background

The Platform Mission Tasks View (PMT View) is a graphic depiction of the system's current, evaluated, mission-based capability. As the testing continuum progresses, and the PMT View is populated, it provides a common, data-driven, and shareable perspective that is a useful reference on multiple levels across the testing continuum. The PMT View is intended to provide a standardized format that is relatively easy to update and share with stakeholders.

The purpose of this appendix is to describe the use of PMT Views to support the Operational Test Reporting process at COMOPTEVFOR, including PMT View usage during PTIP milestone meetings. It is assumed that the program has an established and approved PMT View shell (PV-0 View) that is available following completion of the MBTD process. The process details of the format, content, generation, and approval of the PV-0 are contained within the IEF Checklist, and it is prepared by the test team with the guidance of 01B. Once the test team begins updating the PMT to reflect observed system performance, the PV-0 (PMT View Shell) transitions into a PV-1 (PMT Performance View), which is the primary visual depiction of the status of T&E strategy and results to date.

For programs that completed the MBTD process before the introduction and adoption of PMT Views, the test team should generate a PV-0 using the guidance from the IEF Checklist, and the PV-0 should be approved by the Warfare Division Director and 01B. The PV-0 Shell should be completed prior to commencement of test to allow PV-1 Performance View development and use throughout the test phase and the PTIP process.

2. PMT View Variants

PMT Views are used as the visual depiction of the capability of the system under test as observed during test. As such, the views are allowed to be tailored, to a certain extent, by the test teams in order to "tell the story" of system capability during the SERB and E-SERB and ultimately to outside stakeholders. However, because PMT Views are used across all programs at COTF, standardization of the views ensures the information on the views is consistently interpreted.

The PMT View is not a DODAF-defined view, such as an Operational View (OV), Capability View (CV), or System View (SV). However, for ease of understanding the variants of PMT Views, DODAF-like nomenclature is used for ease of reference of the various products.

The following PMT View variants are the standard views currently defined. This is not an exhaustive list of all the possible use cases, and test teams are encouraged to be creative when developing their tailored methodology to display information. However, these variants are defined to allow for consistent baseline use across all programs. The initial set of PMT View variants used during the PTIP process includes the PV-0, PV-1A, PV-1B, PV-1C, and PV-1D.

a. PV-0 PMT View Shell

The uncolored PMT View, referred to as PV-0, is a set of Excel spreadsheets organized within a workbook. The PV-0 includes an Excel tab for each Effectiveness COI, and a single tab

containing all Suitability COIs. Each of these tabs contains all of the tasks, including all subtasks, measures, and conditions developed in the MBTD process for each COI. The PV-0 is uncolored, but it provides the format to be used to create a set of information views as the test program progresses. Therefore, the PV-0 is a baseline depiction of the Operational Test requirements for a given SUT. It graphically displays the MBTD-developed test design.

One additional spreadsheet tab is created within the initial PV-0. This is the "Summary" tab, containing a 'roll-up' graphical display of each COI, along with the first level tasks of each COI. This tab provides the initial format for the PV-1C view, described below, and provides summary status information to decision makers at a high level, providing a snapshot of the entire system in a single view.

b. PV-1 PMT Performance View

The PV-1 Views are designed to display system performance information. It is CRITICAL that these views be clearly marked as PMT Performance Views (PV-1A/1B/1C/1D) because the designation uses a prescribed and standardized colorization scheme representing specific and defined status for the performance of the system. The color scheme is defined in detail below in paragraph 3. The applicable PV-1 designation will ensure that information displayed will be commonly understood and not be misinterpreted. Examples of PV-1 views are included in the training material for the PTIP course and the template for the SERB and E-SERB briefings.

Four sub-views are defined below for the following specific use cases:

- **PV-1A PMT Performance View (COI Level)**

This view includes the entire PV-0, populated and updated in accordance with ESTABLISHED and UNIFORM GUIDELINES for coloring borders and cells. For some complex programs, this view can require many excel tabs to display all required information. At a minimum, it will contain all effectiveness COIs as individual tabs, as well as the suitability tab.

- **PV-1B PMT Performance View (COI Tasks)**

This view includes only the Task Structure of the PV-0. Specific measures and conditions are excluded from the view, and only the status of the First-Level Tasks and Second/Third level subtasks are presented. This allows a more concise display of high-level capability performance. PV-1B Views are not mandatory for use during the PTIP. They are optional and should be considered for use when their use would facilitate better understanding of system performance. Details desired as a result of the PV-1B review will refer to the PV-1A views for additional information.

- **PV-1C PMT Performance View (System Level)**

This view includes only the "Summary" tab of the PV-1, and includes only the COIs and the First Level Tasks of the COIs. This arrangement allows a single page display of system capability performance, and can be useful for display of high-level information to executive decision makers. Additional details desired as a result of the PV-1C review will refer to other PV-1 views for additional information.

- **PV-1D PMT Performance View (Response Variable Detail Level)**

This view does not have a prescribed format. It will be designed by the team developing the MBTD to graphically display the performance of the SUT in the case of specific RVs. It is intended to provide additional factor performance detail of individual RVs and be used as a supplement to the PMT-1A View.

- **PV-1 Program Information Tab**

The first tab in the spreadsheet is always "Program Information" in order to provide a quick reference for everyone using it. The format and content of this tab may be modified by the test team.

3. PV-1 PMT View Formatting

The content, format, and color scheme of all PV-1 Views are standardized to provide consistency in information presented to decision makers. For example, a green cell border means the same thing in any system PV-1, and a yellow colored cell means the same thing in any system PV-1.

a. PV-1 Measure and Task Fonts

Measure fonts are standardized and should not be modified by the test team:

- OTA Critical Measure (***BOLD ITALIC***)
- Program KPP, KSA, or CTP Measure (<u>UNDERLINED</u>)
- OTA Critical AND Program KPP, KSA or CTP (***<u>BOLD ITALIC UNDERLINED</u>***)
- Not an OTA Critical Measure, KPP, KSA, or CTP (Normal)

Task fonts are standardized and should not be modified by the test team:

- OTA Critical Task (***BOLD ITALIC***)
- Not an OTA Critical Task (Normal)

b. PV-1 Legend

Task and Measure colorization is the method by which system test results are displayed. The standard PV-1 Legend is shown in figure C-1 below, and should be used for all PV-1A, -1B, and -1C views. The PV-1D view is not standardized, so the legend may be modified for that view, however it should remain consistent with the standard PV-1 Legend.

Figure C-1. PV-1 Border/Fill Color/Font Standard Legend
UNCLASSIFIED

PV-1 (PERFORMANCE PMT VIEW) BORDERS - FILL COLOR - FONT LEGEND		
WIPT ASSESSED BORDER COLORS		Trending MEASURE MET or TASK CAPABLE. (Heavy GREEN Border)
		Trending MEASURE MET FOR CERTAIN CONDITIONS or TASK PARTIALLY CAPABLE. (Heavy YELLOW Border)
		Trending MEASURE NOT MET or TASK NOT CAPABLE. (Heavy RED Border)
OTA ASSESSED CELL FILL COLOR		Sufficient data to evaluate as MEASURE SUPPORTS TASK or TASK CAPABLE.
		Sufficient data to evaluate as MEASURE PARTIALLY SUPPORTS TASK or TASK PARTIALLY CAPABLE.
		Sufficient data to evaluate as MEASURE DOES NOT SUPPORT TASK or TASK NOT CAPABLE.
		Response Variables (measures) or Response Variable Factors (conditions)
		Measures and Tasks out of scope for current test phase
MEASURE and TASK FONT FORMATS	M28	Not an OTA Critical Task or Measure, Not a KPP, KSA, CTP
	M29	OTA Critical Task or Measure (BOLD ITALIC)
	M30	Program KPP, KSA, or CTP Measure (UNDERLINED)
	M31	OTA Critical AND Program KPP, KSA or CTP (BOLD ITALIC UNDERLINED)

c. Measure and Task Border Colors

Task and Measure border colors are meant to display the collaborative assessments of the PMO and/or the WIPT. Borders will be used extensively for programs incorporating Capabilities Based Test and Evaluation (CBTE) as their T&E strategy. When CBTE is utilized, the test team should include the border colors determined by the PMO in the PV-1 for the program. Additional guidance regarding border color usage can be obtained from the program Lead Test Engineer or 01C Representative.

d. Measure and Task Fill Colors

Task and Measure fill colors are used to display the Operational Test results developed by the test team during OT.

4. PMT View Usage During the PTIP

PMT Views, like the COI evaluation matrix used before, are the primary display tools used to visually display the results of test during the PTIP. The following general guidance applies:

a. PMT Performance View (PV-1) Development

PV-1 development begins with the PMT View Shell (PV-0) developed during the MBTD process and output from the MBTES tool. In the case of CBTE programs, it is possible that the PV-1 has been utilized by the PMO and/or Developmental Test team prior to the COTF PTIP process beginning. If that is the case, then the most recent PV-1 should be obtained from the PMO and used as the starting point. PMO/DT assessments will be present on the PV-1 in the form of border colorization of measures and tasks within the PV-1. These PMO/DT assessments in no way restrict the COTF analysis or the results obtained. COTF analysis results are displayed on the PV-1 through the use of cell fill colors as described in paragraph 3 above. The combination of the cell borders from PMO/DT results and cell fill colors from COTF PTIP results can provide useful information and perspective that will be of interest to COTF leadership, and could highlight issues worthy of discussion within the test report.

In the case of non-CBTE programs, the uncolored PV-0 will be used as the starting point for PV-1 development. For non-CBTE programs, cell border colorization is not required for PV-1 views.

The PV-1 view will primarily be updated with results prior to and during the CEWG meeting(s) during the PTIP.

b. PMT Performance View (PV-1) Finalization and Approval

Completed PV-1 Views will be presented to the Warfare Division Director during the SERB meeting(s), and approved during the E-SERB with the Admiral.

Based on the test team's recommendations, the PV-1 may or may not be included in the actual test report document. This decision is at the discretion of the Warfare Division Director, and is approved by the Admiral.

c. PMT View Lessons Learned

PMT Views are an effective way to graphically present the capability of a system as observed during operational test. Lessons will be collected and learned as they are used in test reporting. Recommendations for improvement in the design, format, and use of PMT Views are strongly desired. Test teams are encouraged to provide feedback and recommendations via your program 01C representatives and Lead Test Engineers.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX D - Acronyms

00TD	Technical Director
01B	OPTEVFOR Test Design Competency
01C	OPTEVFOR Test Planning and Evaluation Competency
01D	OPTEVFOR Cybersecurity Test and Evaluation Competency
6PP	Six Part Paragraph
ACOTD	Assistant COTD
ACTD	Advanced Concept Technology Demonstration
AO	Action Officer
AOC	Assessment of Operational Capability
AWG	Analysis Working Group
B&G	Blue and Gold
CDD	Capabilities Development Document
CEWG	COI Evaluation Working Group
CNA	Center for Naval Analysis
CNO	Chief of Naval Operations
CO	Commanding Officer
COI	Critical Operational Issue
COMOPTEVFOR	Commander, Operational Test and Evaluation Force
CONOPS	Concept of Operations
COR	Contracting Officer Representative
COTD	Chief Operational Test Director (VX Test Position)
CPD	Capabilities Production Document
CTF	Core Team Facilitator
DC	Desired Capability (associated with JCTD programs)
DCP	Data Collection Plan
DED	Demonstration Execution Document (associated with JCTD programs)
DMOT	Detailed Method of Test
DoD	Department of Defense
DOE	Design of Experiments
DoN	Department of the Navy
DOT&E	Director, Operational Test and Evaluation
DT	Developmental Test
EA	Electronic Attack

E-SERB	Executive SERB
EOA	Early Operational Assessment
FOT&E	Follow-on Test and Evaluation
IEF	Integrated Evaluation Framework
IOC	Initial Operational Capability
IOT&E	Initial Operational Test and Evaluation
IR	Interim Report
IT	Integrated Testing
JCTD	Joint Capabilities Technology Demonstration
KPP	Key Performance Parameter
KSA	Key System Attribute
LMUA	Limited Military Utility Assessment
LOO	Letter of Observation
LTE	Lead Test Engineer
MBTD	Mission Based Test Design
MOA	Memorandum of Agreement
MOE	Measure of Effectiveness
MOS	Measure of Suitability
MS	Milestone
MUA	Military Utility Assessment
NIPRNET	Non-Secure Internet Protocol Router Network
NLT	No Later Than
OA	Operational Assessment
OAR	OTA Assessment Report
OER	OTA Evaluation Report
OFER	OTA Follow-on Evaluation Report
OM	Operational Manager (associated with JCTD programs)
OMAR	OTA Milestone Assessment Report
OMF	Operational Mission Failure
OP	Operational Problem (associated with JCTD programs)
OPCON	Operational Consideration
OPNAV	Office of the Chief of Naval Operations
OPTEVFOR	Operational Test and Evaluation Force
OT	Operational Test
OT&E	Operational Test and Evaluation

OTA	Operational Test Agency
OTC	Operational Test Coordinator (Air Warfare Division Only)
OTD	Operational Test Director
OUA	Operational Utility Assessment
POA&M	Plan of Action and Milestones
PM	Program Manager
PMT	Platform Mission Tasks
PTIP	Post-Test Iterative Process
QRA	Quick Reaction Assessment
RV	Response Variable
SAT	Satisfactory
SERB	System Evaluation Review Board
SH	Section Head
SIPRNET	Secure Internet Protocol Router Network
SITREP	Situational Report
SOP	Standard Operating Procedure
SoS	System of Systems
SUT	System Under Test
T&E	Test and Evaluation
TEIN	Test and Evaluation Identification Number
TEMP	Test and Evaluation Master Plan
TSOT	Total System Operating Time
TTP	Tactics, Techniques, and Procedures
UNSAT	Unsatisfactory
VCD	Verification of Correction of Deficiencies
VTC	Video Teleconference
VX	Test and Evaluation Squadron

THIS PAGE INTENTIONALLY LEFT BLANK