



Making the Most of Scientific Test and Analysis Techniques

William F. Rowell, Ph.D. ■ Seth A. Guldin, MS
Steven N. Thorsen, Ph.D. ■ Darryl K. Ahner, Ph.D., PE

DODI 5000.02 DIRECTS THE USE OF SCIENTIFIC TEST AND ANALYSIS TECHNIQUES (STAT) TO DESIGN an effective and efficient test program for both developmental and operational testing. Unfortunately compliance with this guidance by simply checking the block (“We do STAT!”) provides no assurances that your testing and evaluation (T&E) program will receive the expected benefits in the following key areas:

- **Effectiveness:** the degree to which the testing process is successful in achieving the program’s test objectives (answers the right questions). This goal is the true bottom line of the testing process and one that the structured nature of the STAT process enhances.
- **Efficiency:** the degree to which the testing process is successful in minimizing the use of resources (uses the minimum necessary resources).
- **Defensibility:** the degree to which the analytical results of the testing process can be defended (verify the right question has been answered).

Rowell is a Senior Operations Research (OR) consultant at The Perduco Group supporting the Scientific Test and Analysis Techniques (STAT) Center of Excellence (COE) as a STAT expert at the Air Force Institute of Technology (AFIT) at Wright-Patterson Air Force Base in Ohio. He received his Ph.D. in Operations Research (OR) from the University of Texas at Austin. **Guldin** is an OR analyst at MacAulay-Brown supporting the STAT COE as an assistant STAT expert. He earned his Masters of Engineering in Material Science from Virginia Polytechnic Institute and State University. **Thorsen** serves as Associate Director of STAT COE and earned his Ph.D. in Applied Mathematics from AFIT. **Ahner** is Director of STAT COE and a full Professor in the Department of Operational Sciences at AFIT. He received his Ph.D. in Systems Engineering from Boston University and holds a Professional Engineer license.



- **Resilience:** the degree to which the testing process can achieve the program's test objectives in the face of unexpected changes in testing resources, factors, and conditions (readily adapts as needed to achieve test objectives in the face of testing uncertainty).

This article provides the STAT Center of Excellence (COE) perspective on what you really need to know to integrate STAT, as a process, into your T&E program to realize these benefits.

First, we look at the fundamentals you need to know about STAT to fully understand the rest of the paper. Second, we elaborate on various ways STAT can make a difference in your program by showing how integration of STAT into the testing process enhances its effectiveness, efficiency, defensibility and resilience. Third, we cover best practices for implementing STAT in your program so you can make the most of STAT. Finally, we provide you with a few questions to help you to evaluate how STAT benefits your T&E program and to figure out your next step in making the most of STAT.

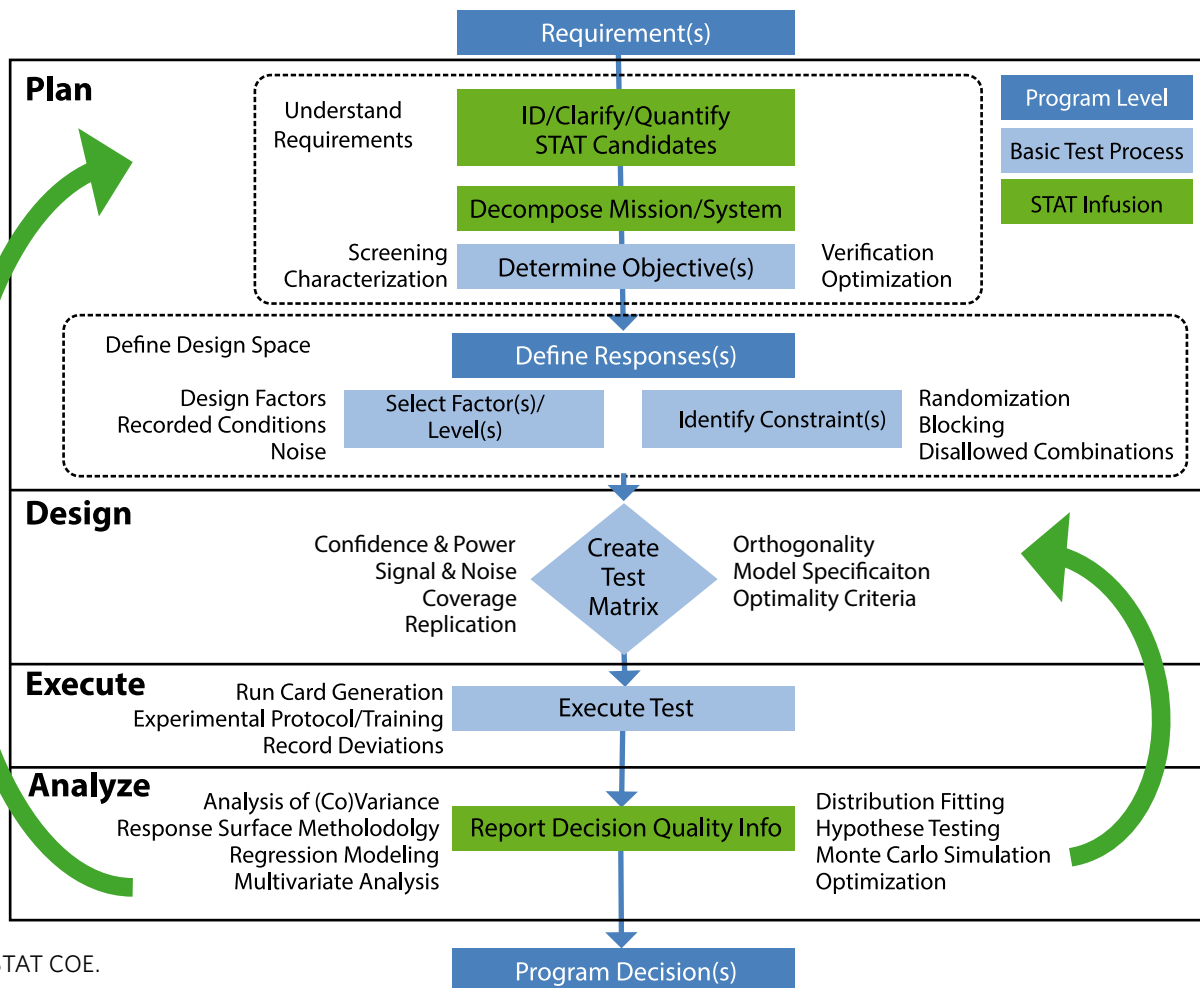
What You Really Need to Know About STAT

Instead of offering a technical, abstract definition of STAT we offer a few concrete ideas to help you understand where STAT may fit into your T&E program.

Despite its use of the word "Techniques" the strength of STAT lies in seeing it as a deliberate process that uses a wide range of rigorous scientific techniques tailored to the T&E process at hand. This paper will use "STAT techniques" when we refer to the technical aspect of STAT, "STAT process" in emphasizing its process aspect, and STAT otherwise. Figure 1 illustrates how STAT is integrated into the T&E event-planning process.

The purpose of the STAT process is to support, not drive the testing process. The results of applying STAT techniques—such as efficient, effective test designs—are not intended to dictate testing decisions but provide information to support better testing decisions and consequently improve the decision quality and defensibility of

Figure 1. STAT in the Test and Evaluation Process Schematic



Source: STAT COE.

testresults. Like any tool, STAT is not meant to be the tail wagging the dog.

DoD 5000.02 notes that scientific test and analysis techniques also includes Design of Experiments (DOE) methodologies. Although DOE is definitely a major STAT technique, it is only one of many scientific test and analysis techniques. STAT techniques are applicable to all kinds of T&E processes (Developmental, Operational, Integrated, Live Fire, Cyber) and to any scale of testing from laboratory bench testing to System of Systems testing. There are STAT techniques for handling both experiments (systematic manipulation of factors to determine their effect on responses) and non-experiments (correlations, surveys, interviews, observations, case studies).

Adoption of STAT by DoD acquisition organizations represents a significant change in culture from the “one-factor-at-a-time” or “we’ve-always-done-it-that-way” approaches to test design.

How STAT Can Make a Difference More Effective Testing

Methodically following the structured STAT process (Figure 1) is essential to the overall effectiveness of your T&E program. In particular, the STAT process focuses up front on developing a set of clear test objectives to drive the need for test resources throughout the testing process. Test objectives derived in this manner accurately reflect the goals of the program for the following reasons:

- Derived from the requirements
- Reflect the focus and purpose of testing
- Further define the scope of testing
- Specific, unbiased, measurable, and of practical consequence

A good test objective for a weapon system would be to characterize its end-to-end mission effectiveness over the operational envelope as opposed to focusing on simply verifying the set of threshold/objective requirements. This continuing focus on explicit test objectives, a hallmark of

the STAT process, virtually guarantees a more effective testing process.

More Efficient Use of Test Resources

The sections below discuss how some of the STAT techniques and tools enhance the efficient use of scarce test resources. The last section highlights recent dollar savings from STAT-COE-supported programs due to more efficient use of test resources.

Optimizing Test Designs

A key efficiency of DOE is the use of software tools such as JMP to design an optimal set of test runs taking into account the large number of sources influencing the test design—factors, values of the factors (levels), allowed combinations of factor values, interaction between factors, maximum number of test runs, and specified level of confidence in test results, to name a few. Use of a tool like JMP will allow efficient trade-offs between the levels of confidence and coverage of all possible test sets. Furthermore, DOE can be used to screen factors, thereby removing needless factors.

Efficient Sequential Testing

Sometimes it is more efficient to first execute a small run of carefully designed test points to collect information about the response variable at selected points in the factor space instead of running a single large set of points. The points in the next set of runs in the sequence would be chosen based on the results from the previous sequence. This deliberate exploration of test points in the test space enables a very efficient test strategy.

Using Combinatorial Optimization (CO) Tools

Consider the need to test complex application screens with:

- A large number of different fields (inputs) in a screen
- Multiple possible values associated with each field
- Many possible test cases required to exhaustively test a screen

STAT tools like the National Institute of Standards and Technology's (NIST) Automated Combinatorial Testing for Software (ACTS) can be used to select sets of test runs that allow test resource managers to optimally tradeoff the number of test runs against the coverage of combinations of factors.

Quantified Resource Impacts of STAT

Table 1 displays documented dollar savings for some of the acquisition programs supported by the STAT COE in 2018.

More Defensible Results

The final output of the test process, the analytical results that support the decision maker, are more defensible be-

cause they are built on the rigor invested in applying STAT across the test planning, design, execution, and analysis phases. Rigor in the test process can be assessed by the way each key component in each phase is handled—its accuracy, exactness, exhaustiveness, meticulousness, and precision. For example, rigor in the planning phase might mean that STAT-related requirements are clear and unambiguous, are quantifiable, and can be accurately measured.

More Resilient Test Designs

Even the most carefully formulated test designs are likely to face a slew of uncertainties before being executed. Potential problems include:

- Immediate gain or loss of test resources
 - Funding (more/less test cases)
 - Test assets (more/less)
 - Instrumentation (change in measurement accuracy)
 - Test range availability (more/less time)
 - Test results data (response/factor values)
- Uncontrollable changes in testing conditions
 - Inability to control value of factor from one test case to another as expected
 - Tests that cannot be executed in the originally planned order

STAT experts working closely with key T&E personnel (Chief Developmental Tester, T&E Manager) can typically provide a complete redesign of the remaining tests consistent with the program's current testing objectives in a matter of hours (or even minutes) instead of days.

STAT Implementation Best Practices

We need to examine best practices for implementing STAT in the areas of people, training, integration, and external support.

People

The most important attribute of the people involved in applying STAT is their willingness to pro-actively learn and implement STAT and to gain a working knowledge of the

Table 1. Documented Savings from Application of STAT

Service	Program	Savings to Date
Navy	Littoral Combat Ship	\$27.7M
Army	Common Infrared Countermeasures	\$2.8M
Air Force	Joint Strike Fighter (JSF)	\$1.2M
Navy	Next Generation Jammer	\$980K
Air Force	Combat Rescue Helicopter	\$420K

Key: K = thousands; M = millions of dollars.
Source: STAT COE.

STAT process instead of waiting until it is needed. STAT techniques continue to grow and evolve over time; if people do not keep their STAT knowledge and skills up to date, their effectiveness in tackling the growing challenges of applying STAT to the increasingly complex T&E process will diminish. Thus, the people providing STAT support to your program should ultimately be full-time STAT experts.

Training

From two major perspectives, training is essential to facilitating the desired change to organizational culture. Key acquisition program personnel and the entire T&E team, at a minimum, need to understand the fundamental STAT concepts that will be employed. Frequently the best way to accomplish this training is to have the STAT support team develop and present “STAT 101” training. The second perspective involves specialized training to fill a specific need, such as a briefing to the test execution team to ensure critical aspects of the test methodology are followed. Additionally, senior leader STAT training may be required.

Integration

The STAT COE’s experience has shown that STAT can make the most difference when the STAT experts are fully integrated into your T&E program. They need to know what’s going on in the program and be invited to participate at all relevant meetings. STAT experts should routinely be part of the discussions about what the requirements mean and how to ensure that program performance is correctly assessed. Ensuring two-way communication with STAT support provides more opportunities for any lingering questions to be answered and increases the chances of a mutual understanding of what needs to be done when changes of direction occur. Working collaboratively with your STAT experts will help them gain a better understanding of program requirements and consequently aid them in helping to develop a rigorous test plan that most effectively covers the entire design space while remaining within budget. When STAT support is viewed and treated as an integral part of your team, it’s remarkable how seemingly intractable test design problems are easily handled.

Early integration of STAT into the program increases the likelihood of better formulation of performance requirements and earlier identification of significant test factors. Rigorously testing performance requirements frequently demands precise formulation. For example, the STAT COE supports a Defense Business System program that had numerous pass/fail user response time requirements already cast in concrete when we engaged the program. Such a rigid requirement formulation does not provide a way to quantify the risk of failure as does a requirement that explicitly captures the expected uncertainty. If we had been engaged with the program during the requirements generation stage, we would have had the opportunity to

recommend that the program consider formulating requirements that explicitly take into account risk, such as 95 percent of all response times must be less than 5 seconds.

Programs sometimes need to accurately predict the number of expensive, long lead test resources that will be required early in the program where there is little past experience to guide the decisions. In this situation, savvy STAT experts know how to make the best use of existing test data to design an efficient, effective test program that can be quickly updated as additional test data becomes available.

An important, yet often overlooked, aspect of integrating STAT into your program is ensuring the contracts of your prime contractor and/or system implementer require them to employ STAT in their own testing activities as well as to support STAT tasks as needed in government testing. Consequently, STAT-related tasks need to be incorporated into the Statement of Work/Performance Work Statement and related contractual documents.

External Support

The primary source available to you for external support is the STAT COE. This center has gained extensive experience providing direct support to a wide variety of DoD acquisition programs. Program managers can request STAT COE support by calling (937) 255-3636 x4736 or via the email address below. To learn more about the STAT COE and its resources (Short Courses, Best Practices, Tools, Test Planning Guides, Tools, Newsletter, Ask a STAT) go to <https://www.AFIT.edu/STAT>. We have found that STAT-trained individuals who have been mentored under the supervision of a STAT COE expert while supporting a T&E program also provide excellent STAT support.

Make the Most of STAT

If you can answer “Yes” to the questions below without blinking an eye, congratulations! Your program is making the most of STAT and we at the STAT COE would appreciate your sharing with us what is working for you.

- Do you believe that your testing process answers the right questions?
- Are you getting the most out of your test resources?
- Do you believe that the rigor of your process ensures that the right question has been answered?
- Can your T&E program quickly adjust your test designs to changes in test resources and testing conditions?

If, on the other hand, you are not satisfied with what STAT is doing for your program or need to learn more, reach out to the STAT COE for help (email: COE@AFIT.edu).

The authors can be contacted at william.rowell.ctr@afit.edu; seth.guldin.ctr@us.af.mil; steven.thorsen@afit.edu; and darryl.ahner@afit.edu.