412TW-TIH-22-02

TEST REPORT AUTHOR'S GUIDE

Office of the Technical Director

NOVEMBER 2022

TECHNICAL INFORMATION HANDBOOK

DISTRIBUTION A. Approved for public release; distribution is unlimited.
412TW-PA-22372

412TH TEST WING
EDWARDS AIR FORCE BASE, CALIFORNIA
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

This technical information handbook (412TW-TIH-22-02, *Test Report Author's Guide*) replaces AFFTC-TIH-14-01, *The Author's Guide to Writing 412th Test Wing Technical Reports* (Reference 1). This handbook was submitted by the Technical Director, 412th Test Wing, Edwards AFB, California 93524-6843.

Prepared by:

CARLSON.DANIEL.C.1394878843 Digitally signed by CARLSON.DANIEL.C.1394878843 Date: 2022.11.06 05:30:17 -08'00'

DANIEL CARLSON
AFTC Detachment 1/MDTF

LIEBMANN.CHRISTOPHER.J.1387297010 Digitally signed by LIEBMANN.CHRISTOPHER.J.1387297010 Date: 2022.11.04 16:26:11 -07'00'

CHRISTOPHER J. LIEBMANN
773rd Test Squadron

MAILEN.STEPHANIE.SHARP.1386619952 Digitally signed by MAILEN.STEPHANIE.SHARP.1386619952 Date: 2022.11.07 10:21:11 -08'00'

STEPHANIE S. MAILEN
445th Test Squadron

MARTIN.JEFFREY.D.1369066478 Digitally signed by MARTIN.JEFFREY.D.1369066478 Date: 2022.11.07 05:04:26 -08'00'

JEFFREY D. MARTIN
773rd Test Squadron

PECK.KEVIN.L.1250713017 Digitally signed by PECK.KEVIN.L.1250713017 Date: 2022.11.07 12:03:33 -08'00'

KEVIN L. PECK
775th Test Squadron

This handbook has been reviewed and is approved for publication: 16 November 2022

OSBURN.DANIEL.W.1241451239 Digitally signed by OSBURN.DANIEL.W.1241451239 Date: 2022.11.16 12:08:09 -08'00'

DANIEL W. OSBURN
Technical Director
412th Test Wing

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to the Department of Defense, Executive Service Directorate (0704-0188). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(FROM - THROUGH)*
16-11-2022	Technical Information Handbook	N/A

4. TITLE AND SUBTITLE

Test Report Author's Guide

5A. CONTRACT NUMBER

5B. GRANT NUMBER

5C. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Carlson, Daniel, Project Engineer
Liebmann, Christopher J., Project Engineer
Mailen, Stephanie S., Project Engineer
Martin, Jeffrey D., Project Engineer
Peck, Kevin L., Project Engineer

5D. PROJECT NUMBER

5E. TASK NUMBER

5F. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
412th Test Wing
Edwards AFB, CA 93524-6843

8. PERFORMING ORGANIZATION REPORT NUMBER
412TW-TIH-22-02

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
412th Test Wing
307 East Popson Ave
Edwards AFB CA 93524

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NO.(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A. Approved for public release; distribution is unlimited. 412TW-PA-22372.

13. SUPPLEMENTARY NOTES
SC: 012100 (This number comes from STINFO via EAFB Technical Library)
CA: 412th Test Wing Edwards AFB CA
Print this document in COLOR.

14. ABSTRACT
The purpose of this handbook is to promote consistency across 412th Test Wing (412 TW) test reporting products. Test reports are the formal products the 412 TW uses to communicate results, conclusions, and recommendations from developmental test efforts.

15. SUBJECT TERMS
Reports; test plans; technical reports; handbook; technical writing; writing instruction manual; AFTC(Air Force Test Center); 412 TW(Test Wing)

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19A. NAME OF RESPONSIBLE PERSON
A. REPORT	B. ABSTRACT	C. THIS PAGE			Daniel W. Osburn Senior Technical Director, 412th Test Wing
U	U	U	Same as Report	40	19B. TELEPHONE NUMBER (INCLUDE AREA CODE) 661-275-4000

STANDARD FORM 298 (REV. 8-98)
PRESCRIBED BY ANSI STD. Z39.18

TEST REPORT AUTHOR'S GUIDE CHANGE LOG

Date (DD-MM-YYYY)	Document	Description of Significant Changes
16-11-2022	412TW-TIH-22-02	Full rewrite of 412TW-TIH-14-01

TABLE OF CONTENTS

INTRODUCTION

PURPOSE AND SCOPE OF THIS HANDBOOK

The purpose of this handbook is to promote consistency across 412th Test Wing (412 TW) test reporting products. Test reports are the formal products the 412 TW uses to communicate test results, conclusions, and recommendations from experimentation and developmental test and evaluation (DT&E) efforts. Most commonly, 412 TW reports recommend whether the system under test (SUT) is ready for the next step in the acquisition process.

This handbook uses "test report" to refer to any test reporting product, including technical reports (TRs) preliminary reports of results (PRRs), and others (as listed in Appendix A). This handbook contains basic test report philosophy and process guidance, and describes TRs in depth, as the TR is the most comprehensive and commonly used type of test report. However, most concepts in this guide are applicable to all types of test reports. This handbook is a companion to the 412 TW TR template, which offers section-by-section content and structure guidance. All authors should read this handbook before using the template. Unit-specific guidance and consultations with technical experts, platform subject matter experts, and the chief engineer should provide additional detail to complement this handbook. For the most current MS Word version of the template, consult your squadron's technical editor or the Technical Publications Office.

The guidance in this handbook was written to complement EdwardsAFBI 99-103, *412 TW Technical Report Program* (Reference 2), which defines the reporting process, policies, and procedures, as well as the various types of reporting products. In all cases, reporting products should comply with current Technical Reporting and Scientific and Technical Information (STINFO) Program regulations and guidance.

This handbook does not include guidance for test report classification markings. Such guidance is provided by DoDM 5200.01, Volume 2, *DoD Information Security Program: Marking of Classified Information* (Reference 3). If a portion of test report content is expected to be classified, test teams should consider whether to classify the entire test report or to place all classified content in an appendix published under separate cover, leaving the bulk of the test report unclassified.

KEY CONCEPTS FOR SUCCESSFUL 412 TW TEST REPORTING

1. **Know the test report's audience(s) and purpose(s).** The report integrated product team (RIPT) should ensure that audience needs are met when structuring the report content. The audience for 412 TW test reports includes customers (usual program office decision makers) and other stakeholders such as the warfighter, operational test and evaluation (OT&E) personnel, vendors, and future DT teams. The RIPT should discuss where the test report fits in the SUT's overall acquisition cycle, consider the next recipient of the SUT (e.g., additional DT, OT, or direct fielding).

   Authors should always select a report type that meets the customer's needs. If a different report type is required from the approved test plan, the author and RIPT should coordinate that change appropriately.

2. **Start early.** Typically, once the last test data are received the deadline for report delivery is set. Therefore, authors should begin their first draft before testing is complete. The technical editor can start building the report shell as soon as the test package has been approved. Authors should identify

and solicit input from RIPT members early, especially the RIPT co-chairs. Appendix C provides an overview of the reporting process.

3. **Leverage prior work wisely, balancing efficiency with innovation.** The RIPT is rarely doing the first-ever test report of its kind and should not reinvent the wheel. Authors should consider using previously vetted language (sources include 412 TW technical documents, flight manuals, etc.) to describe concepts rather than writing original words. New authors should consult experienced personnel for in-depth content and formatting suggestions. Authors are strongly cautioned against blindly copying and pasting content.

   There is a healthy tension between efficiency and innovation. Similar program-to-program formatted/structured reports provide familiarity and are therefore faster for authors to produce and customers to consume. However, an innovative format, plot, report structure, media, test objective, MOP, etc., may communicate test results to the customer so much better that it's worth the extra time and effort. Innovation should be encouraged and not automatically rejected on the basis of unfamiliarity.

4. **Present an unbiased assessment.** Test results, whether positive or negative, should be presented fairly, accurately, and impartially; test reports should advocate neither for nor against the SUT. Disparaging the contractor, PO, test participants, or the 412 TW undermines the credibility of the report.

5. **Content is the highest priority.** Format should support test report content rather than drive it. EdwardsAFBI 99-103 (Reference 2) provides a standardized set of reporting options (summarized in Appendix A of this handbook); however, flexibility is encouraged in order to satisfy the customers' needs. If a new/modified format better communicates the content, authors should propose it to the RIPT.

6. **Be clear, accurate, and concise.** Customers often depend on test reports to make timely, critical decisions. Rather than communicating every detail the team knows, a test report should focus on communicating what the customer needs to know. Wordy, dense writing may be confusing and a waste of the audience's time. Although specialized technical terminology (i.e., jargon) can be useful shorthand within a particular group, these terms are often meaningless or confusing to the larger audience; everyday language should be used whenever possible in test reports.

7. **Substantiate conclusions and recommendations with data.** The conclusions, ratings, and recommendations are key elements in the report. Therefore, authors should ensure conclusions are defensible, replicable, traceable to data, and that they answer the questions posed by the test objectives (Figure 1). When test results seem to point to different conclusions from what the report states, authors should clearly explain the logic or justification for their conclusions. Authors should avoid speculating, exaggerating, or inappropriately extrapolating data to draw hypothetical or indefensible conclusions.

8. **Convey the value of test.** A well-written test report improves customers' insight into the SUT and highlights the value of test and evaluation expertise at the 412 TW. Reports should provide context not readily available, such as background information, robust engineering analysis, required workarounds, and/or military utility. The test reports should detail how well the SUT performed its intended purpose, often with emphasis on the operators' perspective.

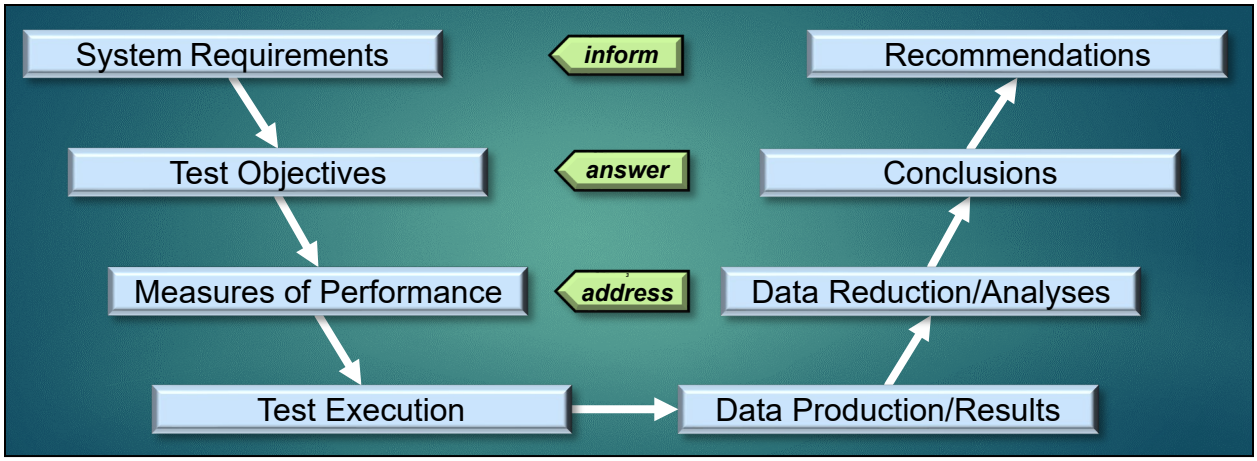

Figure 1 Requirements Traceability

TEST REPORT FRONT MATTER

OUTSIDE FRONT COVER

The 412 TW test reports are official U.S. Government publications; the front covers reflect the professionalism of the 412 TW and the USAF. The cover is readers' first impression of the contents, and a photograph or professional illustration of the aircraft or SUT is encouraged. Cartoons or personalized logos are inappropriate.

- **Document Number –** The official, 412 TW-assigned document number (e.g., 412TW-TR-99-99 or 412TW-PRR-01-01); the number is assigned via the technical editor and the Technical Publications Home Office.
- **Title –** The title should be brief and descriptive of the test project, matching or reflecting the relevant test plan. If at all possible, the test report title should be unclassified, regardless of the document's overall classification in order to facilitate administrative use such as tracking, searches, etc.
- **Authors –** Test reports typically include project engineers and project pilot/operator/aircrew. If the aircrew is not a major contributor, they are not required to be on the cover. Personnel appearing on the cover are typically government; contractor authorship of test reports should be coordinated with the 412 TW/CT office.
- **Report Type –** Indicates the report type as it appears in the Standard Form (SF) 298 Form Field 2.
- **Report Date –** The month and year the report is approved by 412 TW/CC or their delegate.
- **Distribution Statement –** All 412 TW test reports must contain one of the official distribution statements on the cover IAW DoDI 5230.24, *Distribution Statements on Technical Documents* (Reference 4), as implemented by AFI 61-201, *Management of Scientific and Technical Information (STINFO)* (Reference 5). The distribution statement type is selected by the RIPT and customer. Further guidance may be found by searching "distribution statements" at the Defense Technical Information Center (DTIC) website (https://discover.dtic.mil).
- **Document Control Markings** – All documents requiring control markings, such as Controlled Unclassified Information (CUI), must follow DoDM 5200.01 guidelines (Reference 3). The TR template reflects the most recent guidance, and the technical editor will help ensure the control markings are complete and correctly formatted.

INSIDE FRONT COVER (SIGNATURE PAGE)

The purpose of the signature page is to document who wrote the test report and approved its publication. Major contributors must be conversant with the key points of the document and will acknowledge their concurrence by signing the signature page.

The primary author and major contributors should be listed on the left-hand side, with any support contractors listed last and identified as such. The approval authority signature(s) on the right-hand side are listed in EdwardsAFBI 99-103 (Reference 2).

STANDARD FORM (SF) 298

The SF 298 is required (per DoDM 3200.14, Volume 1, *Principles and Operational Parameters of the DoD Scientific and Technical Information Program [STIP]: General Processes* [Reference 6]) for all documents delivered to DTIC, except those classified higher than Secret collateral. The TR template will indicate which fields are required and their proper format.

QUALIFIED REQUESTORS AND EXPORT CONTROL STATEMENTS

The qualified requestors and export control statements (standard and/or program-specific) are required for documents not cleared for public release. The format and wording for these statements are governed by AFI 61-201 (Reference 5).

ACKNOWLEDGMENTS

The optional Acknowledgments section is for author(s)' recognition of individuals who have provided support for test efforts, especially those who went above and beyond to achieve test completion, such as: ground crew who supported long or complex tests; contractors who provided expertise in special equipment or test techniques; range, backshop, and/or depot personnel who facilitated test execution or provided special test equipment; and program management or PO personnel who supported test efforts.

EXECUTIVE SUMMARY

The Executive Summary section is an overview of the TR and targets readers who may not read the entire report, particularly general officers and other program decision-makers. This separate section appears just prior to the Table of Contents and should not exceed one page. The Executive Summary is intended to closely reflect content from the Introduction and Test and Evaluation (T&E) sections, and should be written after those sections have been completed. Much of the content is copied directly (opening paragraph, overall recommendation) or summarized (Background, Test Item Description, Overall Test Results) from the main body of the test report, and should not include any new information (i.e., information that has not been addressed elsewhere in the report). The Executive Summary contains five primary sections/paragraphs:

1. Opening Paragraph (Copied from Introduction Section)
2. Background (Summarized from Introduction Section)
3. Test Item Description (Summarized from Introduction Section)
4. Overall Test Results (Copied from T&E Section)
5. Overall Recommendation (Copied from T&E Section)

Very short or simple TRs may not warrant an Executive Summary. For these reports, the Executive Summary may be omitted with RIPT and 412 TW Technical Director concurrence.

TABLE OF CONTENTS (TOC)

The TOC is included in the TR template. If the TOC does not add value (e.g., in very short reports), it may be omitted.

TEST REPORT MAIN BODY

INTRODUCTION

The Introduction section provides an overview of the test project and includes test scope, background, test objectives, and any test limitations or constraints encountered. This section does not contain test results, conclusions, or recommendations, but should indicate whether or not the overall test objective was met.

Introduction Section Opening Paragraph(s):

This section contains standardized content for easy readability. Generally, a single paragraph is used to convey the basic programmatic context of the test report, and can be copied from the test plan (Section 1.1 Overview) with minor modifications. The following elements should be included:

- Standard introduction sentence
- The overall test objective
- The customer(s)
- Test organization(s), as appropriate. More information on these organizational definitions is found in DoDI 5000.89_DAFI 99-103, *Capabilities-Based Test and Evaluation* (Reference 7):
  - Lead developmental test and evaluation organization
  - Executing test organization
  - Participating test organization(s)
  - Operational test organization(s)
- Test stakeholders
- Test location(s)
- Test date(s)
- Test scope (number of ground/flight test hours, test points, sorties, or other applicable metric[s])
- A statement indicating the status of the overall test objective (Choices for this statement are met, partially met, or not met. [If partially met or not met, a short one- or two-sentence discussion of the limitation should immediately follow. The Limitations and Constraints section should contain the complete discussion.])

Background:

The Background section should provide a summary of relevant program history leading up to the test. The majority of this section may be derived from the test plan's Background section. It may be necessary to include additional information such as any technical concepts important to understanding the test results. Current test results should never be included in the Background section.

It is not uncommon for a test program to undergo a significant change in scope (e.g., changes to program schedule, test objectives, or unplanned software releases) in the time between initial approval of the test package and the end of test execution. Although such changes should have been captured in the test package via the amendment process, major changes in technical scope may be briefly summarized in the Background section of the test report.

Test Item Description:

The Test Item Description (TID) section should focus on what the SUT was designed to do rather than how the SUT actually performed; test results belong in the T&E section. The TID is derived largely from the same section of the test plan. If the TID section of the test plan is large (several pages or more), the system under test can be summarized in this section with in-depth descriptions included in a Detailed TID

appendix. However, if the TID section of the test plan is concise and sufficient, this test report section can be reproduced from the test plan.

If the SUT changed during test, this section should be updated to describe the intermediate and final iterations of the SUT so that the reader can connect the test results to the corresponding test item configuration at the time it was tested. This could include a discussion or tabular presentation of specific test item configurations such as software versions, mission data files, and/or hardware part numbers that were used for specific test events or flights throughout the test program.

Limitations and/or Constraints:

This section addresses items first identified in the Potential Impacts to the Completion Criteria section of the test plan and/or unforeseen impacts discovered during testing. Asset availability or environment issues may be addressed, and should be described without placing blame (e.g., "There was insufficient funding," is preferable to "The PO did not provide funding."). Further examples are provided in Example 1. Once limitations/constraints are addressed here, the test results sections should refer back to the Limitations and Constraints section and not repeat the content.

Limitations prevent a test objective from being fully met. If a limitation resulted in a test objective being only partially met, that test objective should be identified and the limitation(s) described. Even if some GTOs or STOs were not met, the overall test objective may still have been met. If every test objective was fully met, there were no limitations. If the overall test objective was not met as a result of a limitation, this should be extensively described, as it represents a major impact to test effectiveness.

Constraints restrict the scope of activity but do not prevent meeting test objective(s). Constraints can affect several areas or test objectives. Constraints may entail altered statistical confidence or power, or reduced ability to assess military utility. Test objectives may also be partially met; some conclusions/recommendations can be drawn, often with a caveat bounding their scope. If the scope of activity was not restricted beyond what was originally described in the test plan, there were no constraints.

In some cases, not all test points/MOPs/test objectives can be completed prior to the end of a test program. Instead of translating that gap in testing to a limitation/constraint in the TR, those test points/MOPs/objectives may be deleted via a test plan amendment. This could lead to a cleaner test report, but may also risk telling an incomplete story. Because each circumstance is unique, this should be discussed with the RIPT and 412 TW Technical Director to determine the best course of action. If a test plan amendment is the preferred route, the amendment should be started as soon as the team is aware of the problem, rather than waiting until the end of execution.

If there were no limitations or constraints, omit this section from the report.

EXAMPLE 1

LIMITATIONS:

Deferred to Addendum: Due to limited test range availability during the NAVWAR test event, testing of the all-up round was deferred until October 2025. Evaluation of the integration of the all-up round (General Test Objective 2) will be provided in an addendum to this report.

Programmatic Decision: Due to limited test aircraft availability, the JP-X alternative fuel was not tested on the B-10A aircraft with legacy F105 engines installed. This prevented the team from demonstrating the JP-X effects on F105 engine operability (Specific Test Objective [STO] 3.2). As the majority of the B-10A fleet is equipped with F107 engines, the PO accepted the risk of not performing this testing on B-10A with F105 engines.

Test Environment: The weather radar performance (STO 1.5) could not be rated; the lack of weather cells during test precluded assessment of weather cell detection accuracy and wet turbulence detection.

Incomplete Data Collection: Due to corruption of the in-flight data recorder, fault codes and health status of the updated flight control computer were not collected during flight testing, resulting in incomplete collection of the flight control computer data (STO 3.1).

CONSTRAINTS:

Reduced Sample Size: Twenty observations were required per the test plan (Reference 99) to achieve necessary statistical confidence. Schedule constraints resulted in the collection of only 15 observations of the mean angle-of-arrival performance, which resulted in a confidence level of 90 percent instead of the planned 95 percent confidence. Although this constraint affected the mean angle of arrival performance MOP, performance of the electronic warfare suite (STO 1.3) was still able to be rated.

Alternate Data Assessment: The suppression-of-enemy-air-defense (SEAD) mission military utility assessment was constrained due to a lack of range availability. Rather than using an operationally representative scenario, the military utility assessment was based on individual sensor STO ratings and performance instead, which was sufficiently thorough to rate military utility (STO 4.1).

TEST AND EVALUATION (T&E)

The T&E section is the core of the TR and begins with a summary of the test results and overall recommendation. Detailed test results, analyses, conclusions, and recommendations follow. To reduce the author workload and increase readability, the TR template provides multiple examples of how to organize the T&E section.

Fundamental T&E Reporting Elements:

Specification compliance, regression, military utility, deficiency reports (DRs), and recommendations are some fundamental elements that should be understood before writing the T&E section of the TR.

Specification Compliance.

Specification compliance assesses the performance of the SUT against contract specifications or other documented requirements, and is usually a primary driver in the rating of the SUT. Test results may be

deterministic or stochastic depending on the nature of the SUT performance. Details of the data analysis, evaluation criteria, and statistical methods should have been defined in the test planning process, but the RIPT should continue to consult the Statistical Methods Flight on the best way to analyze, present, and interpret the results during execution and reporting. A statement indicating the impact to military utility should be included, particularly when specifications have not been met.

Regression.

Regression testing assesses the impacts of a new system integration on legacy functionality or performance by comparing the current system operation to a baseline, or previous system functionality or performance. Regression testing assumes no measurable/detected change in functionality or performance exists until proven otherwise; this can be expressed as the null hypothesis in a statistical inference test. The alternative hypothesis would then state that a measurable change in functionality/performance exists, and the point estimate of the relevant metric would be bounded with a statistical interval. See the Regression Comparisons section below for guidance on how to present regression test results in the test report. Findings outside the scope of the test objectives should be reported in an Additional Findings section.

Military Utility.

Military utility content describes how well the SUT accomplishes its operational mission. Military utility may factor into a test objective's rating and MOP-level discussions/descriptors. Military utility content should be addressed throughout the TR as each test objective's results are presented, even if the test plan included a dedicated military utility test objective.

In determining the impact of military utility on the test objective's rating, the author should carefully balance the effectiveness of each SUT element as well as the whole of the weapon system against the ability to accomplish the overall mission. This content typically requires substantial input from, or may be written by, the test aircrew and/or aircrew with operationally relevant experience. For example, software deficiency in an avionics subsystem that prevents the sensor operator from using a particular function could inhibit the ability to prosecute a target. In another example, an unexpected flight envelope restriction could prevent the pilot from accomplishing a needed combat maneuver.

Watch Items (WITs), DRs, and Publication Change Requests.

Testing often exposes deficiencies, which may be documented as WITs, DRs, and/or publication change requests.

A WIT is an issue discovered during testing that warrants further observation. A DR is an action tool to report adverse conditions or proposed enhancements discovered during the course of testing or operations. The DRs also serve as risk management tools for the customer. In most cases, WITs should either be closed or elevated to DRs by the time the final report is written. Preliminary reports of results and early test reports in multi-phase test efforts may list WITs only, with subsequent test reports documenting the WITs' final outcomes. Requested improvements to systems that meet specifications/requirements (aka enhancement DRs) should be treated the same as other DRs in the test report. The DRs are generally included in the test results sections of the GTOs/STOs that they affect.

Discrepancies in manuals or other procedural documents may be identified during the course of testing, necessitating publication change requests, which may be formally submitted using AF Form 847 or a similar process IAW T.O. 00-35D-54, *USAF Deficiency Reporting, Investigation, and Resolution (DRI&R)* (Reference 8).

All DRs and publication change requests referenced in the text should be included in a dedicated appendix. In most cases, all DRs referenced in an appendix should be addressed in the test results and include the DR's impact to military utility and recommendation. In cases with multiple minor DRs in one Test Results section, they may be addressed and listed as a group with only one recommendation at the end of the list.

Reports should not include contractor deficiency tracking numbers without having an associated DR number. A 412 TW-approved DR summary format is provided in the TR template. A table of DR numbers, titles, and location(s) discussed in the report (page numbers) is provided if more than two DRs are included; the same guidance applies for publication change requests. If previously reported DRs are relevant and impact the rating of the current test program or are reencountered during testing, it is appropriate to reference the DR again in the new test report.

Recommendations.

Recommendations identify what needs to be fixed, without specifying how fixes should be accomplished. The word "must" is reserved for health and safety issues, and should include information on the potential for harm.

Overall Recommendation – Makes a statement of the 412 TW's position on the programmatic next step of the SUT and answers the purpose of the entire test. The overall recommendation is the final sentence or two at the end of the Test Results Summary paragraph (which also appears as the final paragraph of the Executive Summary).

Specific Recommendation – Makes a targeted statement regarding a particular aspect/component of the SUT or test program that warrants further action such as additional testing or research by customers or stakeholders. Specific recommendations are located in GTO/STO Test Results and/or Additional Findings sections, and typically follow DRs, manuals and/or T.O. changes, or other conclusions as determined by the RIPT.

T&E Section Content:

Test Results Summary.

Opening Paragraph

The opening paragraph of the T&E section summarizes the test results from all completed test objectives and includes the overall test objective rating, supporting evidence, and an overall recommendation (Example 2). Although it appears at the beginning of the T&E section and in the Executive Summary, this paragraph is typically written after the rest of the T&E section has been written.

- **Overall Test Objective Rating** – The first sentence presents the overall results of the test and specifically identifies whether the SUT, in total, met or failed to meet the capability requirements.
- **Supporting Evidence** – The overall test objective rating often involves more than a simple roll-up of subordinate test results; multiple factors and their relative significance should be considered:
  - Test results of GTOs and STOs
    - Specification Compliance
    - Regression
  - Military utility
  - Deficiencies
  - Additional Findings

The author, technical experts, RIPT co-leads, and aircrew should discuss the significance of these factors and use them to determine the overall rating. This is true particularly when test results are mixed. For example, multiple STOs with multiple deficiencies do not automatically equate to the entire SUT being considered unsatisfactory or as having poor military utility. Likewise, one seemingly minor problem may degrade or fully prevent accomplishing mission objectives.

The supporting evidence should be presented in the order of significance, whether positive, mixed, or negative. In cases of mixed results, positive and negative test results should be presented in a balanced way, as shown in paragraph 2 of Example 2 below.

- **Overall Recommendation** – Provide the overall recommendation on the next programmatic step. The recommendation is based on the overall rating and supporting evidence. Common next programmatic steps could include "Proceed to operational test," "Field the SUT," "Continue to develop the SUT for military use," or "Continue to the next phase of DT&E."

EXAMPLE 2

Overall, integration of TF-33 engine control software v3.2 on the B-52H aircraft was satisfactory. Engine response and operability during ground starts, throttle transients, airstarts, and dynamic inflight maneuvers were satisfactory. Post-flight engine data download times were similar to legacy engine control software versions and were therefore satisfactory. **Field engine control software v3.2 on the TF-33 engine for the B-52 aircraft. (R97)**

Overall, integration of TF-33 engine control software on the B-52H aircraft was marginal. Although engine response and operability during ground starts, throttle transients, airstarts, and dynamic inflight maneuvers were satisfactory, post-flight engine data downloads took 250 percent longer to complete than previous engine control software versions, severely impacting post-flight maintenance activities and the ability to support multiple flights in a single day. **Shorten the post-flight engine data download timeline prior to fielding engine control software v3.2 on the TF-33 engine for the B-52 aircraft. (R98)**

Overall, integration of TF-33 engine control software v3.2 on the B-52H aircraft was unsatisfactory. Although engine response and operability during ground starts and airstarts were satisfactory, poor engine operability during inflight throttle transients resulted in several engine stalls, which was unsatisfactory. Additionally, post-flight engine data downloads took 250 percent longer to complete than previous engine control software versions, severely impacting post-flight maintenance activities and the ability to support multiple flights in a single day. **Improve engine operability during inflight throttle transients, shorten the post-flight engine data download timeline, and evaluate any modifications incorporated as a result of the corrective actions. (R99)**

Test Results Summary Table

The test results summary table, also known as the stoplight chart, provides a quick summary of the top-level test results, the completion status of the test objectives and/or MOPs, and a roadmap of the detailed test result discussions that follow in the Test and Evaluation section. Test results should be summarized as descriptors and ratings in accordance with the rating scale(s) that were used (usually the 412 TW Rating Scale, the 412 TW Regression Scale, and the 412 TW Test Completion Scale, shown in the Appendices section).

The level of detail presented in the Test Results Summary Table (specifically the inclusion of MOP-level descriptors or statuses) is dependent on the scope of the test program. Large, complex, or

multi-discipline test reports may only include general or specific test objective-level ratings or statuses in the Test Results Summary Table, as shown in Table 94 of Example 3. Smaller or single-discipline test programs may add specific test objective and MOP-level descriptors and statuses, as shown in Table 95 of Example 3. This flexibility allows authors to balance the level of detailed information presented with the size of the test report so as not to overwhelm the reader. When the Test Results Summary Table at the beginning of the Test and Evaluation section is higher level, a detailed Test Results Summary Table may be presented at the end of the test report body or within an appendix.

EXAMPLE 3

Table 94 Test Results Summary

Test Objective	Rating
Overall Test Objective: Test A/A Weapon Functionality	**Marginal**
STO 1.1: Evaluate AIM-120 Functionality	Marginal
STO 1.2: Demonstrate A/A Gunnery Functionality	Satisfactory
STO 1.3: Demonstrate Weapons Regression	Unchanged
STO 1.4: Collect Data Pump Messages for Contractor	Met
STO 1.5: Characterize A/A Weapon Simulator Flight Suit	Partially Met

Table 95 Test Results Summary

Test Objective/MOP	Descriptor	Rating
Overall Test Objective: Test A/A Weapon Functionality		**Marginal**
STO 1.1: Evaluate AIM-120 Functionality		**Marginal**
Controls and Displays Functionality	Borderline	
WEZ Modeling	Adequate	
STO 1.2: Demonstrate A/A Gunnery Functionality		**Satisfactory**
A/A PVI	Good	
STO 1.3: Demonstrate Weapons Regression		**Unchanged**
Legacy AIM-120 Displays Functionality	Unchanged	
AIM-9X Dynamic Cueing	Degraded	
Launch Cycle Timing	Improved	
STO 1.4: Collect Data Pump Messages for Contractor		**Met**
Data Pump Messages	Complete	
Firehose Data	Partially Complete	
STO 1.5: Characterize A/A Weapon Simulator Flight Suite		**Partially Met**
Suite K-Byte Pressure	Complete	
Close-Level Warning System	Not Tested	

For GTOs/STOs/MOPs in which a rating or descriptor is not required (such as test objectives with the verbs “collect” or “characterize/determine”), authors should use the 412 TW Test Completion Scale to indicate the status of the test objective/MOP without a corresponding rating/descriptor. Test completion status indicates whether the data collected were sufficient to answer the intent of the test objective. The test completion criteria from the test plan typically guide engineering judgment in determining test completion statuses. Rating colors for these types of test results are not necessary.

Any test completion status lower than “Met” should also have a corresponding discussion in the Limitations and Constraints section that describes why the data collected was not sufficient to meet the test objective. A “Partially Complete” MOP does not automatically mean the test objective was “Partially Met;”

the RIPT may still be able to confidently state the test objective was met despite not having a full set of "Complete" MOPs.

Rating Scales and Summary of DRs/AF Form 847s (If Applicable)

The rating scale(s) that were used to rate the test results should be referenced here and included in an appendix (typically the 412 TW Rating Scale, 412 TW Regression Scale, and 412 TW Test Completion Scale, as shown in the APPENDICES section). The number of deficiencies observed during testing should be stated. If deficiencies were identified and resolved during the testing, include that number as well, as this can help demonstrate the value of testing. Refer to the DR summaries and AF Form 847s in an appendix and identify the procedure/program under which the DRs were written, usually T.O. 00-035D-54 (Reference 8). The report may include the T.O.'s recommendation here to correct deficiencies and refer back to it for all subsequent DRs, or it may include DR-specific recommendations in the detailed test results sections instead.

Overall Test Methods and Conditions (If Applicable).

If the test methods and conditions provided in the test plan were followed for most or all test objectives, an Overall Test Methods and Conditions section may prevent needless repetition. The RIPT should determine how much of the test methods and conditions from the test plan would be valuable in the test report. Deviations from the test methods and conditions should be detailed in the relevant specific test objective(s). If test methods and conditions were common to most or all specific test objectives, yet deviated significantly from the test plan, the test methods and conditions may be summarized briefly in this section. Although test plans may include wide data bands, the test report should include the actual test conditions relevant to test results. Detailed test methodology may be included in an appendix or data package.

General and Specific Test Objectives.

General Test Objective (If Applicable)

Every GTO from the test plan should be addressed in the TR. The GTO is presented as a header. The GTO's rating, based on a rollup of the STO ratings under the GTO, is then stated and substantiated with a summary of supporting evidence (including additional findings, if applicable) and impacts to military utility. A test results summary table may be included that presents a summary of ratings and/or statuses of the STOs within the GTO. If there were no GTOs, this section should be omitted. In the event a test objective needs to be reworded for clarity, the new wording should be discussed by the RIPT and the technical review authority (TRA) who reviewed the test plan.

Specific Test Objective

Every test objective from the test plan should be addressed in the TR, although the RIPT may deem it preferable to address only the GTOs if there are many STOs.

Test Methods and Conditions

This section contains a summary of test methods and conditions relevant to a given STO. Although test plans may include wide data bands, the test report should include the actual test conditions relevant to test results. This section should be brief and only contain as-executed conditions. Additional details can be included in an appendix. Test results are not included in this section. If the test methods and conditions for the STO did not deviate from the test plan, that should be captured either here or in the Overall Test Methods and Conditions section. If an Overall Test Methods and Conditions section is included in the report, and there are no test methods or conditions unique to the STO, this section should be omitted.

Test Results

This section contains test results, analyses, conclusions (including ratings), and recommendations. The most significant factors/MOP results and the reasons they substantiate the STO's rating, whether positive (Satisfactory), mixed (Marginal), or negative (Unsatisfactory), should be described first. Impacts to military utility and additional findings relevant to the STO may also be addressed, particularly if they are not discussed at the overall test objective or GTO levels. Depending on the number of MOPs within each STO, a STO-specific test results summary table may be included summarizing the MOP descriptors supporting the STO rating.

Data products should be used to enhance clarity and save readers' time. If numerous data products support a set of test results, a subset of data products (selected based on relevancy or representation of the test results) should be presented in this section, with the remainder presented in a Data Plots and Tables appendix. Appendix B (Multimedia Content) presents guidance on the clear and effective portrayal of information via data products.

The 412 TW uses standard MOP descriptors and test objective rating criteria, including the 412 TW Rating Scale and discipline-specific rating scales (consult the appropriate tech experts). Authors should rate the SUT performance, functionality, etc., not the adequacy of testing. For example, "Radar warning receiver performance was adequate," should be used, whereas "The radar warning receiver test results were adequate," should be avoided. When describing MOP results, the appropriate descriptors should be used, as shown in Example 4. If evaluation criteria or other MOP details in the test plan require better definition, necessary changes may require a test plan amendment or an update to the data analysis plan (DAP), and should be incorporated into the test report.

Borderline results can be particularly challenging for authors to present. In this case, test results are mixed such that there is only a minimum level of acceptable capability, or the capability can only be reached using moderate workarounds. The inherent ambiguity of the descriptor necessitates a more detailed description of the test results and often entails extensive RIPT discussion. Standard rating and descriptor terms (e.g., satisfactory or good) should be avoided in test reports that do not contain formal ratings.

EXAMPLE 4

Satisfactory: Radar warning receiver performance was satisfactory. Angle of arrival error, Correct ID percentage, and response time were good and met specification.

Marginal: Radar warning receiver performance was marginal. Although angle of arrival error and response time performance was good, the Correct ID percentage for acquisition radars was borderline, as 82 percent of the time they were correctly identified (specification was 90 percent) (DR XX9999-99-9999). Unreliable threat identification could result in deploying terminal threat countermeasures when none were required. **Identify and correct Correct ID for acquisition radars. (R3)**

Unsatisfactory: Radar warning receiver performance was unsatisfactory. Mean response time performance was deficient at more than 30 seconds for lethal surface threats, exceeding the specification requirement of 5 seconds (DR XX9999-99-9997, -9998, and -9999). Long response times would lead to delayed deployment of countermeasures and potentially the loss of the aircraft in the mission. Angle of arrival error and Correct ID percentage were good. **(See R2)**

Regression Comparisons

When reporting on regression testing, detailed regression results should include the baseline configuration used as a comparison as well as the baseline rating/descriptor for each regression STO/MOP,

as shown in Example 5. If a single baseline configuration was used for all regression test objectives, that column in the regression summary table may be omitted, and the baseline configuration used can be identified in the title of the table or as a table footnote. For a large, complex, or multi-discipline test report, MOP-level regression results may be omitted, leaving only test objective-level regression results in the regression summary table, as shown in Example 6. The 412 TW uses standard regression statuses, as shown in the 412 TW Regression Scale (see Table 3 in APPENDICES section). Use of 412 TW Rating Scale colors in the Comparison to Legacy column is at the discretion of the RIPT and should only be considered if they better convey the message of the test report. Test results should be captured only for regression statuses other than unchanged. However, when the regression status is unchanged, but the original functionality/performance is less than satisfactory, this may warrant a short description of the regression results which references the original rating.

System improvements discovered during regression testing are rare. A rating of Improved should be used in cases where planned capability changes resulted in a traceable and/or unambiguous improvement to an unchanged part of a system (e.g., an improved software for System A requires less processor time, allowing System B to function with less latency). The presence of statistically defensible improvements to performance of an unchanged part of the system should lead the RIPT to understand why/how the improvements occurred before including them in the test report. Although improved performance/functionality may be favorable for the warfighter, an unintended improvement to the system may indicate that the system designer did not fully understand the implications of the changes made.

If the RIPT lacks confidence in the test data due to setup or sample size differences, etc., or if the analysis method has changed from what was done in the past, the test data should not be compared to a previous baseline configuration. Rather, the test results section should state they could not be compared and that the test data will serve as a new baseline. The RIPT may consider analyzing the data as it had been done previously, and include the analysis for future comparison. In cases where the analysis method change happens after testing begins, a test plan amendment may be warranted.

EXAMPLE 5

Table 96 Regression Test Results Detailed to the MOP Level

Test Objective/MOP	Baseline Information			Regression Status	
	Config.	Descriptor	Rating	MOP	Test Objective
GTO 1: Demonstrate Avionics Regression			**Satisfactory**		**Impacted**
STO 1.1 – Demonstrate Weapons Regression			**Marginal**		**Unchanged**
AIM-120 Controls and Displays Functionality	30R05	Deficient		Unchanged	
Launch Cycle Timing	30R04	Deficient		Improved	
STO 1.2 – Demonstrate ALQ-99 Regression			**Satisfactory**		**Degraded**
Angle of Arrival	Block 3F	Good		Degraded	
Correct ID Percentage	Block 3F	Good		Impacted	
Geo Miss Distance	30R04	Borderline		Unchanged	
Track File Transfer	Block 4	Good		Good*	

* New baseline rating due to updated data analysis method.

EXAMPLE 6

Table 97 Regression Test Results Detailed to the STO Level

Test Objective/MOP	Baseline Rating	Regression Status
GTO 1: Demonstrate Avionics Regression	**Satisfactory**	**Impacted**
STO 1.1 – Demonstrate Weapons Regression	Marginal	Unchanged
STO 1.2 – Demonstrate ALQ-99 Regression	Satisfactory	Degraded

Note: The baseline configuration was 30R05.

Additional Findings (If Applicable).

Any significant discoveries made during testing that were outside the scope of the STOs may be included in an Additional Findings section at the end of the report's T&E section, as shown in Example 7. If the finding was related to one STO, the section may be placed at the end of that STO's Test Results section instead. Additional findings may include deficiencies, performance improvements, test or operational techniques, or other discoveries. Additional findings of significant impact to overall test results should also be referenced in the introductory overall test results paragraph(s) of the T&E section.

EXAMPLE 7

Additional Findings.

The unmodified aircraft required selecting a jettison button to jettison the store. The test configuration required the pilot to activate three switches on the bomb rack unit control panel, in addition to the jettison button (DR XX9999-99-1234). The added three switch actuations would impose additional steps to an emergency jettison procedure that typically would be time critical. **(See R2)**

Though not reported in the Buildup Loading Test Results section, an additional buildup loading sortie was flown. The results of the departure resistance maneuvers flown with loading 31N showed significantly less departure resistance than was expected based on the legacy results. This degradation led to a concern that the angle-of-attack (AOA) cones may have been misaligned relative to the aircraft pitch reference. Subsequently, the AOA cone alignment was checked, and one cone was found to be loose. Both AOA cones were verified to be out of tolerance, and were realigned per T.O. guidance. As a result of the AOA cones being out of tolerance, the results from this buildup loading were not considered to be a valid representation of the effect on aircraft stability by the new store. All future projects should verify correct AOA cone alignment when preparing for high-AOA testing. **(See R2)**

Recommendations (Optional):

A standalone Recommendations section may be useful for reports with six or more individual (vs. "See R2") recommendations. The list is ordered numerically in order of first appearance, not by priority. The recommendations' text should be copy/pasted or cross-referenced from the body of the report and include the page number(s) on which they appear.

References (If Applicable):

References provide the information necessary for a reader to locate and retrieve any sources cited in the test report, such as document number, title, author(s), and publication information. If there are six or more different references in a document, a References section is required. For test reports with five or fewer different references, the reference listings should appear in footnotes. The TR template contains a comprehensive list of reference type examples and provides formatting guidance.

TEST REPORT APPENDICES

Appendices (aka Attachments in some types of test reports) are for supplemental information too detailed/cumbersome to be included in the body of the report. Each appendix should be specific to one type of information. Table 1 includes appendices that are often included in TRs.

Table 1 Common TR Appendices

Appendix	Inclusion
Rating Scales	Required if Ratings are Used
Detailed Test Item Description	Required if in Source Test Plan
DRs and/or Publication Change Requests	Required if DRs Referenced in Report
Detailed Test Methodology/Procedures	Optional
Data Plots and Tables	Optional
Analysis Techniques	Optional
Abbreviations, Acronyms, and Symbols	Required in TRs
Distribution List	Required

Other than the Distribution List and Abbreviations, Acronyms, and Symbols appendices (which are always, respectively, the last and second-to-last appendices in TRs), appendix order should be determined by the RIPT. Refer to the following sections for descriptions of these common appendices.

RATING SCALES

This appendix contains all the rating scales used in the test report, including the 412 TW rating scales and any discipline-specific rating scales. If no ratings are used, a rating criteria appendix is not required.

412 TW Rating Scales:

In general, the 412 TW Rating Scale (Table 2), 412 TW Regression Scale (Table 3), and/or the 412 TW Test Completion Scale (Table 4) are used to assess the overall SUT.

Table 2 412 TW Rating Scale

How Well Does the System Meet Mission and/or Task Requirements?	Changes Recommended for Mission/Task Improvement	MOP Descriptor	Test Objective Rating
Exceeds requirements.	None.	Excellent	Satisfactory
Meets all or a majority of the requirements.	Negligible changes needed to enhance or improve operational test or field use.	Good	Satisfactory
Some requirements met; can do the job, but not as well as it could or should.	Minor changes needed to improve operational test or field use.	Adequate	
Minimum level of acceptable capability and/or some non-critical requirements not met.	Moderate changes needed to reduce risk in operational test or field use.	Borderline	Marginal
One or some of the critical functional requirements were not met.	Substantial changes needed to achieve satisfactory functionality.	Deficient	Unsatisfactory
A majority or all of the functional requirements were not met.	Major changes required to achieve system functionality.	Unusable	
Mission not safe.	Critical changes mandatory.	Unsafe	Unsatisfactory

Table 3 412 TW Regression Scale

How Does System Performance/Functionality Compare with Previous Test Results and Was Overall Capability Affected?	Regression Status
System performance/functionality improved, and overall capability was unaffected or improved.	Improved
No change to system performance/functionality, and overall capability was unaffected.	Unchanged
Minor changes to system performance/functionality, but overall capability was unaffected.	Impacted
System performance/functionality was degraded, and overall capability was affected.	Degraded

Note: The colors in the regression scale are optional and may be omitted in cases where they do not add value.

Table 4 412 TW Test Completion Scale

Were Collected Data Sufficient to Meet the Test Objective?	MOP Status	Test Objective Status
Sufficient data collected; for characterization/determination efforts, collected data were sufficient to establish a performance baseline or attribute of the system.	Complete	Met
Limited data collected (criticality of data not collected is relatively low); for characterization/determination efforts, collected data were sufficient to establish a limited performance baseline or attribute of the system.	Partially Complete	Partially Met
Insufficient data collected (criticality of data not collected is relatively high); for characterization/determination efforts, collected data were not sufficient to establish a performance baseline or attribute of the system.	Insufficient	Not Met
No data collected; for characterization/determination efforts, establishment of a performance baseline or system attribute were not attempted.	Not Tested	

A detailed RGB/Hexadecimal listing of the standard rating colors is provided in Table 5.

Table 5 RGB and Hexadecimal Codes for Standard Rating Colors

Color	Red Value	Green Value	Blue Value	Hexadecimal Code
Blue	46	116	181	2E74B5
Green	30	175	70	1EAF46
Yellow	255	255	0	FFFF00
Red	255	0	0	FF0000

Note: Teams may deviate from these nominal values for purposes of readability.

Additional Rating Scales.

Various organizations across the 412 TW may use discipline- or test-specific rating scales. Contact the appropriate discipline technical expert for guidance on the use of these scales.

DETAILED TEST ITEM DESCRIPTION

The test report Detailed TID appendix is derived almost verbatim from the test plan Detailed TID appendix. If the test report body TID is further abbreviated (as compared to in the source test plan), then the rest of the details should be included in this appendix. If the SUT changes during test, this section should be updated to reflect the final iteration. If there was no detailed test item description in the source test plan, then this appendix may be omitted.

DRS AND/OR PUBLICATION CHANGE REQUESTS

This appendix contains a table listing all DRs and/or publication change requests referenced in the report, followed by summaries (DRs) and/or AF Form 847s (publication change requests). These summaries

and forms should be copied from the deficiency database, but grammatical or wording errors should be corrected in this appendix for the sake of clarity. If no DRs or publication change requests are referenced in the report, this appendix should be omitted.

DETAILED TEST METHODOLOGY/PROCEDURES

This optional appendix is used to document the test methodology, procedures, or flight test techniques that were ultimately used during the test. These may be different from those in the test plan, especially if multiple test package amendments occurred, or if the description is more detailed than that provided in the test plan. This appendix can be useful for future test teams.

DATA PLOTS AND TABLES

This optional appendix contains plots and tables that support the ratings, but are too cumbersome or detailed to include in the main body. For large amounts of tables and plots, that content should instead be included in a data package.

ANALYSIS TECHNIQUES

This optional appendix features analysis techniques that are typically included in a DAP, and which may be valuable to stakeholders.

STATISTICAL ANALYSES

This optional appendix contains statistical analyses used in the test report to ensure that test methodology and results are reproducible and defensible. This section should include the process(es) the data underwent to get the correct subset of information used for the analyses; the statistical methodology used to conduct the analyses; and the results of the analyses, including any plots, graphics, or tables of results. Results should be accompanied by an explanation of what is included in the tables or graphics and their interpretation.

ABBREVIATIONS, ACRONYMS, AND SYMBOLS

Generally, the technical editor will compile and update this appendix (required except for test reports of 24 or fewer pages cover-to-cover without this section) and will ensure that all abbreviations, acronyms, and symbols appearing in text (outside of figures and tables) are defined on first use. This appendix will include all abbreviations, acronyms, and symbols in figures, tables, and text. This appendix is always the next-to-last section of the test report, and should be referred to in a footnote to the first table or figure title in the main body (whichever appears first).

DISTRIBUTION LIST

Generally, the technical editor will compile and update this required appendix in coordination with the RIPT. The distribution list is always the last section of the test report and contains the list of recipients of the final test report approved by the 412 TW and PO; the list is kept updated in the current test report template. Changes to the standard list should be authorized by the CTF's commander or chief engineer.

AUTHOR'S GUIDE REFERENCES

1. AFFTC-TIH-14-01, *The Author's Guide to Writing 412th Test Wing Technical Reports*, 412th Test Wing, Edwards AFB, California, December 2014.

2. EdwardsAFBI 99-103, *412 TW Technical Report Program,* 412th Test Wing, Edwards AFB, California, August 2013. [Certified current August 2021.]

3. DoDM 5200.01, Volume 2, DoD Information Security Program: Marking of Information, DoD, Washington, D.C., February 2012, last updated July 2020.

4. DoDI 5230.24, *Distribution Statements on Technical Documents*, DoD, Washington, D.C., August 2012.

5. AFI 61-201, *Management of Scientific and Technical Information (STINFO),*USAF, Washington, D.C., November 2020.

6. DoDM 3200.14, Volume 1, *Principles and Operational Parameters of the DoD Scientific and Technical Information Program (STIP): General Processes*, DoD, Washington D.C., March 2014.

7. DoDI 5000.89_DAFI99-103, *Capabilities-Based Test and Evaluation*, USAF, Washington, D.C., December 2021. [Corrective Actions applied March 2022.]

8. T.O. 00-35D-54, *USAF Deficiency Reporting, Investigation, and Resolution (DRI&R)*, 558 CBSS/GBHA, Tinker AFB, Oklahoma, April 2021.

9. DAFMAN 63-119, *Mission-Oriented Test Readiness Certification,* USAF, Washington D.C., April 2021.

10. TE 02 EN, *Technical Reporting Checklist,* downloadable Excel file here, last accessed November 2022.

APPENDIX A – REPORT TYPES

This section contains summaries of the report types for the 412 TW that are outlined in EdwardsAFBI 99-103 (Reference 2). The first three are the primary report types:

Capability Report (CR)	Approval Authority: TW Technical Director	DTIC: Yes
Details: Provides the comprehensive, top-level weapons system capability DT&E results to support timely programmatic decisions, with an emphasis on military utility. May be standalone final documentation or followed by one or more TRs, and may be accompanied by one or more data packages. Typical main body length: 25 to 30 pages. Appropriate for: • A program that does not need any detailed analysis, just a rapid assessment of military utility. • A program primarily focused on operational capability of a system-level test. Not Appropriate for: • A results document that is desired on a faster timeline than a TR (the review and approval process is the same). Calling a full TR a CR will not shorten the publishing timeline.		

Preliminary Report of Results (PRR)	Approval Authority: TW Technical Director	DTIC: Yes
Details: The PRR presents early test results and may provide recommendations (e.g., to address deficiencies) for testing that has been completed thus far. Should be followed by one or more TRs. Typical main body length: 5 to 10 pages. Appropriate for: • A customer anticipating a programmatic decision prior to the planned completion of flight test and during the planning process requests preliminary results to support that decision. Should be followed by a full TR. • A significant unanticipated hurdle (such as an aircraft component failure that slips the schedule by months) that prevents a TR from being released at the planned date. If a customer needs preliminary results to support an immediate decision a PRR may be appropriate. Should be followed by a full TR. • The customer who has a program decision deadline that cannot be delayed, and T&E may not be complete in time to provide a TR to meet that deadline. Should be followed by a full TR. Not Appropriate for: • Pausing work on a TR in order to complete a PRR if the TR could be completed a short time later. • Replacing a TR, which provides a complete and more-detailed documentation of T&E than a PRR.		

Technical Report (TR)	Approval Authority: TW Technical Director	DTIC: Yes
Details: Final report of a test and evaluation of results. As described in this document, the TR contains full descriptions of testing, test item, background, test results, analyses, conclusions (including ratings), and recommendations. Alternate formatting styles of TRs such as short, ultrashort, and MFR- and dashboard-style reports may be used; all are considered TRs. Typical main body length: 25 to 30 pages. Appropriate for: • Most circumstances. The TR is the most common 412 TW deliverable. • After a flight test program when the customer needs detailed results and/or recommendations. • Following a PRR that was previously delivered now that flight test is complete. Not Appropriate for: • A data collection-only test program, in which the 412 TW has no analysis requirements.		

The following report types are manuals/handbooks or other occasionally used variants of the main three types of reports:

Progress Report (PR)	Approval Authority: Squadron Commander or Equivalent*	DTIC: No
Details: A PR may be as simple as a periodic update of test points or initial results. It may be made in a memorandum, slide, or email format, and as such usually contains only preliminary information that is of use in program management and planning. The format is agreed to between the CTF chief engineer and the customer. Test Complete Letters (TCLs), Safety-of-Flight Letters (SOFLs), and quick-look reports fall in the PR category and are intended to assist in meeting the customers' needs in a timely and consistent manner. Typical main body length: 1 to 5 pages. Appropriate for: • A customer that has requested updates during flight test to track test point completion and other metrics communicating test program status. Not Appropriate for: • Conveying ratings/conclusions or recommendations. *Consult EdwardsAFBI 99-103 (Reference 2) for the full list.		

PR: Test Complete Letter (TCL)	Approval Authority: Squadron Commander or Equivalent	DTIC: No
Details: The TCL is a type of PR that documents completion of testing and is used when the primary output of the test is data to be used by an outside agency. When no formal conclusions or recommendations are required, a TCL informs the customer that testing is complete. Raw data and/or data packages may precede/follow the TCL. Typical main body length: 1 to 2 pages. Appropriate for: • Data collection efforts. Not appropriate for: • When a formal report is planned. • When a customer is seeking an evaluation or interpretation of the test results.		

PR: Safety of Flight Letter (SOFL)	Approval Authority: Squadron Commander or Equivalent	DTIC: No
Details: The SOFL is a type of PR often used to transmit information to a customer and other flight organizations, such as OT&E, indicating initial functionality or safety checks have been made on a new aircraft system or software, and has been cleared for specified purposes. Typical main body length: 1 to 2 pages.		

PR: Quick Look, Formal/Informal	Approval Authority: Squadron Commander or Equivalent*	DTIC: No
Details: The quick-look is a type of PR that details the initial engineering assessments of a SUT, and should be shared with any stakeholder benefiting from that information. Less formal quick-look assessments may be written immediately following a mission by the discipline or operations engineer to document initial review of test data. The formatting and structure are typically CTF-dependent. Typical main body length: 3 to 15 pages. Appropriate for: • Summarizing flight test events, including participating aircraft, run times, etc. Not Appropriate for: • Supporting acquisition decisions. *If sent outside the RIPT, this report should be approved by squadron leadership, but need not be approved if retained as an internal record of day–to–day test execution.		

Technical Information Handbook (TIH)	Approval Authority: TW Technical Director	DTIC: Yes

Details: A technical document, such as this test report author's guide, whose primary focus is to guide readers through a process. Formatting may be less formal than in reports of results meant to support programmatic decisions.

Appropriate for:
• Documenting processes or other technical information for 412 TW testers or other customers.

Not Appropriate for:
• Conveying test results.

Technical Information Memorandum (TIM)	Approval Authority: TW Technical Director	DTIC: Yes

Details: This report is often used to document and retain technical knowledge for engineering reference. An example might be a description of how a government-created flight test model was developed and validated to generate predictions for future test programs. The TIM may include lessons learned during the model creation process.

Appropriate for:
• Documenting technical results in a detailed form, primarily to benefit future testing.

Not Appropriate for:
• Conveying formal ratings and recommendations that support acquisition decisions.
• Circumventing the publishing timelines associated with TRs.

Data Package	Approval Authority: Squadron Chief Engineer and/or Technical Expert	DTIC: Yes

Details: The data package presents and explains detailed collected data and test results. It may contain photographs, video, .csv, .txt, or other data files, algorithms and data-processing technique details, graphs and plots, statistical analyses and any other technical detail required to understand the full data collection and analysis process. Format as agreed upon by the RIPT.

Appropriate for:
• Providing supplemental data that support results in a TR or PRR.
• Characterization-only test programs, for which no ratings or recommendations are required.

Not appropriate for:
• Circumventing the publishing timelines and detail associated with TRs.
• Supplying required ratings or recommendations.

TR Addendum	Approval Authority: TW Technical Director	DTIC: Yes

Details: Occasionally, a portion of testing is delayed to the point that the RIPT does not wish to postpone publication of the already-completed test results. In this case, the completed test results will be published in a TR, and the delayed test results may be published later in the form of an addendum to this same TR. For example, range or other assets' availability may delay completion of test points. The addendum should address whether and how its test results affect the overall rating and recommendation reported in the TR. The addendum uses TR formatting and has the same final approval authority.

Appropriate for
• Reporting results of additional testing completed after a TR is released.

Not appropriate for:
• Testing is complete, but analysis is not.

TR Amendment	Approval Authority: TW Technical Director	DTIC: Yes
Details: After the TR (and related addendum if applicable) has been published, the RIPT may discover new information that drives changes to ratings/conclusions and/or recommendations.* This new information should be published in an amendment to the TR. Appropriate for: • Modifying ratings/conclusions and/or recommendations based on new information discovered after the TR has been published. Not Appropriate for: • Correcting typographical errors or minor discrepancies. If an error is considered serious enough to require correction, but the ratings/conclusions and/or recommendations in the TR have not changed, consult your technical editor about issuing an errata letter. * The RIPT should discuss whether an addendum or an amendment is appropriate if there is any confusion.		

Recommendation for Dedicated OT&E	Approval Authority: TW Technical Director	DTIC: No
This memorandum documents recommendations for OT&E testing, IAW DAFMAN 63-119, *Mission-Oriented Test Readiness Certification* (Reference 9).		

APPENDIX B – MULTIMEDIA CONTENT

PRESENTATION OF DATA PRODUCTS

Data products in a test report help build a strong, logical foundation for conclusions and recommendations, and they take many forms, including tables, figures (plots or images), and video. Data products should be discussed immediately preceding their presentation, or at the top of the next page when spacing dictates. Because data products are frequently copied from test reports into presentations without the accompanying paragraphs, their meaning should be checked for ambiguity and possible misinterpretation or misuse. The following guidance should help ensure data products are clear and effective.

All Data Products:

1. Choose the most appropriate data product to support the conclusions; consult with the RIPT members as needed.
2. Be selective about when to employ data products in the main body (e.g., new candidates, major issues, etc.); higher-level summary data products, or a single detailed product that is representative of test results are preferred.
3. Employ commonly used software formats (e.g., Adobe, MS Office). Proprietary formats should be converted to standard when possible.
4. Use disciplines' commonly used data products instead of customized products conveying the same information, as familiarity increases reading speed and reduces confusion.
5. Cross-check plotted data points with tabular data (if both are shown).
6. When error bounds for data are known, include them. When they are not known, or are estimated, error bounds are implied by significant digits.
7. Do not use more significant digits than are justified by the resolution and accuracy of the instrumentation system.
8. Include units on all dimensional values and axes.
9. Label axes/rows/columns with easily understood titles (e.g., "Altitude [ft PA]" instead of "AL1234 [ft PA]").
10. Label manipulated raw data (e.g., filtered, down-sampled, brightness, contrast, or color adjusted).
11. Label the starting/entry point "Start" in flow diagrams or decision trees.
12. Explain all plot symbols, line styles, colors, etc. Typically, a legend is used, but callouts such as pointer arrows with floating text boxes or other options are acceptable; these explanations should be on the plot, not in the surrounding paragraphs.
13. Use actual test conditions, not planned test conditions.
14. Explain obvious outliers.
15. State or reference the sign convention (could be in text, legend, axis label, etc.) when presenting axis-system-based values (e.g., positive z-axis down).
16. Make titles for figures, tables, and other media unique and include the most pertinent identifying information, such as test conditions, flight number, or other relevant information.
17. Move text from the column header to the table's title or in-table subtitle in cases where a table with a column that has the same entry for every row (e.g., altitude is 35,000 ft PA for all rows). These constitute metadata, such as:
    a. Aircraft type/tail
    b. Data basis = flight test, ground test, or prediction
    c. Date(s) of data collection
    d. Atmospheric conditions (if relevant)
    e. Aircraft external configuration (stores, flaps/gear position, door position, etc.)
    f. Corrections applied to data (e.g., corrected to Standard Day)

g. Software/Operational Flight Program version(s)
h. Sample rate and cut-off frequency of anti-aliasing filters
i. Post-processing (e.g., data/image/video filters)

18. Use color deliberately to add value, such that it does not distract the reader. Red/magenta should be used with care, as those colors typically denote limits, unsatisfactory, or dangerous conditions.
    a. Choose high-contrast colors (i.e., avoid yellow on a white background)
    b. Some shades of blues and green may be difficult to distinguish
    c. When possible, use a color-blind-friendly palette (search online for examples) or use a variety of shapes to help readers easily distinguish data
19. Construct data products for the type of media on which they will be viewed - specifically regarding size, readability, and use of color.
20. Check both digital and print output for data readability. A resolution of 100 to 300 ppi is recommended.

Tables:

1. Fill in all spaces/cells of a table; use not applicable (N/A) or three hyphens (---) if necessary.
2. Use standard superscripted symbols for footnotes (e.g., *, †, ‡, §). Keep like superscripts consistent between tables (e.g., all inert stores use **).
3. Tables in the main body of the report should generally not exceed one page.
4. If a table spans multiple pages, ensure it remains readable; headers should be repeated as necessary to avoid confusion. The technical editor can help with formatting.

Plots:

1. Do not run data beyond axis lines or extrapolations past data (unless the RIPT agrees that extrapolation is appropriate in a given instance).
2. Show curve-fits of measured data as solid lines; show predictions (including extrapolations) and confidence intervals as dashed lines. Label data fairings to distinguish whether they are "curve fit" (such as hand-drawn line or least squares regression) or were generated by a simulation or model prediction.
3. Select automatically generated curve fits based on theoretical physics of the system's behavior, not just what looks best (e.g., natural logarithmic behavior instead of a 12th order polynomial).
4. Depict thresholds and limits unambiguously (bold green or red lines are excellent choices).
5. If there are notable events in a plot, add arrows or other cues to direct the reader's eyes. With multiple similar plots, highlight or circle important differences.
6. Ensure scales should contain intuitive numbering (e.g., multiples of 1, 2, or 5). Every plot in a series or family of plots should use the same scales. Avoid auto-generated axis scales (e.g., 0, 3.7, 7.4, 11.1)
7. When selecting colors, line patterns, and symbols, consider contrast with the background and differentiation between various data sets, particularly where data nearly or completely overlap.
8. Ensure data plot lines are easy to follow; too many on one chart results in a visual mess, especially if the lines overlap or cross.
9. Select scales to ensure the important details of the trace are discernable; different scales (or ranges) may present all or emphasize only part of the data (e.g., if some large amplitude points drive the scale such that the reader cannot see the relevant details, the data range should be zoomed in to reveal).
10. Use standard symbols (e.g., ■, ⊡, ◈, △, ▲). Keep like data symbols consistent between plots (e.g., all 10,000-foot data use ⊡).
11. Only connect data points with lines if it's appropriate to interpolate results between points. Otherwise, the plot can imply/convey data where none exist.

12. When plotting the results of a rating scale, ensure the scale of the chart matches the range of the rating scale (e.g., the 1-to-10 Cooper-Harper charts should not start at 0 nor go to 11).

Interactive Dashboards:

Interactive dashboards are among multimedia tools that are becoming available to 412 TW authors. These digital formats may present data from different perspectives quickly and in more detail than static displays are capable of providing.

1. Ensure dashboards are reader-friendly (intuitive, consistent, and easy to navigate).
2. Include browser and/or application information (version, iteration, etc.).
3. Solicit early feedback on new design ideas to avoid wasted time and effort.
4. Inclusion of links to databases and/or digital models should be coordinated with the RIPT and the customer to ensure the data are usable.

Images, Animations, and Video:

1. If there are notable features or events, zoom in and/or use callouts (e.g., labels, arrows, etc.) or other cues to direct the attention to significant content.
2. Ensure that key elements remain clear. Consider adjusting the brightness, contrast, and resolution as required, but annotate the video as such.
3. Most CTFs have individuals skilled in vector graphics and photo/video editing that can help prepare high-quality media.
4. Cut the animation/video to an appropriate length: leave a sufficient introduction so the viewers can orient themselves (usually 3 to 5 seconds), and then the relevant event(s). Try to avoid “dead space” or long periods of no movement.
5. For animations/videos in technical reports, avoid superfluous or flashy effects, including music or other artificial audio/video effects. Remarks on captured audio that would reflect poorly on the professionalism of the 412 TW should be considered for removal.

APPENDIX C – REPORTING PROCESS

REPORTING PROCESS AND NOMINAL SCHEDULE

Prior to Test Execution (Following Test Package Approval):

The primary author should request tech editing begin converting the relevant test plan into a customized report template (aka a skeleton or shell). The TR number will be assigned in coordination with the tech editor.

During Test Execution:

Data should be reduced and analyzed during execution, including writing WITs and DRs throughout the test period. Authors should consult their lead engineers and technical experts (outside of RIPT meetings) to ensure data analysis quality and timeliness. It may be beneficial to start an early draft of the T&E section assuming successful data collection and satisfactory test results

Last Day of Test and Last Day Data Received:

The last day of test is the date of the final ground or flight test execution event that falls under the scope of the test report or addendum. The last day test data are received is the date post-test processing by inside or outside organizations is delivered, and marks the starting point of the CTF/customer-coordinated schedule (nominally 30 days, as shown in EdwardsAFBI 99-103 [Reference 2] and Figure C1).

Report Integrated Product (RIPT) Meetings:

The RIPT meetings are held to manage report quality, content, and publishing timeline. These meetings should be held regularly, and as driven by events. The RIPT consists of everyone who will be involved in the reporting process. The primary members of the RIPT include: authors, co-leads (typically the chief engineer and the home office representative), technical experts, project pilots, and technical editor (see EdwardsAFBI 99-103 [Reference 2] for a full list). The first and final RIPT meetings are unique, as detailed in the following sections.

First RIPT Meeting (aka Kickoff Meeting).

This is the initial RIPT meeting, whose purpose is to establish the overall plan for test reporting. This meeting should be held prior to or soon after the start of test execution and address: report content, classification, releasability, and structure; reporting timeline; RIPT roles, responsibilities, and availability; and resource management (for a full checklist, see TE 02 EN, *Technical Reporting Checklist* [Reference 10]).

Final Squadron-Level Coordination Meeting (aka Murder Board).

This is the final RIPT meeting, whose purpose is to ensure all report contributors and squadron-level (both CTF and home office) approvers reach agreement on the report content. At the end of this meeting, all document changes should be completed or identified and any action items assigned so that the report can move forward for final approval. Once all action items have been completed, Tech Publications' final preparation for TW approval begins, including scheduling, readahead packages, and meeting invitations.

Report Signing Brief Checklist

The Report Signing Brief Checklist is a summary document covering the purpose and outcomes of the test including programmatic issues not directly addressed in the TR. The checklist presents concerns or points of contention that may arise to Test Wing leadership, including the 412 Test Wing commander. It is reviewed by the RIPT at the final squadron-level coordination meeting and the Test Wing at the report approval meeting, and finally submitted with the final version of the Executive Summary to the Test Wing commander. The technical editor can supply the current checklist template.

Report Approval Meeting (aka Signing Party).

At this meeting the TW leadership reviews and approves publication of the report. As such, it is highly desired for invitees to attend, particularly those individuals who made a significant technical contribution to the report. The TW leadership or other attendees may request changes be made to the report prior to final approval. After final edits are made and the report has been approved and signed, the final report is distributed to all organizations listed on the report's distribution list. The technical editor coordinates final distribution with the author and chief engineer.

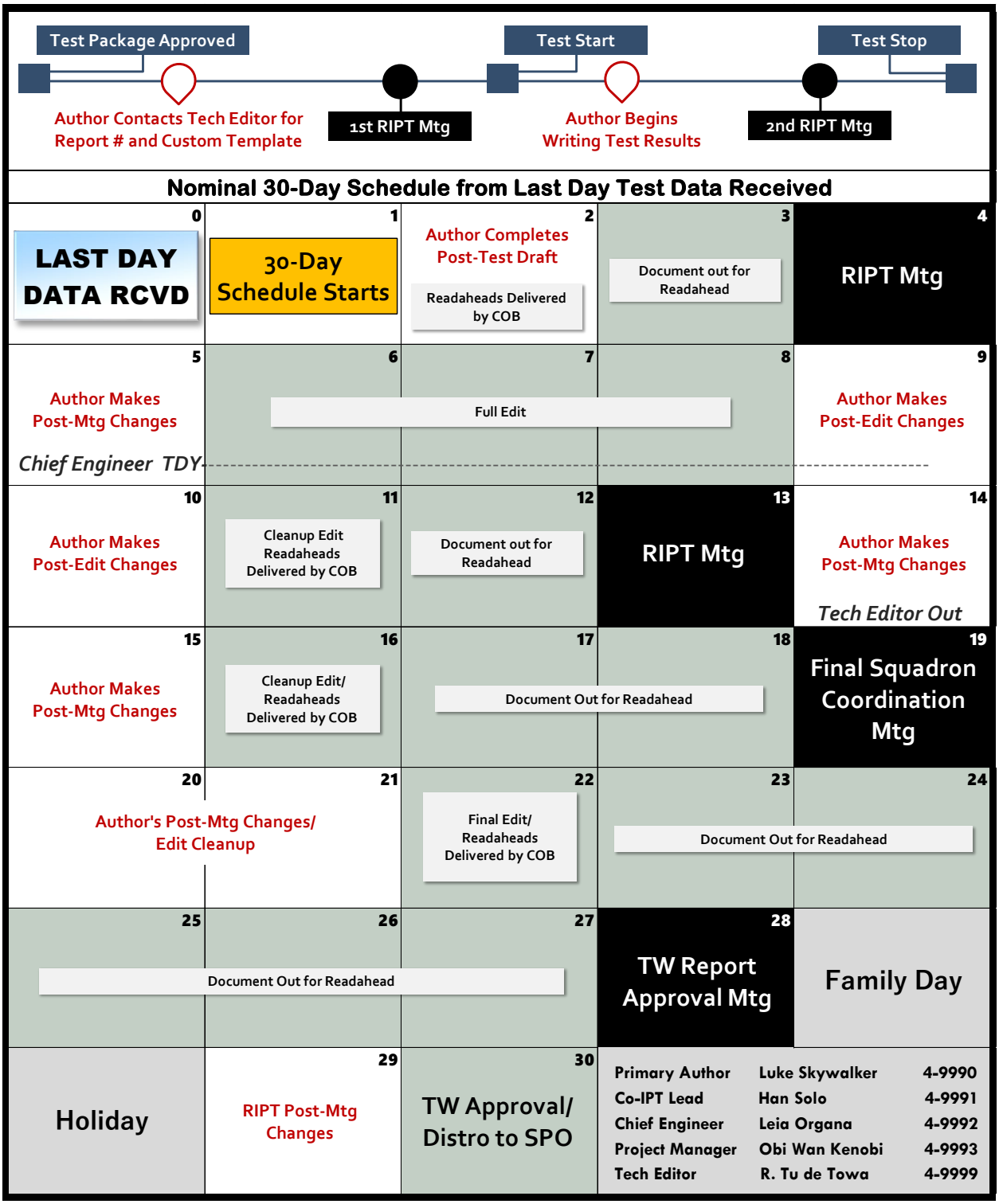

Figure C1 Nominal 30-Day Reporting Schedule

APPENDIX D – REPORTING LESSONS LEARNED

CONFIGURATION CONTROL

Configuration control, also known as version control, is the proper management of changes to documents. Good communication is key; document ownership throughout the project should be clear to all. Co-authors should work with the project's lead technical editor to define a configuration control plan to ensure all changes are incorporated into the document as efficiently as possible. These tips may help improve configuration control:

- Include a revision number or letter in the document file's title, and incrementing the revision as the document changes ownership.
- Avoid the use of relative terms such as "Current," "New," "Draft," or "Final" in document titles, as multiple revisions can make these titles misleading.
- Remind multiple simultaneous reviewers to track their changes for incorporation into the live document later.
- Coordinate simultaneous work on different parts of the report (e.g., the primary author continues writing the appendices while the technical editor works on the main section of the document).
- Maximize use of methods and tools that enable simultaneous content editing such as MS Teams/SharePoint.
- Use version control software such as Git to manage parallel writing, editing, and automatically merging updates.

CONTENT FORMATTING

The authors' focus should be on ensuring the report is technically correct. Technical editors ensure formatting (e.g., font, margins, table borders, and figures) is in accordance with 412 TW guidelines. To simplify the process of report writing and provide for uniform formatting for 412 TW documents, current and customized templates are provided by technical editors.

GRAMMAR/WORDING

The 412 TW test reports are often written as a group effort, with significant oversight from technical experts and RIPT co-chairs, and requiring senior leadership approval. The author should be prepared for wording and grammar changes as Test Wing policy and preferences evolve. Authors should seek out input and insights from the appropriate reviewers early in the process. Additionally, reports with multiple co-authors will be edited for cohesion; specific wording, phraseology, or structure may be modified for the sake of consistency. Spending time/money on trivial changes that do not add value to the customer should be avoided.

APPENDIX E – ABBREVIATIONS, ACRONYMS, AND SYMBOLS

Abbreviation	Definition
412 TW	412th Test Wing
A/A	angle of attack
AFB	Air Force Base
AFI	Air Force Instruction
AFTC	Air Force Test Center
AFTCI	Air Force Test Center Instruction
CR	capability report
CTF	Combined Test Force
CUI	Controlled Unclassified Information
DAP	data analysis plan
DoD	Department of Defense
DoDD	Department of Defense Directive
DoDI	Department of Defense Instruction
DoDM	Department of Defense Manual
DP	data package
DR	deficiency report
DT&E	developmental test and evaluation
DTIC	Defense Technical Information Center
EAFB	Edwards Air Force Base
EdwardsAFBI	Edwards Air Force Base Instruction
EWG	Electronic Warfare Group
ft	feet
GTO	general test objective
IAW	in accordance with
ID	identification
JSE	Joint Simulation Environment
K-BYTE	kilobyte
MOP	measure of performance
MS	Microsoft
N/A	not applicable
O.I.	Operating Instruction
OT&E	operational test and evaluation
PA	pressure altitude
PO	program office
ppi	pixels per inch
PRR	preliminary report of results
PVI	pilot vehicle interface
RIPT	report integrated product team

Abbreviation	Definition
SOC	statement of capability
SEAD	suppression of enemy air defense
STINFO	scientific and technical information
STIP	Scientific and Technical Information Program
STO	specific test objective
SUT	system under test
T.O.	Technical Order
TCL	test complete letter
TID	test item description
TIH	technical information handbook
TIM	technical information memorandum
TP	test plan
TR	technical report
TW	Test Wing
U.S.	United States
USAF	United States Air Force
vs.	versus
WEZ	weapon employment zone
WIT	watch item

APPENDIX F – DISTRIBUTION LIST

Onsite Distribution	Number of Copies	
	Digital	Paper
Edwards AFB Technical Research Library Attn: Alison Vasquez 307 E Popson Ave Edwards AFB CA 93524-6630	0	2
AFTC/HO Email: james.tucker.35@us.af.mil	1	0
Offsite Distribution		
Defense Technical Information Center Email: jackie.l.rike.civ@mail.mil	1	0
Totals:	2	2