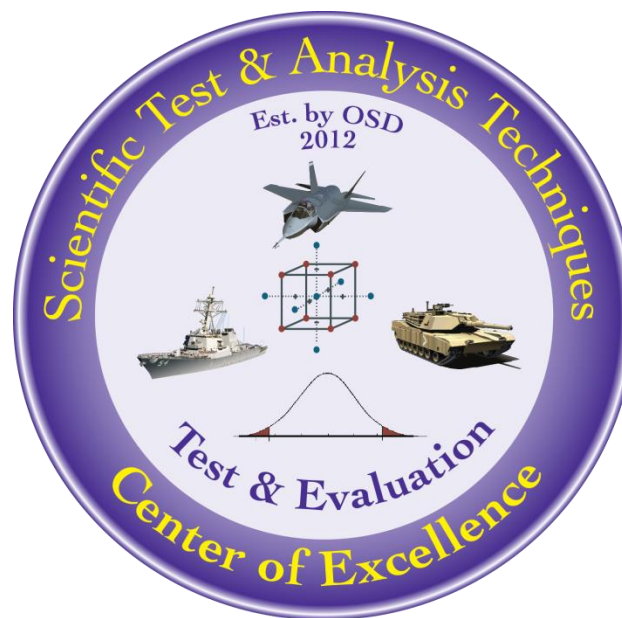


Developing Effective Test Objectives

Authored by: Lenny Truett, PhD

July 2015

Revised: 7 September 2018



The goal of the STAT COE is to assist in developing rigorous, defensible test strategies to more effectively quantify and characterize system performance and provide information that reduces risk. This and other COE products are available at www.AFIT.edu/STAT.

Table of Contents

Introduction	2
Difficult.....	2
Collaborative Approach	3
Unambiguous Objectives	4
Unbiased	5
Specific	5
Measurable	5
Of Practical Consequence	5
Sequential	6
Iterative Process.....	7
Summary and Conclusion	7
References	8

Revision 1, 7 Sep 2018: Formatting and minor typographical/grammatical edits

Introduction

There is a wonderful exchange between Alice and the Cheshire Cat in *Alice's Adventures in Wonderland* by Lewis Carroll;

"Would you tell me, please, which way I ought to go from here?"

"That depends a good deal on where you want to get to," said the Cat.

"I don't much care where –" said Alice.

"Then it doesn't matter which way you go," said the Cat.

"– so long as I get *somewhere*," Alice added as an explanation.

"Oh, you're sure to do that," said the Cat, "if you only walk long enough."

Alice didn't know where she wanted to go, and therefore the cat could not assist her. While Department of Defense (DoD) testing might not be guilty of having no direction, we are all too often guilty of not having good direction. And just like Alice, we do eventually get somewhere, if we test long enough. The problem is that where we get to may not deliver quality information to decision makers.

When the Scientific Test and Analysis (STAT) Center of Excellence begins to work with a program, we are often asked the same question, "Would you tell me, please, which way I ought to go from here?" Before we can be of any "technical" help, we must understand the direction the program needs to go. While the program may have a general idea, there are usually many questions that cannot be easily answered. The direction of any testing program is driven by the objective(s). A great number of difficulties in DoD testing can be directly traced back to the lack of an effective objective. Developing effective test objectives is difficult, collaborative, unambiguous, sequential, and iterative.

Difficult

The first thing to realize is that developing an effective test objective is difficult. There seems to be an assumption that the "best direction" is the most obvious direction. However in practice, an effective objective is almost never obvious. In one of the seminal papers on test planning written by Coleman and Montgomery (1993), they state, "writing the objective is harder than it appears to most experimenters." In his popular DOE textbook, Montgomery (2017) adds, "This (recognition and statement of the problem) may seem to be a rather obvious point, but in practice often neither it is simple to realize that a problem requiring experimentation exists, nor is it simple to develop a clear and generally accepted statement of the problem."

The next factor that contributes to the difficulty of developing an effective test objective is that this skill, along with proper planning, is not usually emphasized, even to highly trained professionals. A 2004 report from the National Research Council stated, "Within the statistical community there has been much more attention paid to the development of efficient techniques for design of experiments than on

the planning process that precedes it.” In a 1993 paper Hahn noted, “Experimental design is both an art and a science. The science deals with the mathematics and formulations of developing experimental plans. This is what most of the literature ... is about. The art of experimental design provides the framework for an effective test program that is maximally responsive ... to the questions that the investigators wish to answer. It deals with the important but seemingly non-statistical topics such as defining the goals of the [test] program.” In a 2013 article, Vining noted, “Most modern textbooks focus primarily on the analysis of the experimental results rather than the proper protocol ... As a result, most modern experimental design textbooks are truly little more than expositions on the analysis of variance (ANOVA).” We will not be successful at developing effective test objectives until we acknowledge that it is not an easy process, and that most practitioners are not properly trained in the “art” of experimental design.

Collaborative Approach

In any acquisition program, there are many stakeholders that have different roles. One way to illustrate the needs of the stakeholders is shown in Figure 1. The test process to address all of their requirements is very complex and no one person could possibly have enough knowledge to adequately represent everyone’s needs.

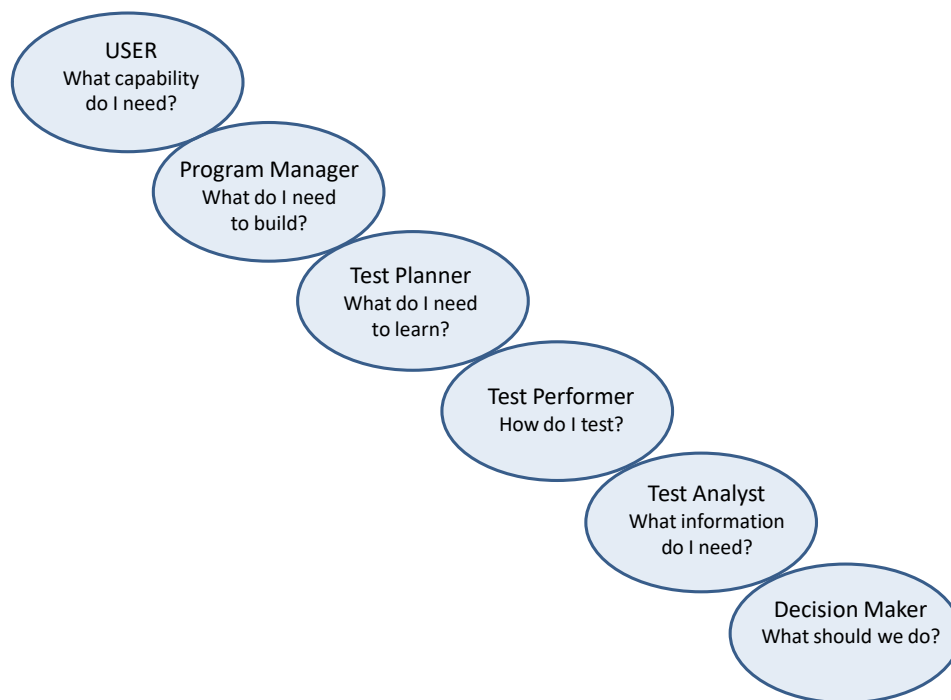


Figure 1. Program Stakeholders’ Needs

Montgomery (2017) writes, “To use the statistical approach in designing and analyzing an experiment, it is necessary for everyone involved in the experiment to have a clear idea in advance of exactly what is to be studied, how the data are to be collected and a least a qualitative understanding of how these data are to be analyzed... It is important to solicit input from all concerned parties... For this reason, a team approach to designing an experiment is recommended.” We recommend establishing a STAT working group with representation from all stakeholders. The purpose of this group is to develop objectives for all testing and ensure that everyone is committed to fulfilling their role in the process to deliver useful information to the decision makers. Box, Hunter, and Hunter (1978) wrote (paraphrased) the statistician or other members of the experimental team should “ensure that all interested parties agree on the objectives, agree on what criteria will determine that the objectives have been reached.” The objectives are the chain that keeps everyone together (Figure 2).

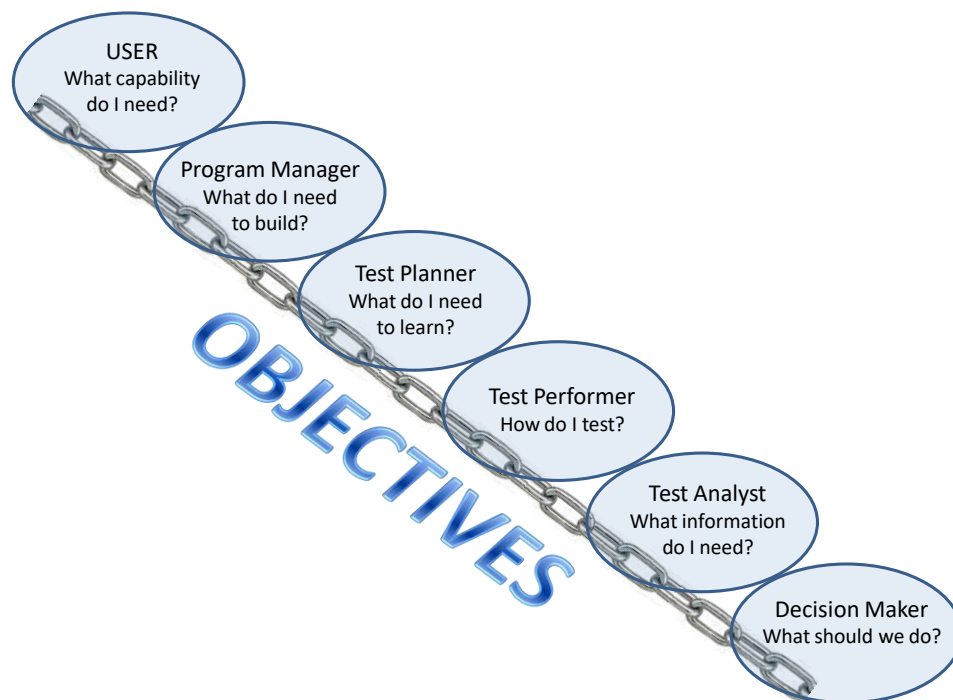


Figure 2. Objectives and Program Stakeholders' Needs

Unambiguous Objectives

No one would intentionally develop an ambiguous objective, but just because people do not intend to be ambiguous doesn't prevent it from happening. If this was not a prevalent problem, we would not have 1.3 million lawyers in the US. It is a common practice to have a lawyer review a contract to ensure that it will produce the desired result. In a similar manner, a test team should have an STAT process expert involved in the development of the objectives to ensure they will effectively inform the decision

makers. To avoid ambiguous objectives, Coleman and Montgomery defined these four guiding principles. Objectives should be *unbiased*, *specific*, *measurable*, and *of practical consequence*.

Unbiased

To be unbiased, the team must encourage participation by knowledgeable and interested people with diverse perspectives (Coleman and Montgomery, 1993). A collaborative approach to developing objectives described above should protect against bias. Any objective that was not developed by a diverse group should be vetted to ensure that it addresses the needs of the entire team.

Specific

To be specific, it should be clear when the objectives have been met. To help define success or failure, the objective should consider aspects such as null and alternative hypothesis, type I and type II error rates, under what conditions, what type of personnel, what locations, etc. Each case is unique, but just like more detail is almost always better in a legal contract, more detail is almost always better in an objective.

Measurable

In the DoD, we generally develop objectives that are measureable, but usually we do not consider how accurately we can measure the response. Often test are designed using a signal to noise ratio, but this just hides the fact the objective has not defined a meaningful difference (signal) and made a reasonable estimate of the standard deviation of the response (noise). Designing a test without a thorough understanding of these measures will get you somewhere, but probably not where you wanted to go.

Of Practical Consequence

This is the aspect most often not considered in DoD testing, and it is actually the most important. Coleman and Montgomery (1993) wrote, **“To be of practical consequence, there should be something that will be done differently as a result of the outcome of the experiment.”** This is often not properly considered because it requires the tester to clearly understand what the user needs (that is to say “what is good enough?”) and what information would cause the decision maker to change the course of the program. The second part is particularly difficult because there are many factors that go into an acquisition decision. As testers, we need to provide the critical information that could justify significant decisions, and not just report on a requirement with a confidence interval. While collaborating on objectives will not ensure that they are of practical consequence, not collaborating will almost certainly guarantee that they will not be.

Sequential

Too often, the objective is far too broad in an attempt to “cover everything”. Following the steps above will help to refine the objective, but the concept of developing **sequential** objectives is also critical.

Montgomery (2017) has an excellent description of this issue in the *Design and Analysis of Experiments*.

An important aspect of problem formulation is the recognition that one large comprehensive experiment is unlikely to answer the key questions satisfactorily. A single comprehensive experiment requires the experimenters to know the answers to a lot of questions, and if they are wrong, the results will be disappointing. This leads to wasting time, materials, and other resources and may result in never answering the original research question satisfactorily. A sequential approach employing a series of smaller experiments, each with a specific objective ... is a better strategy.

This concept is also consistent with the “Shift Left” initiative from DASD(DT&E) (2013). The basic principle is to test what you can, as early as you can, and not try to test everything with one comprehensive test during operational testing. Developing effective objectives for each sequential stage will enable this process by getting the entire team working together and understanding, at least qualitatively, how each phase of testing will be utilized. All testers need to have knowledge of the objectives of the other stages to understand how they contribute to the entire testing process and to ultimately ensure that decision makers will get the information they need. Previously, the objectives were referred to as the chain that holds all of the stakeholders together. It can also be asserted that they are the chain that holds all of the phases of testing together (Figure 3).

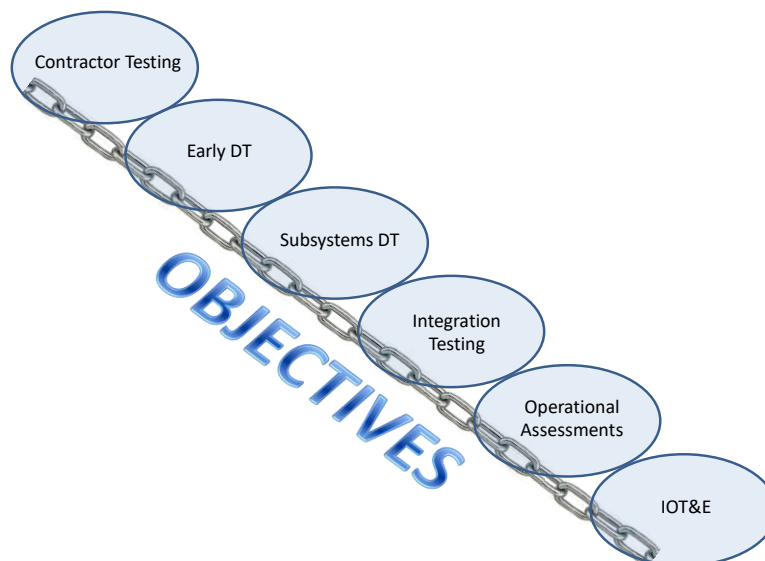


Figure 3. Phases of Testing

Iterative Process

Just because your team has successfully developed detailed sequential objectives doesn't mean the process is complete. Situations change during the course of every test program and this will require the entire team to re-examine the objectives of the remaining tests. The objectives may need to be updated because of budget, schedule, capability limitations, requirement changes, new discoveries, etc. The point is that as the program evolves, the test objects should also evolve. The earlier quote from Box et. al (1978), continues by adding "if objectives change, all interested parties will be made aware of that fact and will agree on the new objectives and criteria." Getting agreement from all stakeholders can be a difficult but it is absolutely critical to the process. Objectives should be codified in a document, and any updates should be discussed and approved in writing by all stakeholders. This concept is not unfamiliar to the DoD because stakeholders must approve all updates to the Test and Evaluation Master Plan (TEMP) and do so on a recurring basis. Any objectives that are not iteratively updated may become obsolete and any updates that are not approved by all stakeholders may become ineffective.

Summary and Conclusion

Developing effective test objectives is difficult, collaborative, unambiguous, sequential, and iterative. The objectives should be developed by a team representing all stakeholders and all phases of testing. Everyone should have at least a qualitative idea of the entire process. The objectives must be unbiased, specific, measurable, and of practical consequence. They should detail a sequential approach to testing and should be regularly updated by all stakeholders as the program evolves. The process of developing and updating objectives should be a team effort guided by a STAT process expert. This proposed methodology should result in a living document, signed by all stakeholders, that helps guide the testing process and facilitate the delivery of defensible and actionable information to the decision makers.

References

Box, George E.P., Hunter William G. and Hunter J. Stewart, 1978, *Statistics for Experimenters*, New York: John Wiley & Sons, Inc., 1978.

Coleman, David E. and Montgomery, Douglas C., "A Systematic Approach to Planning for a Designed Industrial Experiment," *Technometrics*, vol. 35, no. 1, 1993, pp. 1-12.

Hahn. G.J., "Discussion: Experimental Design Review and Comment," *Technometrics*, vol. 35, no. 1, 1993, pp. 15-17.

Montgomery, Douglas C., *Design and Analysis of Experiments*, 9th ed, John Wiley and Sons, Inc, 2017.

National Research Council of the National Academies, "Improved Operational Testing and Evaluation, Phase II Report," The National Academy Press: Washington D.C., 2004.

Vining, Geoff, "Technical Advice: Experimental Protocol and Basic Principles of Experimental Design," *Quality Engineering*, vol. 25, 2013, pp. 307-311.

<http://www.dau.mil/publications/DefenseATL/DATLFiles/Sep-Oct2013/Hutchison.pdf>