Air Force

Analyst's Handbook

On Understanding the Nature of Analysis

by
Christopher A. Feuchter
January 2000

Office of Aerospace Studies
Air Force Materiel Command (AFMC) OAS/DR
3550 Aberdeen Avenue SE
Kirtland AFB, NM 87117-5776
www.oas.kirtland.af.mil

Preface

The goal of this handbook is to provide a clear description of the basic elements and practices of operations analysis as it is conducted in the Air Force. It is the product of the Air Force analysis community as a whole and is formally sponsored by the Air Force Analysis Community Steering Group. It is intended to be both an introduction for newcomers and a reference for career operations analysts. It was undertaken with the belief that good analysis practices are teachable and to a large extent can be standardized—despite the fact that analysis often thrives on ingenuity and innovation. If there will be a measure of success for this handbook, it will be in its use. With that in mind I have tried to give practical advice illustrated with real-life blunders and successes. I realize that such a handbook is likely to generate a variety of responses—it is the nature of analysts to have strong opinions. I extend an invitation to you, the reader, to provide whatever feedback you deem appropriate.

While I undertook this handbook to give back to the community much of what it gave me, there is no way I could have managed it by myself. Mary Benze, Keenan Kloeppel, Larry Looper, and Maj Rich Roberts were kind enough to contribute chapters respectively on cost analysis (Chapter 12), model VV&A (Chapter 9), modeling human behavior (Chapter 10), and executing the analysis (Chapter 8). Additionally, I thank the hardy (if small) corps of reviewers who worked to keep me honest and out of too much trouble: Col John Andrew (AFIT), Lt Col Joe Auletta (OAS), Patty O'Brien (AFSPC), Capt John Dulin (AFPOA), Rich Freet (ACC), Elaine Goyette (Mitre), Jim Haile (OAS), Bruce Merrill (OAS), Lt Col Roxann Oyler (OAS), and Dr Roy Rice (Teledyne Brown). They made many fine suggestions.

Christopher A. Feuchter
Technical Advisor
Office of Aerospace Studies

Contents

Figures

Tables

1 What's It All About, Alfie?

Basic Dogma

Operations research or OR—also referred to as operations analysis, operations research analysis, systems analysis, and system engineering to name a few aliases—is the investigation of the processes and interactions of human and hardware systems within selected scenarios. OR's goal is to provide information for making decisions. The nature of that information is insight (a more definitive word would be inappropriate) into possible alternative futures. The information generally provides a better understanding of interactions of selected system aspects and the operating environment. In many instances OR-based decisions result in lower cost and/or better performing solutions.

OR relies on three procedural components: Modeling, Simulation, and Analysis. These are collectively abbreviated MS&A. In practice these terms are frequently—and disconcertingly—used interchangeably. We will use the definitions given below:

- ***Modeling*** is the process of creating representations of systems and processes and their interactions.
- ***Simulation*** is the process of using ("executing") models to investigate system performance.
- ***Analysis*** is the process of using the results ("outputs") of the simulations to identify and display important relationships among systems, their processes, and measures of how well they perform.

Data ("inputs") are critical to MS&A. Modeling considers the nature of available data when shaping the representations of systems and processes, and simulation cannot go forward without appropriate sets of model inputs.

Cost estimating ("costing") also is integral to MS&A. Cost estimating is the process of estimating how much money is required to build, test, operate, and maintain the systems being studied. There can be no cost-effectiveness assessments if there are no cost estimates.

Purpose of This Handbook

This volume focuses on the "nature of OR." More specifically, it will address the process of OR as it relates to complex analyses such as Analyses of Alternatives (AoAs), mission analyses, and determination of force structure. This is done with the belief that simpler analyses tend to be microcosms of the complex analyses. Where this is not true, we will attempt to illuminate the differences.

There are many fine books and journals that address the tools of OR; there is little easily accessible material in the literature that addresses the details of the process, especially the military OR process. The OR process involves:

- Defining the problem and its scope
- Selecting practical methodologies and criteria for investigating the issues
- Defining ground rules and assumptions
- Integrating the methodologies in a manner that will provide the needed insight
- Balancing the methodologies against schedule and resources
- Organizing and managing the OR team
- Extracting and presenting the results

The acquisition of these skills has most often been through hard-earned experience. We will facilitate that learning process by sharing some of the knowledge and experience of the Air Force analysis community.

Plan of Attack

We have organized our discussion of OR and MS&A into fourteen chapters. As much as possible we will illustrate concepts with examples taken from life, drawing liberally from the personal experiences, anecdotes, and aphorisms every analyst maintains to illustrate their hard won wisdom. Many of these will appear in sidebars throughout the handbook. They are intended to be instructive and amusing.

Systems and Entities

OR studies systems. The concept of a system is simple: a system is a collection of interrelated elements performing one or more functions. System elements are generally referred to as entities. Entities need not be related in any way except through their common purpose or effect.

Systems may be created by humans. They may also occur naturally, e.g., a weather system. The beauty of the concept of a system is that almost anything can be thought of as a system. Something as lowly as a potato peeler is a "peeling system." Something as complex as the collection of our schools is our educational system. Even the entities of a system can be systems—one of the entities of our highway system is the interstate highway system.

The state of a system describes the values of the variables required to characterize the nature of the system at a particular point in time. If our purpose is to address the size of our educational system, total enrollment might adequately describe state of the system. However if we want to specify its effectiveness, scores from a series of standardized tests are more appropriate state variables.

First Principle of OR

It's possible to do a study fast and cheap, fast and good, or cheap and good. You cannot do a study fast, cheap, and good. Not ever.

Second Principle of OR

Every universal truth in OR has at least one exception.

Ways to Study a System

Direct Experiment

Conceptually, the simplest way to study a system is by experimenting with it directly. In many instances this is difficult or impossible. To begin, the system may not exist. If it does, it may be impractical or overly expensive to use it in experiments. The alternative to system experimentation is experimentation with a system model. Such a model may be either physical or mathematical.

Experimentation with a System Model

Physical Models

An often-convenient way to study a system is with a physical model. Physical models may range from balsa wood mockups to breadboard fabrications to working prototypes. We will not consider physical models in this handbook, focusing instead on mathematical models. That does not imply that physical models are not important in some types of Air Force analysis.

Mathematical Models

The heart of MS&A system modeling is the mathematical model. Just as a physical model can take a variety of forms, a mathematical model may range from a simple equation to a large, complex computer code. Mathematical models are either deterministic or stochastic. Deterministic models allow a computed solution. Stochastic models do not, because they have, at least implicitly, some random aspects (models can be classified in other ways as well; we'll discuss this later in the chapter).

Why We Model

Practically, we model when modeling is better than experimenting with the system(s) in question. The major reasons why modeling is the better or only alternative are discussed below.

Problem Complexity

Experimentation with complex systems and scenarios may be impossible or impractical. We cannot, for example, engage in real combat in order to test the performance of new military systems (although we analyze actual combat experiences extensively). Nor can we necessarily disrupt the operation of an existing system to try new operational concepts. In these and similar situations, the application of OR is the only way to obtain information about system performance or the value of new ways of doing business.

Evolution and Tempo

We live in the present; we live at real time. In studying a system there may be a need to study its past or future performance or to study it in non-real time. It is easy to imagine a system—say a traffic control system—and ask how it will perform as traffic volume increases with future population growth. Or it may be necessary to slow down the pace of a system to understand cause and effect within the system. These tasks cannot be done by direct observation of the system.

Limited Funds and Time

Large scale system testing is expensive and time-consuming. Insufficient funding and time often lead directly to modeling.

Safety Concerns

Modeling is often used to reduce the risk to life and property that would accompany system testing or actual system implementation. An excellent example is a proposed nuclear waste disposal site. In this case, the only alternative to studying the system with MS&A is actual storage at the site. This is not considered an acceptable option.

Lack of Access

Systems lacking access are either conceptual (see Chapter 10) or physically inaccessible for experimentation. In either case, experimentation with the system is impossible, although if the system is real it might be observable. Unexploited hostile threat systems are good examples of this latter situation. Such systems lead immediately to the need for a model.

Classifying Models

A model is a physical, mathematical, or otherwise logical representation of a system, entity, phenomenon, or process. Some models have inputs and outputs and are run; others serve only to facilitate thinking about how to do the analysis. We frequently describe models by certain of their characteristics. It is useful to have a nominal understanding of some of these descriptive terms, as they are useful in conceptualizing the basic purposes and structures of models. The list in this section is not complete. Neither are the terms necessarily mutually exclusive.

Executable vs. Non-Executable

In the OR community the word model will generally conjure up a vision of a mathematically based model requiring inputs and run—usually on a computer—to produce outputs. These models are clearly executable, and indeed they are the typical means of producing analytical results. However, before the executable model is written or before it is integrated into a study's methodology, it will probably be part of a non-executable model. Non-executable models are constructed to help visualize relationships and processes, to establish logical consistency, or to insure identification of all relevant elements of a study. Often they will take the form of a simple flow chart connecting models, data, and processes. They can also be schedules, outlines, lists, circuit diagrams, decision trees, or even a physical mockup. The value of non-executable models is often overlooked by inexperienced members of the OR community.

Static vs. Dynamic

If the situation described by a model evolves with time, the model is said to be dynamic. If not, the model is said to be static. Most OR models are dynamic. A model that portrays the interaction of a missile and the aircraft it is attacking is a dynamic model. Static models describe a situation at a point in time. They tend to be statistical models. For example, a model of a bomb's miss distance based on drawing randomly from a probability distribution is a static model of miss distance. It should be obvious that a dynamic model can incorporate static aspects.

Deterministic vs. Stochastic

A deterministic model is a model that contains no random aspects. Once a set of inputs has been selected for a deterministic model, the model outputs will be the same every time the model is run. Deterministic models usually perform mathematical computations, for example, integrating differential equations or solving systems of linear equations. A model that draws random numbers to determine its course is a stochastic model. Stochastic models can be as simple as a one server queuing model or as complex as a theater campaign model. A stochastic model will often have deterministic components.

Discrete vs. Continuous

A model is discrete if its variables occupy distinct (discrete) states. For instance, an event has or hasn't occurred—a missile has or has not been fired. Complex models are generally event-oriented and will thus be discrete. Nothing is explicitly modeled between events in a discrete model. A model is continuous if its variables are continuous, usually as a function of time. Lanchester's force-on-force differential equations constitute a continuous model. It is possible for a model to have both discrete and continuous attributes.

Event-Stepped vs. Time-Stepped

Discrete models are typically either event-stepped or time-stepped. An event-stepped model moves ahead in time by moving from the current event to the next scheduled event. This requires that the time of occurrence of each modeled event be determined in proper time sequence with all other events. For this to happen, all events are preprogrammed, generated by the occurrence of previous events, or occur randomly. A time-stepped model, on the other hand, moves forward in time by stepping the model in discrete time increments. When an event occurs in a time-stepped model, it is considered to occur at the time step at which it was noted. Some time-stepped models can vary the incremental time step, making it longer or shorter to adapt to the situation.

Descriptive vs. Optimizing

Descriptive models are models that mathematically describe “real world” situations. They are typically used to evaluate the impact of changes to physical or procedural aspects of systems or scenarios. Optimizing models are models that are designed to find the best solution(s) to specific problems. Linear and non-linear programming models are optimizing.

Prescriptive

Prescriptive models are designed to recommend a course of action. Scheduling models are examples of prescriptive models. Prescriptive models may be either descriptive or optimizing models.

Potential MS&A Pitfalls

OR represents the best means we have for providing decision-making information, but as good as OR may be it is not without potential pitfalls under the best of circumstances. Below we have provided a brief list of what can go wrong—even when good analysts use good models. Some of the problems are inherent to the MS&A process. The analyst introduces others due to lack of access to information, inexperience, or inattentiveness.

- The model has undiscovered errors. Virtually all complex models contain errors. Serious errors tend to be found and corrected quickly, but later modifications to the model can create new serious errors.
- The model is poorly documented. Poor documentation leads to a poor understanding of the model’s capabilities and limitations. This increases the likelihood of inappropriate model usage.
- Model fidelity is not matched to the question. The model may not represent the real world with enough fidelity to accurately answer the question. Conversely, the model may possess such detail that it obscures the bigger picture being sought.
- The assumptions made by the model user are not consistent with the assumptions built into the model or input data. This can occur because of inadequate documentation. Analyst inexperience with the model or lack of diligence and care are also causes.
- The model is used outside its validated input parameter range. Models are tested for limited ranges of input values. Using the model in unanticipated ways can result in these limits being exceeded with potentially significant consequences.
- Suitable inputs are not available. This is essentially a variation on the well-known garbage in, garbage out problem. The best analyst running the best model still needs good inputs.
- Model-user interfaces are complex and lead to errors in preparation of model input or capture of model output. User-friendly interfaces are typically an afterthought.
- Model fidelity is not consistent throughout the model. What the modeler understands well tends to be well modeled; what is poorly understood is necessarily poorly modeled. The consequences of these discrepancies are often difficult to disentangle.
- Results are misinterpreted or used out of context. Once the analysis has been briefed, the results tend to take on their own life. People often see only what they want to see. The only precautions against this are fully annotated briefing charts, or better, a thoroughly documented report.

Model Verification, Validation and Accreditation (VV&A)

For a model to be useful it first must perform as it was designed to perform. Second, it must represent the real-world situation it is emulating “well enough.” Models meeting the first criterion are said to be verified. Models meeting the second are said to be validated. A third hurdle for a model is accreditation. Accreditation is the affirmation or decision that the model is suitable for a particular usage. While there are no absolute VV&A standards, this collection of processes when conscientiously executed legitimizes selection and usage of the model in question. Unfortunately in practice, VV&A is often given short shrift or it is totally neglected. Clearly when this is the case, the authority of the analysis is called into question.

2 And the Answer Is...

Legitimate and Not-So-Savory Uses of OR

When I first started thinking about this handbook, it seemed worthwhile to create a taxonomy of the uses of Air Force OR. Before I was finished I had two classification schemes: one based on the type of questions asked and a second based on the generic issues that generate the questions. These schemes appear below. I also wanted to discuss post-decision analysis—the "make my decision right" analysis. This discussion concludes the chapter.

Question by Type

In Chapter 1 we indicated that OR is performed to aid in making decisions. These decisions are generally about selecting solutions to problems. Solutions may be either materiel or non-materiel. Materiel solutions involve building or modifying hardware—things. Non-materiel solutions involve altering organizational structures and/or processes to solve a problem. In either case four basic types of analytic questions can arise:

- Can I do it? (feasibility)
- Does it make sense to do it? (suitability)
- What is the best answer? (comparability)
- Is my answer widely applicable? (sensitivity)

It is common for a study to address more than one of these questions. For example, in a formal Analysis of Alternatives each alternative will be examined for suitability, compared to the other alternatives, and frequently, examined for its sensitivity to changing inputs.

Feasibility Analysis (Will It Work? Can I Make It Work Better?)

At some point early in the life of a proposed solution to a problem, the question should be asked: Will it work? Analysis done to answer this question is called feasibility analysis. Often the solution will need several increasingly sophisticated feasibility analyses during its development as the understanding of both the problem and the proposed solution mature. This is especially true for solutions that involve cutting edge technologies.

Suitability Analysis (It Works, but is It Suitable to the Task?)

Once a solution to a problem is judged feasible, it should be examined to determine if it is practical (suitable to the task). There are many aspects of suitability: Can I afford it? Can I produce it? Can I staff it? Can I maintain it? Can I deploy it? Is it survivable? While suitability can often be examined analytically, frequently actual testing is called on at some point to back up the analysis.

Comparative Analysis (How Does It Stack Up against the Competition?)

The heart of choosing a solution lies not in examining a single potential solution, but in comparing different potential solutions. This is done formally in the acquisition process through Analyses of Alternatives. Both effectiveness and cost analysis are essential to most comparative analyses, as effectiveness results without cost results paint a limited picture. Comparative analysis is treated in detail in later chapters.

Sensitivity Analysis (How Robust Is My Answer?)

A potential solution should initially be examined relative to the most likely scenarios it will face. The scenarios can consider the nature of hostile threats and their locations and concepts of operation, terrain, foliage, weather, time of day, climate, etc. However, these initial scenarios often are quite limited in diversity. Once demonstrated successfully in these scenarios, the investigation should be expanded. Sensitivity analysis is designed to determine if "the solution" is applicable over expected ranges of critical parameters.

Question by Issue

Specific feasibility, suitability, comparative, and sensitivity questions do not appear out of thin air. They arise in conjunction with generic Air Force issues that are dealt with over and over again. The five issues below are comprehensive, but certainly not mutually independent.

- Making spending decisions
- Investigating the potential of new technologies and guiding technology development
- Assessing threats and countermeasures
- Developing strategies, tactics and operational concepts
- Planning force structure or organizational changes

Making Spending Decisions

What do I buy? How do I deal with funding realities? These typical spending decisions may not be mutually exclusive, and clearly they are made at every level from the President down to individual organizations. In the "What do I buy?" case there is a general agreement that something must be bought. Analysis determines the most cost-effective options from a list of likely candidates. In the "How do I deal…?" case, the decision is usually programmatic. Can I afford A and B? How do I react to changes in funding? What is my spending plan? These are all crucial decisions.

Investigating the Potential of New Technologies and Guiding Technology Development

Military laboratories, either directly or through contractors, investigate many technologies with potential military applications. In the initial stages of technology development, the analysis supporting this potential is usually sketchy at best. As the technology is developed, it becomes necessary to make detailed assessments of just how valuable the technology will be. This can be investigated through MS&A, through testing of experimental hardware, or a combination of the two. If the technology continues to look promising, MS&A can be used to do sensitivity analyses of the technology to indicate what aspects of the technology—volume, weight, power, accuracy, ruggedness, etc.—need to be pursued during the development process. Of course, throughout the development of a technology, implicit or explicit affordability decisions are made regarding continued development.

Assessing Threats and Countermeasures

The job of assessing threats and countermeasures on friendly systems is never-ending. Such assessments influence our procurement decisions and operational planning processes at every level. If the enemy can do this, then I must react in such and such a way. Unfortunately, the business of gathering information about an enemy is fraught with difficulties. The information being assessed is often fragmentary and subject to error. In such a situation the analyst is quite useful. Analysts can assess the consistency of the available data, make testable predictions for later investigation, set performance limits based on specific assumptions, etc. The intelligence community and the MAJCOMs are the prime players in this area.

Developing Strategies, Tactics and Operational Concepts

Military OR was developed during World War II to answer operational questions. Some of these questions required predictive answers, for example, what is the best search pattern in given circumstances? Others required analyzing collected data to compare different practices in a search for improvement. The operational issues today are much the same, even though computers have replaced paper and pencil and analytic techniques have matured.

Planning Force Structure or Organizational Changes

Changes in force and organizational structures are most often made in response to changing outside influences: goals, missions, threat, technology, budget. They are also made to achieve improved internal efficiency. The Quadrennial Defense Review (QDR) is an example of a formal consideration of such changes. The QDR is supported by analysis conducted by the services. Considerable analytic resources are expended, because service responsibilities and budgets are at stake.

Third Principle of OR

The only effective analysis is analysis that precedes a decision.

The Dark Side of Analysis

Whether we like it or not, every analyst runs the risk of being asked to perform analysis to support a predetermined decision. Implicitly or explicitly, the answer has been decided, and the analyst is expected to show why it was the right answer. In the analysis community this is known not as analysis, but by the technical term of intellectual prostitution.

Obviously, such a situation may be fraught with pitfalls. All too often the requester is your boss or someone higher in your chain of command. How can you do an honest job as an analyst and avoid unfavorable consequences? Perhaps you can't, but we would recommend the following approach to an inappropriate request for analysis.

Begin by analyzing the situation as honestly as possible. This provides the requester with the best available—if not most desired—answer. This answer may be acceptable if it supports the desired outcome, or if it shows that the requester is supporting an untenable position. If neither of these outcomes occurs, the demand for positive analytical answers will be repeated.

Your next step is to use the insights obtained from the original analysis to identify and analyze specific conditions or scenarios favorable to "the answer." You can then present these results in proper context along with the original results.

At this point you have done pretty much all that you can do. It is then up to the requester to determine how your information is used or misused. If you follow this path and your efforts are regarded in a negative light, it may be time to tie your star to someone who understands the proper role of analysis.

3 The Complete Analyst

What Makes a Good Analyst?

There are few professions that can benefit from a roughly equal balance of artistic and technical aptitude. Operations analysis is one. Artistic aptitude is central to the creation of imaginative methodologies to deal with difficult new problems. Technical aptitude allows these methodologies to be implemented. Not all analysts are blessed in both areas, and it is most often the artistic side that is the weaker. This is due to the natural attractiveness of analysis to those with strong technical aptitude and the existence of an educational system to teach the techniques of analysis. Certainly aptitude and education are prerequisites for the complete analyst, yet they share the additional traits discussed below. Are there complete analysts, or are they myth? Yes, Virginia, they are.

Has Passion

Every complete analyst has two great passions: a burning curiosity to know the unknown and a strong desire to produce the "best possible" truth from the available data and tools. Without these passions the analyst is at best a technician, albeit sometimes a capable one. These passions can be learned; it is not clear that they can be taught.

Minimizes Effects of Personal Biases

All of us are biased. The trick is not to let our biases affect the design of a study or the presentation of its results. A complete analyst does this by being indifferent to the study's outcome. Designing an analysis methodology that treats all alternatives fairly manifests this indifference. The design of such methodologies is discussed in Chapter 7.

Interdisciplinary Perspective

Just as analysts are well served by having both artistic and technical aptitudes, they are equally well served by having an interdisciplinary perspective. This means that they are open to using whatever tools and skills can be brought to a problem, regardless of their original application or place of origin. It further means that they actively seek the most appropriate tools and skills that can be applied to a problem. The greater the experience of the analyst, the more varied the toolkit. The experience comes from formal education, from work experience, and from a network of fellow analysts who are willing to share their knowledge.

> **Reality**
>
> **The antithesis of the interdisciplinary approach is embodied in the old saw, "If your only tool is a hammer, every problem looks like a nail." Unfortunately, much of the community believes it has just such a limited toolkit (see Chapter 7).**

Education

There is no single course of study that marks the complete analyst. In truth, suitability for OR resides more in aptitudes, passions and the ability to think logically than in a specific formal education.

Historically, OR began in earnest during World War II. The initial practitioners were in large part professionals, often lawyers. With time, these pioneers gave way to analysts educated as mathematicians, statisticians, physical scientists, or engineers. More recently, analysts with a formal OR education have joined the mix in increasing numbers. Analysis does require significant facility with mathematics, and to a lesser degree with physics and engineering. It is also helpful to have familiarity with the techniques of cost analysis. In specific circumstances, psychology, medicine, and other disciplines can also play a role.

Sense of History

The complete analyst approaches every new task by becoming familiar with related past and ongoing analyses. These efforts can provide insight to the problem, identify useful contacts, suggest reasonable assumptions and detail lessons learned—even provide entire methodologies and databases. A search for these analyses can be done through personal contacts, Internet searches or through online databases such as the Defense Technical Information Center (DTIC) and the National Technical Information Service (NTIS).

Recognizes Reality

No analyst is an island. Every complete analyst must be willing to acknowledge when an impasse has been reached and seek additional help in resolving a problem. In the same vein, the analyst must be realistic when assessing resource needs and the ability to meet schedules. Over-optimism is a fault even among basically solid analysts (see Chapter 7).

Attention to Detail

No analyst can be complete without a strong interest in detail. Every step of an analysis must be understood and examined repeatedly as the analysis progresses. The practice of analysis is not for the strictly big picture person.

Imagination

Analysts are constantly faced with new challenges: providing new data, evaluating a new system, solving a practical problem. Each of these challenges provides opportunities for the exercise of mental ingenuity—the imagination to adapt old methodologies or visualize new methodologies. It is here that the artistry of OR comes to the fore. It is here that the complete analyst's mettle is proved.

History

The prevailing attitude today is that everything done in the past is ancient history with no present-day relevance. This is reinforced by government and contractor workforces that change so rapidly that corporate memories are effectively limited to three or four years (see Chapter 14).

"Every time history repeats itself the price goes up."

- Anonymous

"Those who cannot remember the past are condemned to repeat it."

- George Santayana

Likes to Teach

The OR profession for many years had a resemblance to a medieval guild with masters and apprentices. Knowledge was passed from mentors to their pupils through a combination of observation and tutelage. That system no longer exists. The gray heads have retired and the guild has been replaced with a university education. This is unfortunate, for while the techniques of OR can be taught well by academics, the artistic aspects of OR are taught poorly or not at all. There will always be a need for experienced analysts to guide the apprentices, and a willingness to do so is a hallmark of the complete analyst.

Good Communication Skills

An analyst who cannot clearly and concisely communicate ideas bears a significant liability. The complete analyst must be able to effectively make points *mano a mano* in meetings, prepare and deliver quality briefings, and provide clear detailed documentation when appropriate (see Chapters 13 and 14).

The Way It Was

Twenty years after I signed on as an analyst I was *still* learning from four practicing civilian analysts who were in the office the day I first arrived for work. Today in the same office there are no practicing civilian analysts with more than 10 years experience.

4 OR is a Team Sport

Organizing the Disorganized

In today's complex world the lone analyst slaving away in a garret is an anachronism. Current problems are unrelentingly multidisciplinary, and one person cannot be an expert in all areas. You need engineers and technologists to assure accurate portrayal of systems. You need cost analysts to estimate the costs of competing concepts and the intelligence community to create scenarios. Operators must develop operational concepts and modelers have to work their magic. One quickly realizes that OR is a team sport.

Like any team, an OR team can play well or poorly. Putting a successful team together requires astute selection of personnel, team chemistry and good leadership. You need to "recruit" members who have the necessary knowledge and skills, create an environment in which they can work together productively, and set and enforce team goals and standards.

In this chapter we discuss how to build such a team and make it work. We will focus on teams addressing complex analyses such as Analyses of Alternatives and mission analyses. The problems of less encompassing analyses can often be recognized as being variations on a theme.

Team Organization

In this section we describe a team organizational structure that has been successfully employed for many years. At the top of the structure is the study leader. This individual should be an experienced manager, preferably one with at least a modest analytical background. The study leader's lieutenants are the chairs of panels that address the major aspects of the study. The use of study panels helps ensure that a responsible person leads each important study area. The focus of these panels will vary with the type of study. A common set of panels used in AoAs is shown in Figure 4-1. The arrows in the figure indicate typical

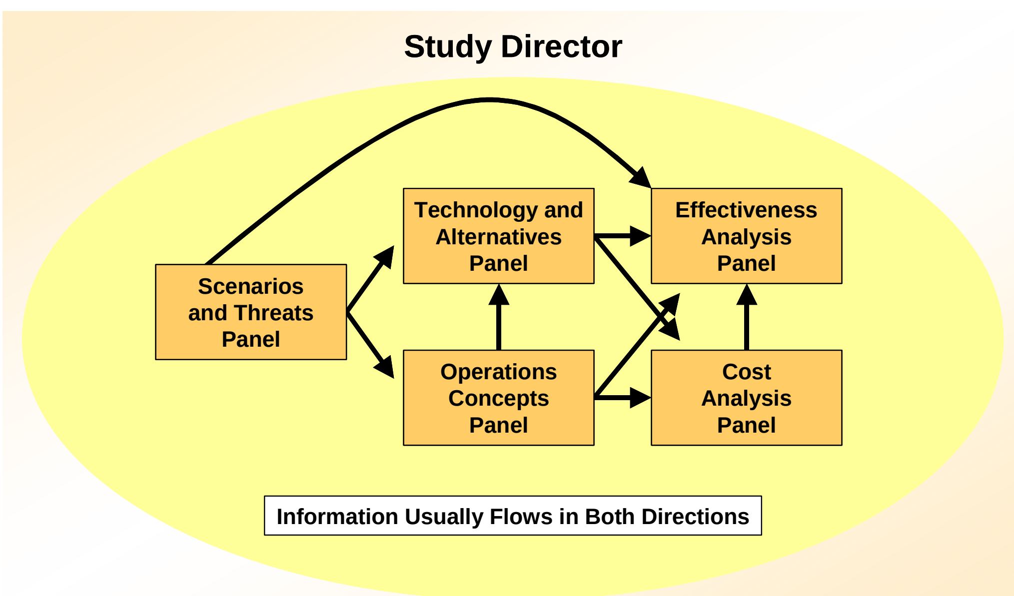

Figure 4-1 : A Common Panel Structure for AoAs

panel interactions.

The panel concept is wonderfully adaptable. Depending on circumstances panels can be split (e.g., separate the technology and alternative panels), panel functions can be combined (e.g., marry operations concepts and alternatives), or entirely new panels can be created (e.g., test, requirements or doctrine panels). The final choice of panel structure nominally belongs to the study director, but it is usually a decision by committee. At a minimum the panels should mirror the major aspects of the study. Thus the creation of panels offers an important early opportunity to contemplate the study's processes and goals.

Every panel must have enough members to achieve a "critical mass" of knowledge, interest and motivation. Typically, a small group with critical mass is more productive than a larger group with critical mass. This suggests starting with a small panel and adding members as necessary.

Participation

The participants in a study should ideally be selected based on their qualifications. Consideration should be given to those with relevant experience, knowledge, and interest in the subject matter. These individuals are frequently referred to as subject matter experts (SMEs) and/or stakeholders. More often than not these people are available. In some cases they have an organizational charter to support your study. In other cases they participate because they are heavily invested in the outcome of the study. When the most qualified are not available, the next most qualified should be sought.

People with diverse viewpoints on controversial issues should be enlisted to participate in a study. Partisans, by their nature, will defend their ideas and detect flaws in the ideas of their competition. This makes for interesting and often acrimonious meetings during the study, but it allows issues to be raised and resolved early that might otherwise lie in wait along the briefing trail.

Leadership

Successful study leadership requires a person with:

- Involvement
- A willingness to lead
- A willingness to learn
- An understanding of the study's background
- An overall plan
- Flexibility to change the plan
- Communication skills
- An understanding of human motivations
- An ability to make reasonable decisions
- Attention to detail
- Political astuteness

Although desirable, in the strange world in which we live, this person may not be the titular study leader. It may be a government analyst, perhaps the effectiveness panel chair; or it could be a contractor serving as the *de facto* study lead. It can even be a shared lead.

Communication

If studies today weren't complex enough, they are further complicated by the wide geographic dispersion of their participants. The customer for study results may be in Washington, the responsible co-study leads in Colorado and Virginia, the principal contractor in California, and…you get the idea. Without frequent and timely communication among the parties, serious problems are a certainty.

For better or worse, the responsibility for good communication usually rests with one person. That person is ideally the study director, but the job may be delegated to any responsible government or contractor party. The tools available for communicating are discussed below.

Meetings

The most potentially effective forum for promoting productive interactions among group members is the face-to-face meeting. Unlike other forums, meetings allow a group to interact over days if necessary. Meetings are generally democratic and give everyone a stake in what is happening. They permit easy interchange of ideas, both vocally and visually, and they allow participants' states of mind to be assessed.

Perhaps most importantly, they encourage the development of personal relationships among the participants. These relationships foster trust and are the grease that keep the wheels turning after everyone has gone home.

Of course, meetings come with a price. They can be notoriously unproductive if not guided by a good chair with a well-conceived agenda. They also frequently involve expensive and time-consuming travel, requiring as much as a three-day investment for a one-day meeting.

Videoconferences

Videoconferencing has long been touted as an inexpensive replacement for meetings. It is inexpensive when compared with the cost of travel, and it is a replacement. However, it lacks the opportunity for personal interactions that characterize a face-to-face gathering and does not permit the often-indispensable sidebar meetings.

Videoconferencing is a good means of keeping a group in contact between meetings. With a well-conceived agenda and a capable chair, videoconferencing can be a productive forum.

Teleconferences

Teleconferences are inexpensive and easy to organize on the spur of the moment. Like videoconferences, they work well as a means of keeping a group informed between face-to-face meetings. With planning, a teleconference can be as productive as a videoconference. Printed and graphical material can be distributed in advance using e-mail, the Internet and fax machines. As always, an agenda and an effective chair are essential to making the teleconference productive.

E-Mail and Web Sites

In the last few years e-mail has become a primary means of communication among study participants. It is timely. It allows the transmission of large amounts of text and graphics in a form that can be immediately manipulated electronically. It allows the simultaneous distribution of material to large groups.

This latter ability is also the Achilles heel of e-mail—it is too easy to send. This results in e-mail being sent because it can be sent. This often leads those so inundated to overlook important material buried in the avalanche. The tendency to overkill with information reaches a zenith in the "reflector" that re-emails everything it receives to the entire reflector membership.

Maintaining information on a web site is an effective alternative means of speedy communication that avoids e-mail's principal drawback by requiring each recipient to actively seek desired information.

Minutes and Action Items

Whatever group forum is used, it is extremely helpful to keep meeting minutes and distribute them to all participants. Documenting arguments, agreements, discussions, decisions, and action items in writing provides a record that supports later study documentation and reduces the potential for misunderstandings. Electronic briefings and documents can also be included in the minutes. Minutes are best distributed by e-mail and should provide for coordination by the recipients.

> ## When Potholes Attack
>
> **When potholes attack and threaten to swallow the scope, credibility, or schedule of a study, the end user of the results must be informed immediately of the difficulties and consulted as to the best course of corrective action. This has the potential to be unpleasant, but it usually beats hiding the bad news.**

Potholes Swallowed My Study

No study runs smoothly. Data arrive late; key people leave the study; there is an unexpected trimming of the schedule or the funding. In most such cases, there are no rules for coping. The answer is to be flexible and make the best of the bad situation in whatever way you can.

Loss of Resources

Studies are frequently disrupted by loss of key personnel. People move to new positions, are assigned new duties, or retire. Often a qualified replacement can be found and brought up to speed quickly. At other times it is possible to redistribute the workload among the remaining panel members. Infrequently, the loss will result in a significant delay in the progress of the work.

The loss of money tends to be more disruptive. Such a loss typically reduces available contractor support or limits travel. The ability to find replacement funds is usually a good indicator of the priority of your work.

Unavoidable Delays

The study director has limited influence over the workload of those from external organizations that are assigned to support the study. The director may, however, have no influence over the work of others whose inputs are needed. Thus it is common to encounter delays in receiving promised products.

Sometimes a personal appeal to the right party can help in such a situation. At other times a management-to-management discussion proves useful. If such appeals are unsuccessful, expect either a slip in the study schedule or the need to compensate with adjustments in the study methodology.

Increased Scope

Often a study will be well underway before the exact extent of the required work can be known. Needless to say, seldom does the scope of work decrease with the additional understanding of the issues that comes with time. The increased understanding generally suggests more work. Typical causes for the increase are the need to consider new alternatives or scenarios, the need to produce a more detailed or higher fidelity analysis, or the need to develop new models or methodology. Since increases in scope are so common, it is wise to try to allow for it in the initial planning for the study.

Internal Conflicts

In any study there will be conflicts due to the divergent viewpoints of participants. It is always desirable to resolve these conflicts fairly and quickly. Unsatisfied parties to a conflict are a disruptive force both within the study and to the presentation and acceptance of study results.

The first attempt at resolution is usually an appeal to logic. Unfortunately, if logic were the answer, it is unlikely that there would be a disagreement. This makes compromise the principal tool for solving disputes. However compromise always has an associated cost. It entails either bending a principle (you give a little, I'll give a little) or an increase in work (we'll satisfy your concerns by analyzing more cases). Because it's hard for most people to bend their principles, guess what generally happens.

Given that analyzing more cases is a practical solution, there is a bright side to the compromise. The additional cases often represent extreme conditions or points of view. Thus when it comes time to brief the study results, even the supporters of the additional cases will not demand that they be briefed for fear of embarrassment. Isn't working with people fun?

External Pressure

It is not at all uncommon to encounter advice and suggestions from "interested parties" not directly involved with the study. Often it is useful, as for example when it clarifies the concerns of the decision maker or exposes you to new points of view. However, it may be no more than meddling motivated by less than altruistic motives. This again offers opportunities to practice your people skills and may again require some form of compromise.

5 Eh?

Answering the Right Questions—The First Time

Old joke: a fellow is walking down the street late at night. He finds a drunk on his hands and knees under a street lamp searching for something.

"What did you lose?" the fellow asks.

"My keys," the drunk replies.

Being a Good Samaritan, the fellow joins the search. After a few minutes he says, "I don't think they're here."

"Of course not," replies the drunk, pointing to the dark side of the street. "I lost them over there."

"Then why are you looking here?" demands the fellow, more than a little put out.

"Because," says the drunk, "This is where the light is."

Update: instead of the drunk, think analyst; instead of the Good Samaritan think General Officer. The analyst comes back to brief the general six months after receiving a tasking. The general says, "You haven't answered my question." The analyst replies, "I know, General. I couldn't answer your question, so I picked one I could answer."

Plain and simple, it is a waste of everyone's time to answer the wrong question. It's far better to shed minimal light on the right question, than to supply the universal answer to the wrong one.

Ensure the Right Question is Answered

Identifying the right question to answer can be daunting (see sidebar). For starters, the asker may not understand the problem well enough to ask the right question. If they do understand, they may be careless in their phrasing, especially if the question is composed hastily in a briefing or other participatory forum. And hey, it happens—they may intentionally ask the wrong question in order to get an answer they like. Of course, we analysts sometimes share the blame. I'm sure at least once an analyst somewhere has misinterpreted clear directions.

So what precautions can we take? There are three things we can do: examine the overall problem and specific question(s) in enough detail to be sure we understand the issues, document our resultant plan of attack in detail, and do everything possible to get the poser of the question(s) to validate the plan.

In the course of this process, you are likely to find that your increased understanding calls for modifying or clarifying the original question to provide useful insights to the questioner. You may even find (as I describe in the sidebar),

True Story

I did a study many years ago for an Air Staff general. His request was made to a general in my command and it eventually worked its way down several layers to a captain who relayed it to me. There was nothing in writing and the verbal statement could basically be paraphrased as 'Different people give me different answers when I ask what is needed to kill target *x*. What does it really take?' Try as I might I was unable to get a more definitive statement.

Three of us worked it for a year. Happy with our answer, but still not sure we had the right question, I took the briefing to the general. With shaking knees (he was renowned for summarily dispatching analysts) I began. Three slides into my presentation, he said, "So you really did answer the question!" He was so delighted that he then proceeded to help me through the rest of my pitch.

that the question must be interpreted before it is answered because the questioner formulates the query in a very general sense. If any of these circumstances arise, the last step of the three-step process becomes crucial.

Of course, not all guidance is oral or off the cuff. In many cases there is written direction in the form of formal documentation. For AoAs there should be a Mission Needs Statement (MNS) from which measures of effectiveness are derived, an Acquisition Decision Memorandum (ADM) framing the questions, and a Program Management Directive (PMD) directing how the study is to be organized. These and other acquisition documents such as the Test and Evaluation Master Plan (TEMP) and the Operations Requirement Document (ORD) will contain useful guidance. There may also be point papers, briefings, documentation from other studies, correspondence, statements of work, etc., which pertain to the question. You should be as familiar as possible with all of this material.

Know Your Constraints

In addition to answering the right question(s), you must identify, understand and document the constraints under which your study must operate. These constraints might be exterior to the study (study direction, politics, schedules), or they may be introduced during the course of the study (study scope, availability of models and databases, etc.). In either event, constraints can go a long way toward defining the study's context and content.

Schedule Constraints

Schedule constraints are among the most common constraints. They frequently reduce the quality of a study by limiting the available analytic options. The study may suffer as a result because it is based on what can be done in the available time, rather than what ought to be done. Analysts can take comfort, however, by realizing that an on-time decision based on a less than perfect analysis is almost always preferable to a late decision based on "perfect" analysis. Thus, the glass is at least half full.

In reality, schedule constraints tend to slip more often than not, but this is basically useless knowledge because you can never count on this happening.

Budget Constraints

Budget constraints are very much like schedule constraints. They limit how much work can be accomplished on a study. Typically budget constraints limit contractor support and travel. Like schedule constraints, budget constraints tend to slip, at least when the study is thought to be important.

Model and Data Availability

Many studies tend to push the analytic envelope because they are attempting to study a completely new problem or new aspects of an old problem. This may require modifying existing models and/or databases or generating new ones. Time and dollars are needed to make this happen. Thus data availability issues are a variation of schedule and budget constraints. However, assessing their impact requires estimating the work necessary to modify or generate the new/updated models and data. This is difficult, even for an experienced analyst. History has shown that all such estimates should be regarded as highly optimistic.

Cost and Performance Thresholds

Cost and/or performance thresholds are levied with the goal of producing an effective and affordable product. Thresholds can be mandated by politics, the service, or DoD. They specify that a solution to a problem must exceed threshold effectiveness for less than a threshold cost. These thresholds imply the use of either iterative or parametric cost and effectiveness methodologies to secure success, since cost and effectiveness are functions of one another.

Security Issues

Many studies must deal with classified information. Receiving, storing, processing and distributing classified information can become costly and time consuming for many studies, especially those dealing with special access information. Potential problems that may arise from working with classified information must be considered early in a study.

> ## Science vs. Art
>
> **Some analysts like everything neatly packaged: use these models and databases, apply this methodology, and answer these questions. Others prefer the challenge of an ill-defined problem, the solution of which allows/requires an imaginative approach, often with new models and methodologies.**
>
> **Call these two types what you will—crank turner vs. innovator, practitioner vs. artist—just don't expect the former to function well in the less structured environment. OR is like every other field: some of us are incompetent, most are competent, but only a few push the envelope.**

Alternative Solutions

Virtually all analyses make comparisons among alternatives. They can be a simple comparison of Alternative *A* with Alternative *B*; they can compare many alternatives to each other; or they may be parametric analyses of one or more alternatives. In any of these situations, we must consider the resources needed to execute the study. In effect we try to thread a needle, examining enough alternatives to be thorough while not drowning in a sea of marginal alternatives. It is not always easy.

Fortunately or otherwise, experience has shown that having too many alternatives—rather than too few—is the greater danger. Some are forced into consideration through politics. Others are initially included because their capabilities or cost are not yet well enough defined to allow their exclusion. When facing a long list of alternatives (or a large trade space), we often seek early purging of the list for verifiable causes. Possible legitimate causes are:

- Non-compliance with study guidance
- Non-compliance with treaties or other national policy
- Unacceptably high cost
- Unacceptably poor performance
- Inability to meet initial/final operational (IOC/FOC) dates

Evidence for the last three shortcomings may come from previous studies, expert judgment, or early cost or effectiveness results. Because these criteria are open to interpretation, a disciplined approach for downsizing the set of alternatives should be followed to forestall second-guessing. This includes documenting the rationale for excluding particular alternatives. For the same reason, it is important to document the alternatives well. At a minimum, the source and nature of the inputs describing every alternative should be included in these descriptions. To minimize overstatement of alternative capabilities, all alternative descriptions should be made available to competing alternative advocates for peer review.

In any comparative or parametric analysis a base case should always be one alternative. The base case usually represents an attainable situation or an existing situation that is funded and operating. The base case offers a yardstick against which to measure any improvements projected for the other alternatives.

New or revised alternatives may need to be included after the analysis is underway. These latecomers are generally conceptual solutions based on immature technologies that are still being fine-tuned. Also, often some alternatives initially discarded for cause may need to be reconsidered as the study progresses. The reconsideration generally arises due to new information.

Write a Study Plan

Every study deserves a written study plan. A study plan serves as a blueprint for what is to be done. At a minimum it will discuss the study's background, goals, scenarios, assumptions, methodology (including any decision criteria), sources of data, resources available for execution, and schedule. The plan may also list participants, describe the study's administrative organization and oversight, discuss the results of previous studies, or anything else of significance. Generally, simple studies will have relatively short plans; complex studies need longer and more detailed plans.

Living Document

Part of every study plan is based on assumptions: the availability of certain people, the existence of needed data, the relative importance of parameters—the list is endless. This means few, if any, study plans can be followed as originally written. We handle this by revising the initial study plan as necessary to reflect new facts and insights as they become available. In recognition of the need to revise, we refer to the study plan as a living document.

Most changes to the study plan are internal to the functioning of the study: a different source of data must be used, an existing model requires unanticipated modifications, or a key person must be replaced. These types of changes are usually transparent to the study's oversight (the requestor, the sponsor, review groups, etc.) and are handled internally. The problem is addressed and the study plan appropriately modified. However, when the change involves study scope or schedule or funding, solutions should be raised to the appropriate oversight level for concurrence before the study plan (and consequently, the study) is modified.

Terms of Reference

A Terms of Reference (TOR) is a simple, frequently used study plan format. It is adequate to support studies of limited scope or serve as a prospectus for larger studies. It will usually be from four to ten pages in length. Different organizations may use other names for similar formats. The name is not important; the content is.

A typical example of TOR content and organization is shown here:

- Title
- Purpose (Why is the study being done?)
- Objective (What do we hope to accomplish?)
- Background
  - Tasking
  - History
  - Previous studies
  - Alternative solutions-if appropriate

- Approval (methodology)
  - Tasks
  - Inputs (models and data bases)
  - Outputs
- Available Resources
  - Personnel
  - Travel and other expenses
- Programmed reviews
- Required documentation
- Schedule

AoA Study Plan

More complex studies require a more comprehensive study plan. The format that follows is taken from the AoA Handbook. AoAs employ cost and operational effectiveness analyses to compare possible materiel solutions to operational deficiencies. This format was designed specifically with the AoA in mind, but it should adapt to other types of studies. A valuable feature of this study plan is that it quickly converts into the final report with the addition of a few sections. This effectively makes the study plan a partial first draft of the final report.

1. Introduction
   1.1. Background
   1.2. Purpose
   1.3. Scope
2. Acquisition Issues
   2.1. Mission Need
   2.2. Scenarios
   2.3. Threats
   2.4. Environment
   2.5. Constraints and Assumptions
3. Alternatives
   3.1. Description of Alternatives
   3.2. Nonviable Alternatives
   3.3. Operations Concepts
4. Determination of Effectiveness Measures
   4.1. Mission Tasks
   4.2. Measures of Effectiveness
   4.3. Measures of Performance
5. Effectiveness Analysis
   5.1. Effectiveness Methodology
   5.2. Models, Simulations, and Data
   5.3. Effectiveness Sensitivity Analysis
6. Cost Analysis
   6.1. Life Cycle Cost Methodology
   6.2. Models and Data
   6.3. Cost Risk Methodology
7. Cost-Effectiveness Comparisons
   7.1. Cost-Effectiveness Methodology and Presentations
   7.2. Cost-Effectiveness Criteria for Screening Alternatives
8. Organization and Management
   8.1. Study Team/Organization
   8.2. AoA Review Process
   8.3. Schedule

A. Acronyms
B. References
C. Other Appendices as Necessary

> **"I have never yet seen any plan which has not been mended by the observations of those who were much inferior in understanding to the person who took the lead in the business."**
>
> ***- Edmund Burke***

Circulate the Study Plan for Feedback

If too many cooks spoil the broth, then too few often leave it lacking. Circulate study plans comment; such exposure improves the plan and helps with buy-in from all concerned parties. Unfortunately in today's busy environment, the feedback may be minimal, late or even non-existent. While this is not what you should hope for, a non-response essentially implies concurrence. It is far better to have implicit concurrence than to be accused later of withholding the plan from examination.

Once the preliminary reviews are completed, the final review should be from the recipient—usually identical to the requestor—of the study. Typically the presentation for this review will be a briefing. It may also be a simple face-to-face discussion or a point paper. Whatever form it takes, try to be sure that you, as the responsible party, prepare the presentation and receive any feedback directly.

6 Methodology to Your Madness I

Taking the Measure of "Goodness"

The problem of how to measure "goodness" is a challenging problem in OR. In part, this is because there is a perception in the community that goodness is in the eye of the beholder—that "it's a matter of opinion." It generally is not. It is a matter of judgment, but it must be informed judgment supported by analysis of quantitative (and occasionally qualitative) measures, including cost, that clearly relate to one or more specific goals. Specific goals and relevant measures make for informed decisions. The absence of either guarantees quite the opposite.

Goals

Unless the goal is to get from here to there as quickly as possible, the speed of an aircraft is a poor measure of aircraft goodness. If on the other hand, we are talking about a combat aircraft and speed enhances aircraft survivability, speed is not the goal—survivability is the goal and speed is a way of improving survivability. But wait! Is survivability the goal, or is an improved rate of killing targets the goal? Improved aircraft survivability affects how fast targets are killed because aircraft that survive fly additional sorties. Or could it be that the real goal is shortening the war? Will I shorten the war if I kill targets faster?

I think you see that goals are hierarchical. So what level of the hierarchy do we measure goodness against? Of course the answer is…it depends. In general, the lower the goal in the hierarchy, the easier it is to assess. Aircraft speed is usually an input to an analysis. Assessing the effect of speed on survivability requires real work by the analyst, especially when this effect must be separated from the effects of radar cross section, vulnerability to defenses, and defensive countermeasures. To further assess the effect of speed on the length of the war sets the analytical bar even higher.

The escalation of analytical difficulty encountered as we move up the hierarchy of goals encourages the analyst to work as near to the bottom of the hierarchy as possible. Even an adventuresome and industrious analyst realizes that time is money and there is always one more problem to tackle. Understandably, however, decision makers want analysis to deal with goals further up the hierarchy. To buy an airplane because it goes fast is an ugly rationale to hang your hat on. To buy it because it can shorten the next war is huge. How do we find the middle ground between these desires? Is there one?

> **"Goodness is the only investment that never fails."**
>
> ***- Henry David Thoreau***
>
> **"No one can be good for long if goodness is not in demand."**
>
> ***- Bertolt Brecht***

Proof

It is the analyst's job to prove—for self and especially the target audience—that the results provide information upon which to base a sound decision. To prove is to convince, nothing more. When we create a proof for a decision maker we must satisfy a sense of propriety. The easiest way to determine what the decision maker thinks is proper is to ask. It is best to get an initial reading at the start of the study through formal direction and seek heading checks as the study progresses. However, the issues may not be understood well enough at the study's start to allow the decision maker to provide that direction. Thus, direction is taken when it is available, and often it changes as the decision maker's understanding of the issues improves.

Should the decision maker fail to provide guidance on a required level of proof, the analyst must decide based on personal experience and available resources and information. In this event, it is a good exercise to try to visualize the content of the final briefing. Put yourself in the shoes of those who will receive the results. Ask the simple question, what information will allow the decision maker to make a sound decision? Try to design the study methodology to provide that information.

It is hard to give general advice, but clearly decisions with inexpensive consequences can be made with a lower standard of proof than decisions with expensive consequences. Put another way, don't spend $20 on a $10 decision, nor only a $10,000 on a $1,000,000,000 decision.

Military Worth

We hinted above at the idea of military worth as a basis for making a decision. Military worth is a concept that we first addressed in our AoA Handbook. Military worth, while not formally defined, carries the connotation of providing a generally accepted military advantage. Thus the aircraft speed mentioned above does not measure military worth, but aircraft survivability and length of the war do.

> ### Tell Me Why
> **Many studies are done with too little funding or too tight a schedule to allow confidence in the results. Yet such constrained analyses may easily lead to bad decisions that delay implementing solutions, reduce solution effectiveness and add substantially to their cost. Why is the support of adequate analysis such a hard sell?**

In our AoA Handbook we identified six categories of military worth. When looking for measures of goodness, these six categories should always be considered first. The importance of selecting measures in these categories increases as one ascends the hierarchy of goals being evaluated:

- Time to accomplish objectives
- Targets placed at risk
- Targets negated
- Level of collateral damage
- Friendly survivors
- Required resources

Time to Accomplish Objectives

This category can provide measures at all levels of the hierarchy. The ultimate objective of war is to win. Winning faster normally means fewer lives lost, less materiel expended, and fewer dollars spent. However, the question "how long?" can also apply to airlifting a given amount of materiel, acquiring a new system, repairing aircraft, or whatever.

Targets Placed at Risk

There are two relevant definitions of targets at risk. In one case a target is at risk if, and only if, we possess a way to damage or kill it. Thus a target will not be at risk if we cannot find it, cannot risk attacking it, or do not have a suitable weapon to use against it. The second definition is favored by the Electronic

Warfare (EW) Partnership. In this definition a target is at risk when an attacking aircraft arrives undamaged at its weapon release point. This definition is used to measure the performance of non-lethal alternatives that improve the lethality of another system. For example, an EW capability, by increasing aircraft survivability, has the potential to increase targets placed at risk.

Targets Negated

Targets negated is an obvious measure that introduces complexities not considered in determining targets at risk. A target negated is usually, but not always, equivalent to a target killed. For example, a SAM site that stays dormant because of fear of being killed has certainly been negated. Using targets negated may involve modeling the interaction of munition and target; delivery system survivability; intelligence, surveillance, and reconnaissance (ISR) systems; airlift capacity; basing considerations; the physical environment; etc.

Level of Collateral Damage

For humanitarian and political reasons, there is always concern about the level of collateral damage—human and property—caused by attacking a target. Limiting collateral damage has taken on more importance as military targets have been intentionally integrated into civilian surroundings to deter attack.

Friendly Survivors

There are studies that consider alternatives that are both non-lethal and non-lethality-enhancing. Two examples are the Combat Survivor Evader Locator (CSEL) aircrew survival radio AoA and the Joint Precision Approach and Landing System (JPALS) AoA. In such cases, the number of "survivors" associated with each alternative may best measure military worth. For CSEL the obvious measure is the number of downed aircrew members recovered. For JPALS, the measure could be the percent of successful landings achieved.

Required Resources

The advantages of reducing the number of personnel or the skills or materiel they need to perform a function are easily understood: dollar savings, freeing of personnel for other functions, an increased "tooth to tail ratio," reduced training requirements, reduced deployment footprint, etc. Each of these consequences provides a clear military advantage.

Cost and Cost-Effectiveness

No analyst can afford to ignore cost (pun intended). Cost is most certainly a measure of goodness, but a measure that must eventually be combined with measures of effectiveness to be fairly assessed. We have devoted Chapter 12 to discussing cost analysis. The melding of cost and effectiveness, i.e., cost-effectiveness, is discussed at the end of this chapter.

Tasks

Goals are reached through performing tasks. Tasks should be general statements of what must be done to achieve the goals. For example, consider the goal of improving a developmental fighter aircraft's survivability in an IR-guided SAM environment. Some tasks that could be performed to achieve this are:

- Increasing aircraft hardness
- Increasing aircraft countermeasures against IR missiles
- Delaying/denying detection by the air defense

For each of these tasks, there will be specific alternative ways of performing the tasks. A partial list of generic alternatives for this example could contain the following:

- Reducing aircraft radar signature (to hinder detection)
- Reducing aircraft IR signature (to hinder detection)
- Increasing functional redundancy on the aircraft (to reduce vulnerable area)
- Hardening the aircraft with armor (to reduce vulnerable area)
- Improving missile warning capabilities (to improve use of IR countermeasures)
- Increasing number of IR countermeasures (to increase the number of missiles defended against)

An example of a specific alternative would be adding functional redundancy by building the aircraft with two engines instead of one. Other critical system elements could also be candidates for duplication. The obvious question (at least for an analyst) is how effective is each of these alternatives? For this we must discuss measures of effectiveness (MOEs). MOEs (when evaluated) quantitatively—and occasionally qualitatively— describe how well tasks are performed. For this reason MOEs must be derived from the tasks. The schematic methodology of MOE development is depicted in Figure 6-1.

Measures of Effectiveness (MOEs)

Details of measuring proficiency in performing a task are contained in the MOEs associated with the task. Each alternative is evaluated against each MOE, and the results are used to compare the alternatives.

Our experience supporting AoAs has taught us that developing a good set of MOEs is usually a harrowing business. This follows from the following observed human frailties:

- People think about MOEs independently of the tasks to be supported—to the point that they decide on MOEs before the tasks are defined ("We always use that as an MOE!" Can you spell irrelevant?)
- People think that if two MOEs are good, ten are better ("You can't have too many MOEs!" Wanna bet?)
- People think that if a possible MOE can be calculated, it should be ("But it's so easy to determine kumquats per quart!" And your point is?)

In general:

- MOEs are quantitative (e.g., how many targets are held at risk?), but they may be qualitative (e.g., does the solution provide a day-night capability?)
- Each MOE supports at least one task and each task will be supported by at least one MOE
- MOEs may support other MOEs as well as tasks
- MOEs must be independent of the nature of the alternatives, as all alternatives are evaluated using all MOEs
- MOEs should not be strongly correlated with one another (to avoid overemphasizing particular aspects of the alternatives)
- MOEs are MOEs only in relation to a task (no quantity is inherently an MOE)
- MOEs are often supported by one or more measures of performance (MOPs)

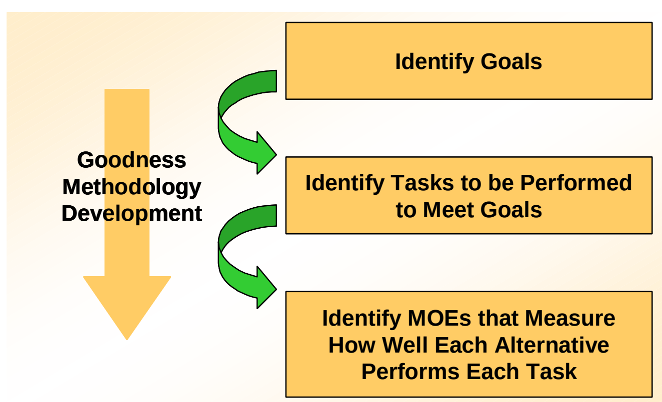

Figure 6-1 : Tasks are Derived from Goals; MOEs are Derived from Tasks

No! No! No!

Cost is never an MOE. Cost does not measure effectiveness. If it did we would not talk about cost-effectiveness, but only effectiveness.

Cost is cost is cost.

Thank you.

Ideally, MOEs should normally represent raw quantities—numbers of something—or frequencies of occurrence. Attempts to disguise these quantities through a mathematical transformation (for example, through normalization), no matter how well meaning, reduces the information content and may be regarded as "tampering with the evidence." This same reasoning applies to the use of MOEs defined as ratios. A ratio essentially "hides" both quantities in the ratio.

Results from MOEs not only make it possible to compare alternatives, they also can be used to investigate performance sensitivities to variations of key assumptions and input values. Such analyses help define requirements (ORD requirements in the case of an AoA). These results can also be used to investigate the robustness (stability of performance) of alternatives whose defining parameters may vary significantly due to uncertainties.

Measures of Performance (MOPs)

A measure of performance is basically an input to the determination of an MOE value for an alternative. As such it is characteristic of the alternative in question. An MOP is typically a quantitative measure of a system characteristic (e.g., range, velocity, mass, scan rate, weapon load-out, etc.). For AoAs, MOPs may be directly or indirectly reflected in system performance parameters in the ORD. MOPs and the methodology for evaluating their impact on MOEs frequently help determine ORD requirements.

Cost-Effectiveness Comparisons

The ultimate measure of goodness for an analyst (and the decision maker) is cost-effectiveness. Cost-effectiveness comparisons simultaneously consider alternatives' cost and effectiveness.

As consumers, we are all familiar with the concept of cost-effectiveness. Whether buying laundry detergent or a home, we collect data on cost and make

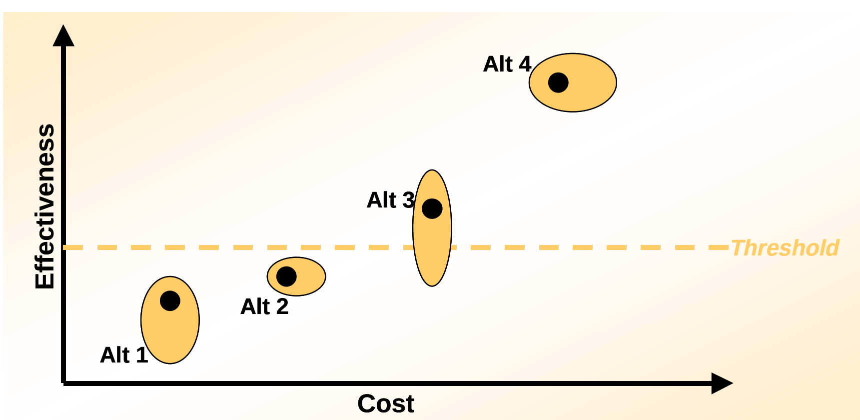

Figure 6-2 : Dilemma 1: Is the Increase in Effectiveness Worth the Increase in Cost?

assessments on how well the alternatives will meet our needs (how "effective" they are). With data in hand, we make our comparisons and select a winner. In OR the process is essentially the same, although usually more formal.

While this kind of cost-effectiveness "analysis" is quite sensible, from experience we are also aware of its inherent difficulties: the need to determine if additional effectiveness is worth additional cost and the need to assess the relative values ("weights") of different measures of effectiveness. The first of these problems is illustrated in Figure 6-2 and the second in Figure 6-3.

From Figure 6-2 we could safely conclude that we would not select Alternative *1* or *2*, but the issue is not clear for Alternative *3* and Alternative *4*. Alternative *3* and Alternative *4* will be chosen if the increase in effectiveness is judged to be worth the cost. The decision may be somewhat easier if there is a minimum acceptable threshold of effectiveness (as for example, from an ORD). However, the threshold may be exceeded by more than one alternative as illustrated, and having a threshold does not eliminate the option of "buying nothing" if all alternatives meeting the threshold are deemed too costly.

Figure 6-3 shows the second type of dilemma. In this illustration, if MOEs *a* through *c* have equal weight, there is little to differentiate among the choices. If on the other hand the MOEs are not weighted equally, then the three may differ substantially in overall effectiveness. The question is "Who makes the judgment?" We would leave that decision to the decision maker. If we assign weights to the individual MOEs, we are effectively taking the final decision away from the decision maker. That is not our job.

In the following section we will focus on the cost-effectiveness comparison process, what it should and shouldn't be and how to make sense of it. Our guiding principle will be that the one and only goal of the process is to identify the most promising candidate(s) for consideration by decision makers (this does not mean that we cannot present additional qualitative information that bears on the decision).

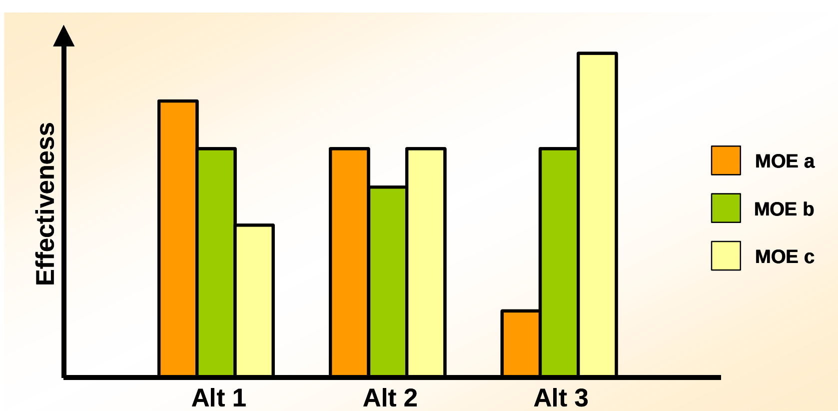

Figure 6-3 : Dilemma 2: Do These Three Alternatives Really Have Significant Differences in Overall Effectiveness?

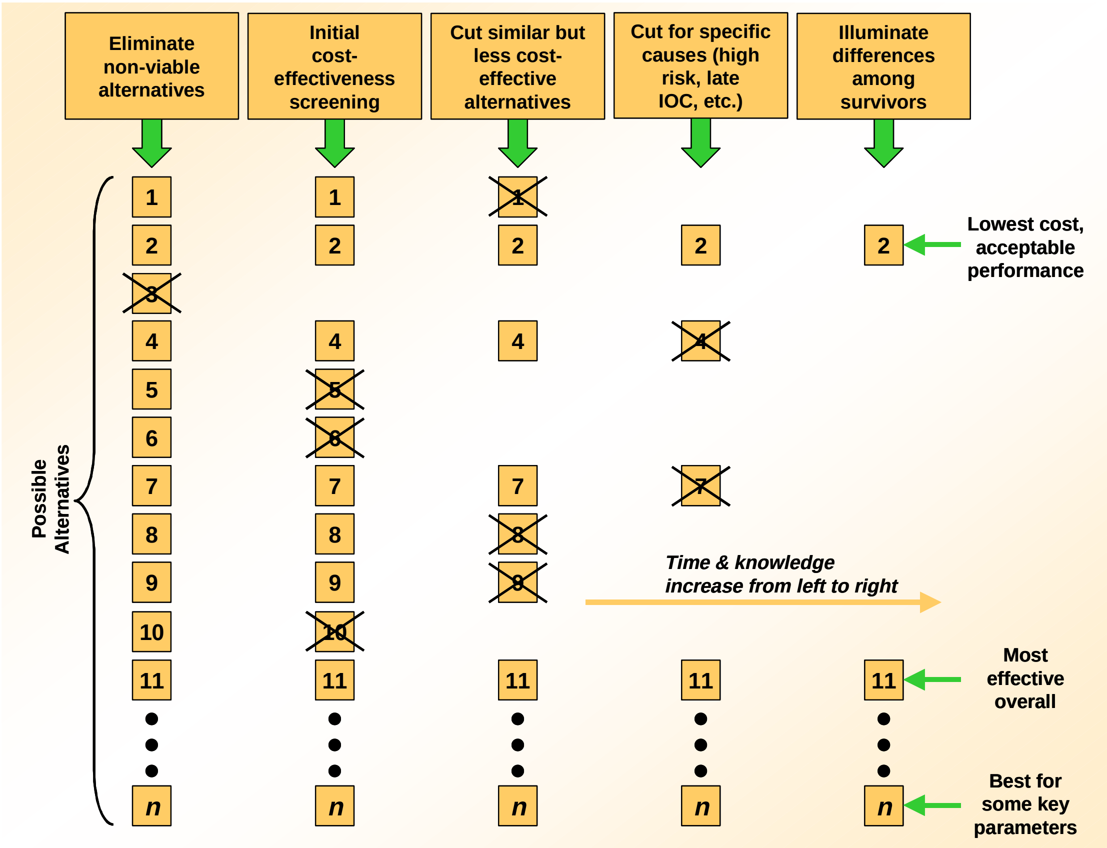

Figure 6-4 : An Example of the Art of Eliminating Alternatives in an Initial AoA (following Milestone I)

Equal Effectiveness or Equal Cost?

Equal Effectiveness

Cost-effectiveness comparisons are made most easily if all alternatives are configured to produce equal effectiveness. The analysis is then reduced to a simple cost comparison. Unfortunately, equal effectiveness is usually difficult—maybe impossible—to define because of the number and complexity of the analysis issues.

For example, suppose an analysis is comparing alternative munitions' effectiveness against a class of targets. We might propose equal effectiveness as killing a fixed percentage of the targets in a fixed time. While this may sound reasonable, it raises questions:

- What if some munitions require more sorties to meet the goal than the force can generate?
- What if the delivery of the different types of munitions results in significantly different aircraft attrition rates?
- What if the delivery of the different types of munitions results in different rates of kill of other targets because of a shift of resources?

Most often, these and other significant "what ifs" will arise and erase any perception of equal effectiveness.

Equal Cost

An alternative cost-effectiveness approach is the equal cost approach. In this instance, a straightforward comparison of alternatives is possible because all alternatives are designed with equal cost. In general, however, this is as difficult to implement as equal effectiveness.

We can see this using the same goal proposed for the equal effectiveness discussion: killing a class of targets in theater. We will assume that it is possible to set a fixed value for life cycle cost and calculate the number of munitions bought for each alternative based on this value. Unfortunately, we have to face "what ifs" similar to those raised in the equal effectiveness case. This does not mean that meaningful equal cost comparisons can never be made, only that it is rarely practicable.

Effectiveness vs. Cost

The obvious alternative to the equal effectiveness or equal cost ideal is a scatter plot of effectiveness versus cost (or vise versa) as in Figure 6-2. As we have implied, however, this seldom gives an unambiguous answer. Worse, it implies that the dilemma illustrated in Figure 6-3 has been solved and effectiveness has been successfully reduced to a single number through aggregating—a practice we strongly discourage unless it meets the criteria discussed at the end of this chapter.

Remember: there is no requirement for an analysis to identify a single most cost-effective solution.

The Art of Eliminating Alternatives

Figure 6-4 suggests how an analysis reduces the original set of alternatives to a small number of serious contenders. There is no formula for doing this. It is an art whose practice benefits from experience, and each analysis must adapt its methods to circumstances. One constant, however, is the need to document the reasons each alternative is eliminated. This documentation will serve you very well in the event the results of the analysis become controversial.

In all analyses, the study team's understanding of the issues and the techniques to deal with them increases as the study progresses. The same is the true for alternatives, especially when many alternatives are initially poorly defined concepts. As the analysis progresses, these concepts are often reengineered to reflect better understanding of requirements, technologies, threats and scenarios. Improved performance and lower cost usually accompany these changes—thus alternative cost and effectiveness are moving targets.

The uncertainty can be limited by setting a cutoff date for concept redefinition. However, remember that the charter of most analyses is to find the most cost-effective alternatives, not the most cost-effective alternatives defined up to an arbitrary time. Thus, the analysis should revisit discarded alternatives from time to time when new information promises to significantly increase their attractiveness. This is most important when a large number of concepts have been screened early in the analysis.

Non-Viable Alternatives

The first screening is to eliminate non-viable alternatives, alternatives that do not adhere to the ground rules of the study. You should identify these alternatives in the study plan and indicate the reasons for their elimination. Criteria defining non-viability are frequently defined in the written guidance for the analysis. These criteria most often reflect political considerations: the environment, world opinion, treaty compliance, desired IOC, etc.

In special circumstances, one or more non-viable alternatives may be carried forward to provide desirable reference points. It should be noted that the baseline (usually the status quo) is almost always retained throughout the study as a reference point, independent of any shortcomings it may have.

Preliminary Screening

When a preliminary screening is necessary, it is usually done with limited data derived for alternatives whose definitions are still in transition. This suggests erring on the conservative side by giving alternatives the benefit of any doubt.

The exact screening criteria will depend upon available analysis resources, the number of alternatives to be carried forward, the perceived uncertainty in cost and effectiveness estimates, and a host of other factors such as similarity of alternatives, advocacy for alternatives, and technology maturity. Other factors that might be considered are sensitivity of system performance to key assumptions, vulnerability to countermeasures, flexibility in future scenarios, contributions to longer-term goals, reliability and maintainability, and time phasing of resource requirements.

Later Screening

As the study progresses and more reliable cost and effectiveness data become available, there will be opportunities to do additional ad hoc screening. This is typically done on a case-by-case basis using any appropriate criteria. For example, an alternative may be demonstrated to be more costly or less effective than the others. If it has no redeeming qualities, it can be removed. Another system could be very sensitive to a key parameter, indicating excessive risk in performance…it may go as well.

Final Selection

There comes a time when the remaining alternatives under consideration all have positive attributes which makes them attractive in some way (think of a scatter plot similar to Figure 6-2). They are true contenders. The next step is to find a way to clearly state for the decision maker the advantages and disadvantages of each. The more clearly this story is told, the easier it becomes to understand the differences. Even with all cost and effectiveness results in hand, it is not unusual for this final story to take a week or more of intense effort to develop. Rational thinking is critical during this period.

In some cases this final assessment may point to a single "recommended winner." In others, no such clear-cut conclusion emerges. In either event, the decision maker will have the best available information and understanding of the alternatives that the study can provide.

Cost-Effectiveness Dos and Don'ts

Do Sensitivity Analysis

Alternatives whose performance is stable over a range of conditions are more adaptable than those lacking such stability. Alternatives are typically defined with certain assumptions made about their performance parameters: weight, volume, power consumption, speed, accuracy, impact angle, etc.—whatever is appropriate to each alternative. These "monolithic" alternatives are then assessed against threats and scenarios under a set of assumptions. This provides very specific cost and performance estimates, but does little to assess

Your Baby's Ugly

An interesting aspect of all alternatives is that they have "parents" whose fondest vision is seeing their baby grow to up to be a star. It can be traumatic telling them that their baby is ugly. Doing so provides a wonderful opportunity for you to practice your people skills.

Table 6-1 : Cost Effectiveness Matrix

MOE and LCC Summary									LCC $(M)
MTs:	1 Air Superiority			2 Supportability		3 Interoperability			
MOEs:	1-1	1-2	1-3	2-1	2-2	3-1	3-2	3-3	
Alt 1									
Alt 2									
Alt 3									
Alt 4									

the stability of an alternative's performance to changes in system parameters, threats, scenarios, and assumptions.

Stability can only be investigated through sensitivity analyses in which the most likely critical parameters are varied. This form of parametric analysis can often reveal strengths and flaws in an alternative's performance that are valuable in making decisions. Sensitivity analysis should be performed whenever time and resources allow. Of course, it is always necessary to balance the amount of sensitivity analysis against its potential value and the available resources.

Provide the Basic Cost and Effectiveness Data

Provide basic life cycle cost and MOE effectiveness data for all candidate alternatives that have been analyzed. Table 6-1 shows a straightforward format for presentation. By their nature, these data are fundamental to understanding the logic of any additional winnowing of alternatives.

> **And Let Me Say Again...**
>
> **Devise an unambiguous, simple to understand format for the presentation of the comparison of surviving alternatives. It's what you are being paid to do. Yes, you may be lucky. It may be easy. Experience promises otherwise. Give yourself time to be creative. Doing it well can be the difference between being a hero or a goat come the briefing.**

Avoid Using Ratios for Comparisons

Ratios—cost/kill, kills/sortie, etc.—are frequently proposed for comparing alternatives. Unfortunately, ratios can be misleading because they frequently hide necessary information.

As an example, suppose that one alternative kills 0.01 target per sortie and a second alternative kills 0.1 target per sortie. The second alternative is ten times better than the first. That sounds significant, but is it? The truth is, we can't tell from the ratio alone. If there are 10 targets to be killed, the answer is likely to be a resounding yes—100 sorties may be acceptable, but probably not 1000. However, if there are 1,000 targets to be killed, the answer is almost certainly no, for we are looking at very large numbers of sorties even for the better alternative. By using the ratio instead of the numbers of sorties required, there has been a loss of understanding without any corresponding gain. Essentially, this use of ratios does not meet our criteria for aggregating results (discussed later).

Avoid Weighting MOEs

In the roll up process, a frequent issue is whether to weight the MOEs. Weighting assigns different factors (weights) to different MOEs. It is a seductive idea: clearly not all MOEs are created equal. A difficulty with weighting, however, is that your weights may not be a decision maker's weights. Why should they be? Your backgrounds are not the same and your perspectives are not the same. By weighting, you are proclaiming your judgment to be superior to that of the decision maker.

We strongly discourage weighting. Almost invariably, weighting is an attempt (conscious or otherwise) to avoid thinking through alternative methods of presenting the results in a clear manner. Good alternative presentations almost always can be found. Take the time to look for them.

Don't Roll-Up (Aggregate) the Results Unless…

Once the MOE evaluations have been presented, it may also make sense to "roll up" these results. Rolling up results describes any process that aggregates results for individual alternatives. A roll up allows comparing the alternatives using a smaller number of measures. The advantage of having a smaller number of measures carries the obvious disadvantage that information, and along with it potential insight, is lost in the roll up process. Our advice is to aggregate when the rationale to do it is sound. This means when:

- The aggregation arises naturally from relationships among the MOEs
- The significance of the aggregates is clear
- The aggregates tell a clearer story than the individual MOEs

These are difficult criteria to meet, but nothing less makes good sense. The message is: don't aggregate just to aggregate.

7 Methodology to Your Madness II

Completing Your Methodology

Every study must have a methodology. This methodology defines what analysis is to be done and, at least in broad terms, how to do it. In Chapter 5 we discussed the ins and outs of answering the right questions. In Chapter 6 we addressed in somewhat theoretical terms how to compare alternative solutions to identified needs through identifying tasks and corresponding MOEs. In this chapter the discussion shifts to the calculation of the MOEs. For those of you with a short memory, Figure 7-1 provides a map of where we've been and where we are going.

Defining the Calculation Methodology

While we talked at length in the last chapter about developing MOEs to measure how well tasks were accomplished, we said nothing about calculating them. From a practical standpoint, the ability to calculate the MOEs is critical to designing a viable analysis methodology. In most instances, of course, the ability to calculate an MOE is considered as part and parcel of the selection process. The primary questions we ask are:

- Do I have (or can I create) models that will allow me to calculate the quantity in question?
- Can I beg, borrow or steal the data I will need to execute the models once I have them?
- Does my schedule afford me the time to do what must be done?
- Do I have access to the personnel to do what must be done and can I afford them?

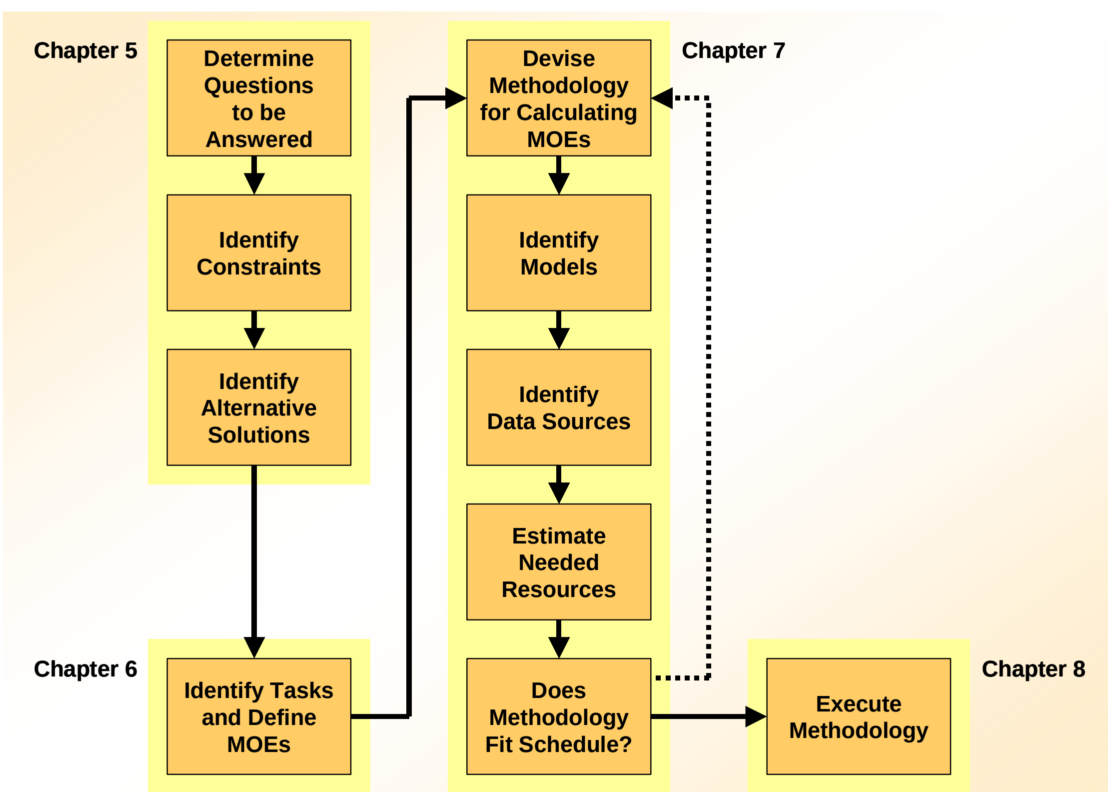

Figure 7-1 : A Map of Analysis Methodology with Chapter References

Pieces Is Pieces

Perhaps the most brilliant man I had the pleasure of working with came to the US from Germany after WW II via Project Paperclip. He was a natural-born teacher and on the day he retired he gave a talk to the assembled office. His message was simple: ***you solve complex problems by breaking them into small, manageable pieces and solving the pieces.***

In his opinion that was the most valuable advice he could leave with us. It's obvious, but it is so often ignored. Heed it.

I can tell you from experience that the most common initial answer given to each of these questions is maybe. You think you can, but…you won't know until things get a little further along (there is quite a bit of art to analysis). Fortunately, for the experienced analyst the final answer is usually yes.

Final negative answers frequently cause big problems in the form of schedule slips and cost overruns. There are, of course, steps you can take to reduce the risk of negative answers.

Do Your Homework

The first step in the planning stages of a study is to review the documentation of past related studies. This frequently turns out to be unproductive, but having done it you can confidently tell the decision makers that you have checked that box (and they may ask). However, when the review is productive, good things happen:

- You become party to others' thinking, learning from their successes and mistakes
- You become party to others' analysis results, acquiring insights into your problem long before you have your own results to digest
- You identify the authors who did the work; often they or their organizations can be important sources of up-to-date information concerning your area of interest

The Defense Technical Information Center (DTIC) makes identifying and acquiring potentially useful reports easier. DTIC archives thousands of defense-related reports every year, cataloging them in a searchable database and providing copies of reports to qualified DTIC users.

Unfortunately, much useful material is never sent to DTIC. Thus, it may be prudent to avail yourself of other government and commercial databases, for example, the National Technical Information System (NTIS). Other sources for this type of information are model managers, web sites, program offices, etc.

Fourth Principle of OR

Model selection is made only after knowing what is to be calculated.

Every Problem a Nail

The problem of selecting the model(s) to be used in calculating an MOE may be trivial. For example, there may only be a single model with the requisite capabilities. On the other hand there are often several potential models to choose from, each with its own merits and demerits. Faced with a choice, the analyst is tempted to select a model not on its suitability, but on its familiarity or ease of access. This tendency is known as the "every problem a nail" syndrome—if your only tool is a hammer, then every problem looks like a nail.

Human nature is to stick with what we know. This is often expressed as "better the devil you know than the devil you don't." If you think it's traumatic giving up that old pair of sneakers, try abandoning your favorite model. Feel that shudder? But if we've got a model we know and believe in, why change? HELLO! BECAUSE IT MAY NOT BE THE BEST MODEL FOR THE JOB!

A poor choice of models can call the entire study into question. Always be prepared to answer honestly why each model was chosen.

Fidelity of Calculation

I studied astrophysics in graduate school and we often used the expression astrophysical accuracy. This was generally taken to mean to the nearest order of magnitude (power of ten). While occasionally military OR should claim no better than astrophysical accuracy, we generally feel more comfortable believing our errors are in the 10-20% range.

The truth is, we seldom have good estimates of our real errors when a complex model generates the results—model statistics not withstanding. If we translate this to practical analysis, most analysts would concede that our results are usually relative. Thus when System *A* rates a normalized 0.8 and system *B* a 0.6, the most we usually would claim is that *A* may be significantly better than *B* (many exceptions to this "rule" occur when dealing with models that have been well calibrated through testing or practical experience; physics and engineering models often fall into this category).

Because of this frequent uncertainty in our answers, we want to be careful not to "over calculate" parts of the answers. For example, if we are adding or comparing two numbers it makes no sense to work hard to estimate one number to the nearest tenth because we can, and estimate the second to the nearest 10 because we can't do better. Use good judgment in determining how good your calculations should be.

On Eating Well

I worked for years with a Texas Aggie who was a wellhead of aphorisms. One of his favorites was the simple admonishment: don't strain at a gnat and swallow an elephant.

This covers a multitude of situations, from reducing a 1% error while ignoring a 20% error elsewhere to modeling in meticulous detail—not because it is needed—but because it can be done.

Balancing Resources and Requirements

It is the nature of analysts to be optimists. They habitually promise more than they can deliver. They may have Seabee mentality—the difficult we do immediately, the impossible takes a little longer—but may lack the corresponding ability to produce. This attitude can be infectious and tends to be positive force—but not when developing the methodology for a study. Such optimism almost guarantees a methodology that, while usually executable, will not be executable with the resources at hand. The result will be schedule slippage.

I call here upon the First Principle of OR. It's possible to do a study fast and cheap, fast and good, or cheap and good. You cannot do a study fast, cheap, and good. Also recall the old saw: If it seems too good to be true, it probably is.

The reason for this optimism is simple: despite all their personal experiences to the contrary, analysts assume that they are in control and that none of the following will occur:

- Temporary or permanent loss of key personnel during the study
- Unexplained model results, which usually mean unanticipated time and effort spent in model debugging
- Unexpected modeling results, which usually result in modifications to the methodology

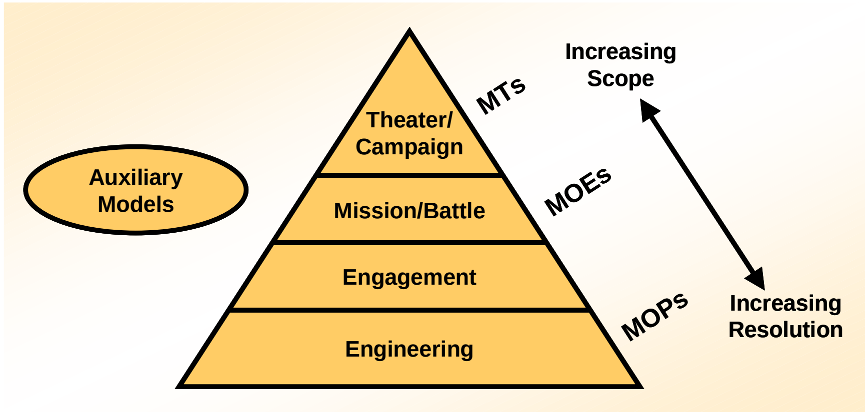

Figure 7-2 : The Model Pyramid

- Models that run more slowly than anticipated
- High-level redirection of the study (to include expanding the study's boundaries)
- Delays in receiving data (or worse, unavailability of data)
- Greater problem complexity than anticipated
- Self-serving proponents of alternatives who obstruct the analysis

The list goes on. What can happen that reduces the need for resources? In truth, there is not much. All this leads to my rule of thumb: whatever period the analyst says is needed to complete an analysis, double it. And for luck, keep your fingers crossed.

Minimize Single Points of Failure

Any complex undertaking should strive to minimize single points of failure. Any statement of the form "we're in big trouble if…." indicates such a point.

- We're in big trouble if Carol quits (is pulled off the study, gets promoted…)
- We're in big trouble if the model isn't finished (runs longer than expected…)
- We're in big trouble if the data aren't available (are late, aren't blessed…)

You get the idea. It is never possible to avoid all single point failures, but careful planning in the form of contingency planning can reduce their number and the risks they represent. Get Carol a backup; know where to get a faster computer (or model); identify alternative sources of critical data.

Selecting Models and Data

Analysis Pyramid

For years the military analysis community has based its model taxonomy around the imaginary hierarchical pyramid shown in Figure 7-2.

The base of this pyramid rests on "engineering" models, models that are generally considered to be high fidelity models for calculating engineering values.

One step up from the base of engineering models are the engagement models. These are generally considered to model one-on-one and few-on-few engagements of weapon systems (e.g., aircraft vs. aircraft).

Higher in the hierarchy are the many-on-many mission-level models that pit force-on-force.

At the top are the campaign models that play all the forces in a theater against one another.

This pyramidal representation is useful in visualizing what model capabilities are available—except that there are a number of models that do not really fit this classification scheme. Route planning models, scheduling models, reconnaissance models, certain optimizing models, etc., are not engineering models and they are certainly not engagement, mission, or campaign models. They are, however, important models in many analyses, often supporting the engagement and higher-level models. Until we revise the classification scheme, I suggest that you think of these models as "auxiliary" models sitting off to the side of the pyramid (as pictured).

Standard Analysis Toolkit Models

The M&S world is always in a state of flux. Existing models are constantly being updated to deal with new realities, and new models are being written to take advantage of new methodologies or to model new capabilities or technologies.

To introduce order to this ongoing disorder, the Air Force has made a concerted effort to reduce the number of models in use and to adequately manage the configurations of the "surviving" models. This effort has resulted in the development of the Air Force Suite of Models for Analysis (AFSOM-A), originally called the Standard AF Analysis Toolkit.

Initially, a limited number of existing engagement, mission, and campaign models were grand fathered into this toolkit. Additional models are being added under a structured evaluation process that ensures inclusion of needed capabilities while avoiding duplication of existing capabilities. Some of these additions will be auxiliary models.

Use of toolkit models is strongly encouraged. Standardized models will have a history of verification and validation (V&V) that will buttress their selection. Thus their use will reduce "model sniping"—the practice of calling an analysis into question by attacking the model.

Ideally, the standard toolkit model functionality will transition as appropriate to JMASS, JWARS or JSIMS—the so-called J-models. However, for the immediate future the toolkit will survive and grow.

There is no proscription against using non-toolkit models, but use of a non-toolkit model when a suitable toolkit model is available will, at a minimum, generally (and rightly) raise eyebrows.

Model and Data Availability and Suitability

The models chosen for the analysis must be compatible with the MOEs being calculated, e.g., you'll need a campaign model if length of the war is to be calculated. Whatever models are chosen, suitable data to support them must be available.

Selecting Threats and Scenarios

The selection of threats and scenarios is critical to many analyses in two ways. First, the threat (scenario, hostile concept of operations, and data) is the yardstick against which friendly alternatives are measured. If the threat isn't appropriate (like analyzing anti-submarine warfare in the Sahara), then conclusions drawn from the analysis will be meaningless at best and misleading at worst.

Second, choice of threats can significantly affect your overall credibility. Often the easiest way to bias the outcome (intentionally or not) is to "cook" the threat. Subtle changes in the adversary's concept of operations can make all the difference between buying F-15s or F-22s. Many a Government Accounting Office (GAO) or Inspector General

(IG) investigation has focused on the selection of the threat. Because of this, your threats are likely to be one of the most scrutinized aspects of your study.

The DoD 5000 regulations, AF 10-601, and DIA 55-3 all insist that major studies impacting requirements or acquisition decisions use Defense Intelligence Agency (DIA) validated or approved threats. This is a case where plagiarism isn't only allowed: it's encouraged. You are expected to use existing threats if you can—expanding, changing, and supplementing them only when you must. Fortunately, there are places to go for threats that have the equivalent of the "Good Housekeeping" seal of approval.

Your retailer for intelligence is your MAJCOM intelligence office or AFMC Director of Intelligence. Their first job is to work with you to find existing, appropriate, DIA approved threats in the intelligence community's virtual library. The Multi-Spectral Force Deployment (MSFD) products, which are produced by the National Air Intelligence Center (NAIC), are a good first product line to look at.

If you need more, their second job is to help you document what you need in "intel-speak," and then work with the intelligence community to produce it.

This is often easier said than done, especially in terms of meeting your schedule. The earlier you ask, the more likely you will get the inputs you need, when you need them (if you have money to throw their way, that can help speed up the process on occasion).

Once any new material is produced, it must be vetted across the intelligence community so that they will stand behind the threat you used. This is true even if you think the changes you have made are minor. This path runs from your MAJCOM intelligence office through the 497th IG near Washington, DC. They decide who needs to review the threat, and help insure it gets done. If your threat is complex, controversial, or significantly stretches the baseline into new territory, this may not be a quick process. Involve your intelligence representative early, and allow as much time as you can for this review process. Usually, you can work this concurrently with other aspects of your analysis.

Practical Example

Consulting an expert on a particular input I was told that he could only guess at the value. My reply, of course, was that his guess was better than my guess. Who says those experts don't earn their pay?

- Maj Rich Roberts

Identify Data Sources

There are few things worse for an analyst than to have the credibility of the models challenged at a briefing; having your data challenged comes close. This challenge may be aimed at specific values of one or more important parameters, but it is more likely to be aimed at your data sources. The discussion usually goes like this: "Where did you get the values for (fill in as appropriate)?" Answer: "They came from (fill in as appropriate). Comeback, dripping with incredulity: "Don't you know that (fill in as appropriate) is the accepted source for those numbers?" Answer: "Doh!"

Models need data, but not just any data. Therefore when devising the methodology it is of critical importance to ensure that adequate data to support the chosen models will be available. These data may:

- Come from one or more standard databases
- Be taken from authoritative, non-standard databases

- Be provided by an authoritative individual or organization
- Be derived from other model results
- Be deduced or invented, based on best available information (as may be the case for parametric input values)

The trouble is that plans are just plans. There are often mines in the data fields. For example:

- A data source vanishes ("I know what I promised, but my Colonel says…")
- The promised data are late ("Sorry, we didn't backup our hard drive.")
- Conditions are imposed on the use of data ("You can have the data, but don't tell anyone where they came from!")
- The data must be transformed (or interpreted, or checked, or blessed, or whatever) before they can be used
- Your data are trashed in advance ("The General already said anything based on those numbers would be unbelievable!")

The message is: like a good scout, be prepared. Data woes are the number one bane of the analyst. Do your data planning carefully and cultivate your sources.

Hard Truth

Good judgment comes from experience. Experience comes from bad judgment.

8 Slice and Dice

Executing the Methodology

Theseus came upon the Minotaur asleep and fell upon him, pinning him to the ground; and with his fists—he had no other weapon—he battered the monster to death.

Unfortunately for Theseus, killing the Minotaur was not the end of his labors; he still had to find his way out of the labyrinth. Fortunately, Theseus had learned the secret of the labyrinth from Daedalus, the designer of the impossible maze: he tied a ball of string at the entrance and carried it, winding and unwinding it as he went. To exit, Theseus simply retraced his path; Theseus escaped because he had a complete record of where he had been.

Perhaps the most important thing an analyst can do during the execution of the methodology is to keep accurate documentation of what was done. Things that should be recorded:

- Keep a journal. Write down what you did, when you did it, why you did it, and whether it worked or not.
- Keep an archive of data files, intermediate versions of code, and benchmark results. Back up your hard drive. There are two kinds of computer users; those who have lost data when their hard drive crashed and those who will lose data when their hard drive crashes.
- Keep a list of people you worked with and what they provided you.
- Document your code as you go. It is difficult to get motivated to go back and document 2000 lines of code that you finally got working. It's even more difficult to document 2000 lines of code two months later when you get around to it.

The Four Phases of Execution

In all the studies I've been involved with, I've found that the execution of the methodology can be divided into four generalized phases: the planning phase, the data collection/modeling phase, the "uh-oh" phase, and the production phase.

In the planning phase, you determine how many runs you need to do, what the runs will be, and how you will get the results you need from the model's output data. The data collection/modeling phase is where you will collect and/or develop the resources to conduct the planned runs. The uh-oh phase is where you find out the data wasn't what you expected, the model doesn't do what you thought, etc. Finally, in the production phase, you produce the information that is used to prepare your results.

The ultimate goal in executing the methodology is to ensure the uh-oh phase occurs *before* the production phase.

> **Analysts!**
>
> **I apologize profusely for letting this chapter's author sully this handbook by mentioning actual analysis techniques, but he insisted it was necessary. I swear that's all they think of. Design of experiments and response surface methodology (discussed later) are techniques allowing an analyst to extract a specific level of information about a system with the minimum number of simulation runs.**

Do You Really Need All Those Runs? (Planning Phase)

Consider the following. You have a stochastic model with five inputs you've decided to vary for your study. The study director wants to use four settings for each of the inputs to get some decent curves and make some great charts.

Each run takes 20 minutes and produces about 500 kilobytes of results. Typically, this model requires 15 replications at each setting to account for variability between runs. After 213 days of continuous computer time (over 15,000 runs) you will be able to begin analysis on your 7.6 gigabytes of results. Somehow, you need to reduce the number of runs.

You now face a dilemma: you can't live with six months of runs, but you don't want to throw out any runs that may be needed. The key word here is "needed." If you use a design of experiments (DOE) methodology, you'll be able to focus on the runs you need. The following lists some things to think about:

- *Do you think the model behaves linearly with respect to your inputs?* It only takes two points to specify a linear relationship; putting more points on the same line doesn't add any information. Set up a two-level design of experiments to get all the linear effects and 2-way interactions; throw in the point at the center of your design and you can test for curvature. If the model is nonlinear, at least you can reuse the results from this two-level design.
- *Do you need results that are predictive or merely comparative?* You don't get something for nothing; predictive results usually require a lot of runs. To reduce the number of runs, consider doing some screening runs with a fractional design of experiments to determine which inputs are driving the model's response and also which inputs drive the variance of the response; these are the inputs you need to focus on. If you only need comparative results, this fractional design might be all that is needed.
- *Are you trying to optimize on the output of your model?* If you are looking for the "best" settings for the inputs within a fixed range of values, a normal DOE could be used. If you don't know *a priori* where the optimum may be, try using Response Surface Methodology.

When planning a run matrix, I usually start with an idea of how much computer runtime I'm willing to live with and back out the number of runs. I consider if I can divide the runs across multiple computers (multiple computers allow you to reduce runtime or increase total runs). If I'm using new data or a new model, I expect to have to do a lot of my runs over at least once, but more likely two or three times.

The planning phase is also where you need to develop your data storage and retrieval methods. Here are some things to ask yourself:

- *Do any data require preprocessing?* Some data may not arrive in the format required by your model. Sometimes all you needed to do was ask and you could have gotten them in the format you needed. You really don't want to be doing unit conversions if you don't have to.
- *How will you tie your input files to your run matrix to your output data?* When you are analyzing your results, and run #432 did something strange, you need to be able to track down and investigate the batch file and the input files.
- *Do your results require any post-processing?* Your model may be generating a database of inputs for another model, or you may need to calculate the mean and standard error of an output if you are doing replications.
- *Are you using a spreadsheet or a database program?* Most spreadsheets have nice graphics and slick features like pivot tables, however they also have a relatively small limit on the number of rows of data. You can usually work around this limitation by storing your data in a database and linking your spreadsheet to the data table.
- *Is your hard drive big enough to hold all your results?* You only want to imagine what happens if your hard drive is full after two days of a three-day run. You backed up your data, right?
- *If you are splitting up your runs across multiple computers, how will you ensure they use the same data? How will you process the results?*

Dwight D. Eisenhower said, "In preparing for battle I have always found that plans are useless, but planning is indispensable." Remember that your plan is only a placeholder for what you eventually end up doing. Without planning your analysis, how can you effectively test your data and models?

Anything Not Worth Doing is Worth Not Doing Well (Data/Modeling Phase)

If the model always did what you needed and the data to run the model always existed, the "crank turning" aspects of analysis would be reduced to an academic exercise. The reality of analysis is that you almost always need to modify your model to work with your data, or you'll need to find data that don't exist and get it to work with your model.

In some cases, you'll need to develop your own model because one doesn't exist, or the existing model is so poorly documented and/or written that it is just easier to start anew. The title of this section is for those of you that are going to develop code: don't spend a lot of effort modeling something that won't matter, don't build a model that uses data that won't exist. Things that are not worth modeling are best handled by modeling assumptions. A good example of this appears below.

I once wrote a model that predicted aircraft availability for a squadron of aircraft. These aircraft required 2 days of scheduled maintenance after each 100 flying hours; once an aircraft exceeded its 100 hours it was not allowed to fly until it had this maintenance. A simple queuing problem I thought, since the base only had two hanger spaces to do this work. I soon learned that the scheduler that decides which aircraft fly each day would never send an aircraft over its 100 hour mark unless there was a maintenance dock to tow it into when it landed.

> **"It is the mark of an educated mind to rest satisfied with the degree of precision which the nature of the subject admits and not to seek exactness where only an approximation is possible."**
>
> ***- Aristotle***

My simple queuing problem became a complex scheduling problem that, if I were able to model successfully, would not impact the availability. My fix for this problem was to model scheduled maintenance by assuming that aircraft were always able to get a maintenance dock when they needed it.

I discovered this error because I took the time to verify my conceptual model with real operators (people who know what happens) before I started to write code. My conceptual model was a simple logic flow diagram that I used to show the maintainers how my model would fly/break/fix the aircraft. This also was a great opportunity to find out what data existed, who owned them, what they looked like, and how to get them.

It Worked Fine with the Test Data (Uh-Oh Phase)

There's something about real data. They can break a stable model, they can thwart your hypothesis tests, and they can refuse to be imported into your database.

Real data can change every year. Ask your data source how often they update and if they save the old data. I once worked on a study that carried over into the next year. When I needed to update my database, I found out the data coding scheme had been changed; I had to rebuild my entire database.

Real data sometimes fail to materialize. If you identified data that you absolutely must have, work it early and work it often. Investigate alternate sources for the data.

Real data can be wrong, it is just as important to validate (certify) your data as it is to validate your models. If you can use data with some credentials (pedigree), by all means do so and save yourself the headache. But real data is not the only thing that can cause things to go wrong. Another culprit is the hidden assumption.

A hidden assumption is an assumption (embedded in models or data) you unknowingly agreed to and discovered later. For example, you used an air-to-air engagement model to examine the on-board radar range requirement for a new air superiority fighter. Only later did you discover that the model always cued your on-board radar to the bogie at 100 miles to simulate an AWACS.

If you must develop your own model, be aware that software development is an event driven process while most analysis is schedule driven. Your best bet for reducing schedule problems is to start by developing a simple yet flexible model and maintaining an archival trail of improving model versions.

This sounds like motherhood and apple pie, but several times I have observed the problems that can happen when you try to develop the complex model before the simple one. You usually struggle to find all of your syntax errors because you have so much new code. Then, after a couple of weeks you come to the realization that your data structures are wrong or that the model will probably take a week to run. Your only recourses then are to start over because you don't have any previous versions to fall back on, or to chop out big parts of your model (making the model simple) so you can figure out what's going on and maybe fix it. You entered the labyrinth without your ball of string.

Another common problem in analysis is using an existing model outside its limits. Daedalus was imprisoned in the labyrinth with his son Icarus for revealing its secret to Theseus. They eventually escaped the maze to find themselves stranded on the island. Daedalus invented his famous wings and told his son not to fly too low to keep the feathers dry nor too high so the wax won't melt.

We all know what happened. The problem is you don't have a Daedalus to tell you the limitations of your models. I've found that most model documentation tells you what you can do, but not what you can't. What I usually do in this situation is build up some test cases to determine and document if and when the model starts to behave badly. This process can be time consuming, but you must discover these limitations before you commit to model runs you cannot do or to collecting data you cannot use. Of course, you won't really know if there is a problem until you have the real data.

> **"The only way to discover the limits of the possible is to go beyond them into the impossible."**
>
> ***- Arthur C. Clarke***

Measure Twice, Cut Once (Production Phase)

We've all heard this carpenter's saying, but it also applies to reducing the time and trouble it takes to perform your production runs. If your model supports it and your hard disk is large enough, you should automate your runs using a script or batch file. Here are some things that are on my measure twice list:

- Do some single test runs at what you think will be the extreme or most stressing input values.

You don't want to find out that half of your runs stopped early because your weapon stockpile inputs were too small.

- If you are using a stochastic model, do some replications to get a feel for the variability between runs. If you don't do enough replications the first time, it is a lot of work to do more because of database and filename issues, random number seeds, etc. Saying the war lasts 60 days ± 30 days won't get you any acclaim, either.
- Exercise your batch file. I usually write my batch files to rename output files and move them into different data directories after each run. Invariably, I'll have a typo, or forget to create a directory. Usually, if the batch file works for a couple of runs it will work for all---usually. If you're doing a lot of runs, it's worthwhile to write a small program to create your batch file for you. This will improve the odds of any batch file problems showing up right away and make them easier to fix.
- Verify your data paths. This is partly done with your test runs but only for your input files. You also need to make sure that your output files can be post-processed. Can your databases read the files? Are the results formatted correctly? If your model outputs 0.0001 as 0.00 but you needed it to output 1E-4, you are still in the uh-oh phase.

In my experience, most inexperienced analysts tend to start "turning the crank" too soon, resulting in many false starts to the analysis. The less time you spend doing analysis that won't be used, the more time you will have to check your results, answer the inevitable what-if questions, and prepare your report.

Am I Done Yet?

Only Robinson Crusoe had everything done by Friday.

Your production runs may be done, but you are not. It's time to perform a few checks on your results before you invest a lot of time processing output data, writing reports, and making briefing slides. There's still a chance something went wrong.

Do a "look right" test. Look in your output file directories. Does it look like all the files you expected to see are there? Do they all appear to have data in them? If three files have a 100-byte file size and the rest are 500 kilobytes you may have a problem. Open up a few output files at random and page through them. Do they look like you thought they would?

If your output files look okay, compute your metrics and give them a "look right" test. Are the results consistent with the inputs? Are the results believable? Can you explain all counter intuitive results? If you answered no, you are not done yet because if your customers can't understand your results, they won't use them. Check the output data files and your post-processing methods for errors.

If you can identify the specific run(s) with problems, check your batch file and input files for errors. Try to find out if the problem is the data or the model. Look for patterns that may give you a clue. Are the problem runs at extreme values? Were they all run on a different computer? Is there a replication acting as an outlier? Explain the problem to another analyst. At times fresh eyes see the obvious (or you see it yourself during your explanation). Don't overlook the possibility that the model may be correct and your intuition is wrong.

If everything appears to be fine, you may want to invest in some insurance. Two ways to do this are by making sensitivity runs, and by getting feedback from your subject matter experts.

Sensitivity runs are useful in assessing how robust your results are to uncertainty. Even though you used real data, the data may (and usually does) have significant uncertainties. If your analysis shows Alternative *A* is much better than every other alternative, try degrading Alternative *A* to find out what it would take for the second best alternative to be preferred. Then, consider whether this degraded situation would be likely to occur. Sensitivity analysis is especially useful in studies that are contentious.

Lastly, get some feedback from someone you believe understands the problem being studied. This feedback will not only give you a "warm fuzzy" (you hope), but it will also allow you to test and improve your presentation formats as you prepare your final report.

9 Making Sure You're Sure

Data Accreditation and Model VV&A

by Keenan Kloeppel, Office of Aerospace Studies

ENIAC, generally recognized as the first electronic digital computer, was built to calculate ballistic tables for the large number of new guns developed during WW II. While it became operational too late to support the war effort, ENIAC introduced the U.S. military to the computer age. Within a few years, military decision-making began the move from being an art based on simple calculations and judgment to quantitative comparisons based on computer modeling of alternatives. The introduction of computer modeling allowed "building" and "testing" each alternative without the expense of building and testing each alternative.

This new paradigm had the effect of shifting the standard argument from the rather personal "I don't trust your judgment" to the more esoteric "I don't trust your computer model."

As a result, people began to ask pointed questions. Who developed the model? What expertise do they have? Did they consult with the "right" people? Where did they get their data? What are the limitations of the model? Has the model been checked against real test data? And on and on…

Nevertheless, this was progress, for it is easier to demonstrate the believability of a computer model than the infallibility of one's judgment.

The use of modeling was nowhere more prevalent than in the selection of major weapon systems for acquisition. Typically, a major command would support a new weapon system by running computer simulations to show its worthiness, then promote the results to skeptical audiences. Instead of discussing the consequences of the results, audiences would focus the discussion on the credibility of the results. All the above model and data questions (and more) would be asked and answered—usually to no one's satisfaction. Thus, reaching conclusions about the weapon system remained a very contentious process.

As a result, processes for enhancing the credibility of data and models have been developed. Data certification is aimed at guaranteeing the use of suitable data, and verification, validation and accreditation (VV&A) help insure the use of acceptable models. Very simply,

- ***Data certification*** is the process of obtaining expert consensus that data are the best available data for specified uses
- ***Verification*** is the process of establishing that a model's computer code does what it was designed to do
- ***Validation*** is the process of establishing that

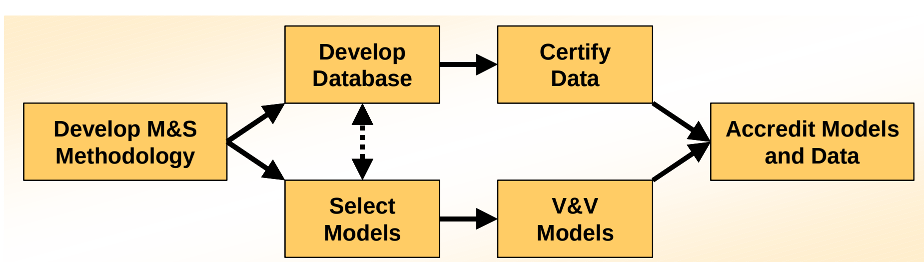

Figure 9-1 : Steps in the Data Accreditation and Model V&V Processes

the computer code satisfactorily models reality

- ***Accreditation*** is the decision to accept data and a model or models for a specific application, based on the results of data certification and model V&V

The interrelationship of the data accreditation and model VV&A processes are depicted in Figure 9-1. Both processes depend on the study's methodology, but neither depends on the analysis done during the study. Practically, neither the data nor the models can be selected entirely independently of the each other as different models doing similar functions may require different input data.

In this chapter we explore some of the practical aspects of data certification and model VV&A. It is best to look at these two practices not as guarantees of correctness, but as insurance against grievous errors.

Data Certification

The groundwork for an analysis requires identifying and acquiring the data that will be used in the analysis. These data may consist of scenario definitions, environmental databases (e.g., climate or terrain), system parameters, decision rules, study assumptions, etc. The identification should be done by experts dissecting the study's analytical methodology in detail. Acquisition of the data is usually parceled out among study team members as appropriate.

Expert Sources

If the study is being conducted for a large effort such as an Analysis of Alternatives (AoA), a Data Validation Working Group (DVWG) is formed to determine and pursue data. The DVWG consists of experts from relevant areas. They work together to identify and document the information necessary to conduct the study. Their documentation covers their processes and results; it ensures that all of the analysts and modelers have a common data reference.

For a smaller scale study, an analyst typically takes on the job of identifying required data through consultations with experts and from reading authoritative reports. Once assembled, the data can be reviewed as necessary. Again the data acquisition process and results are documented.

Consistency with Similar Studies (Pedigree)

A pedigreed database is a database whose ancestors are properly documented (and presumably revered). There is a current effort underway to maintain a centralized pedigreed database suitable for many, if not most studies. This database would consist of all the known threat scenarios and would contain other up-to-date system information (P_k values, target sets, vulnerability, mobility rates, etc.) based on expert knowledge. Such a database enables all studies to have access to the best and most current information from a central location. It can save the time and expenses of convening an expert group for each study and could eliminate data discrepancies between studies.

Theoretically, and even in practice, this idea has merit. However, it is not hard to imagine significant disagreements among experts about the "best" database values. Nor is it hard to understand that almost every study generates requirements for data that will not be in the database (why conduct a study that didn't examine something new?). A possible solution to the first limitation might be to allow contrasting expert opinions into the database. The areas of the database that show disagreement can then be dealt with on a case-by-case basis.

V&V of Models and Simulations

As shown in Figure 9-1, models must be chosen concurrently with data acquisition. Ensuring that the models will do a proper job constitutes the verification and validation of the models.

Necessary Capabilities of the Models and Simulations (M&S)

At a minimum, each model must be able to do all of the following:

Accept All Relevant Data

The data certification effort determines the information that is important for the study. The models must be able to use the data that are relevant to the piece of the puzzle being addressed.

Simulate the Problem

Does the model represent what we need represented? It might be able to accurately represent an existing aircraft, but can it be adapted to handle one of the proposed systems (which may not even be an aircraft)?

Generate the MOEs

The study team identifies the measures of effectiveness (MOEs; see Chapter 6) used in the study to evaluate how well alternatives meet mission needs. These metrics are crucial for judging potential solutions. The models must be able to provide output in a form that can be used to calculate MOE values for each system simulated.

Be Assessed to Be Correct

The models may simulate the problem, but is the simulation correct? If it compares favorably with test data, then we can say it is "correct." If there are no test data for comparison, then we must find other ways to validate the model. These might include:

- Demonstrating the code is based on "proven" concepts and the interconnections are logical
- Showing that an empirical code (or a portion of it) can be favorably compared to a "validated" physics code
- Having experts review the code and its results

Documentation

Does the model have a user manual or online help? Is this documentation clear and does it provide detailed information on the model options that are available? Documentation often can be used to help verify model capabilities and may also be used to validate its correctness. It may provide information about the extent of the configuration management of the code and the capabilities and limitations of each of its configurations. This may help determine which version is best suited for the analysis.

Instant V&V: A True Story (I Think)

This story was told to me with a straight face by an astrophysics professor I had in grad school back in the early 60's. A Ph.D. candidate submitted his dissertation based on hundreds of pages of computer output—the results of numerically modeling stellar atmospheres—then a hot new topic. Not being fools, his committee asked why they should have faith in his results. The ever-resourceful graduate student returned the next day with each page of computer output bearing the words "These results certified by IBM." End of discussion. Who was going to argue with Big Blue?

How to Use

The entire scope of the analysis effort needs to be laid out to determine how well the models will fit together to provide the answers. In what area of the study is the model to be used? Will the output of the model be used as input to another model? Does the model need to accept input from another model? In a complex study, all of the candidate models need to be viewed as a piece of the overall analytic puzzle. This is illustrated schematically in Figure 9-2.

How It Works

Try to keep in mind what needs to be accomplished. How does the model receive input—files, graphical user interface (GUI), some other way? How long does it take to set up a problem—an hour, a week, a year? What platforms will the model run on? How long does it take for each run?

The model must not be too complicated to support the effort. Don't be impressed if a model has unneeded bells and whistles. Choose the simplest and most straightforward solution.

Inherent Assumptions, Limitations, and Errors

Some assumptions, limitations and errors simply won't matter for some simulations. In other situations they can make the model unsuitable. It is very important to understand the inherent implications of any model and determine if the impact is critical to the problem.

What types of assumptions are inherent—models only clear weather, doesn't model multi-path radar signals? What types of limitations are inherent—calculations are accurate only within a given range of speed, temperature? What types of errors are inherent—fails if you play an F-16 and a Scud B together?

Practical vs. Impractical VV&A

The processes described in this chapter are meant to be practical applications of data certification and VV&A performed for an analysis. An in-depth, independent V&V (IV&V) of a model is a separate action that should be left to the model manager. If an IVV has already been done, the results can be used to support a study's VV&A effort. Because IVV is both expensive and time consuming, it is often delayed until a model has matured and found its niche in the analysis community.

Prior V&V

Has this type of V&V evaluation been done before? If it has, and the analysis is similar, then there is no need to repeat history. Plagiarism is allowed. There is an ongoing effort to create an M&S repository, where all previous V&V efforts can be accessed. This will save a lot of time if the problem that needs to be solved is similar to one already done. The write-up for that previous effort can be used in the new VV&A report.

Accreditation of Modeling and Simulation Data

If the models have the necessary capabilities, they can be evaluated to determine a level of trust in the answers they provide.

Describe Model V&V and Data Certification (VV&C)

The V&V report for each model extensively documents the compliance with the items discussed above. The subsequent evaluation can either be a quantitative grading where each section

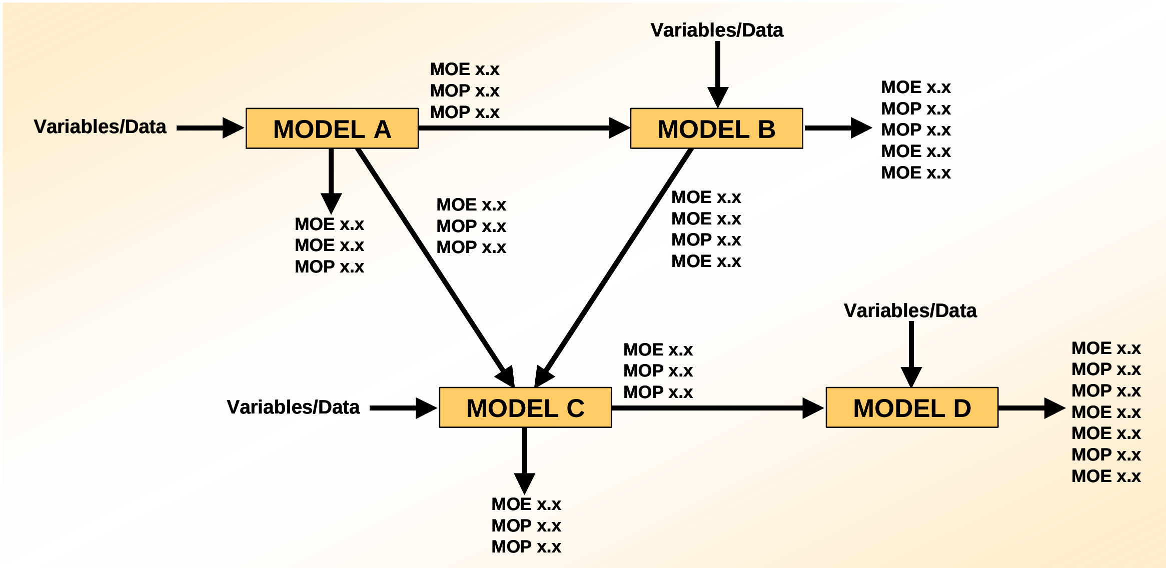

Figure 9-2 : Schematic Indicating Linkage of Data, Models, and MOEs & MOPs

receives points for how well it rates (say on a 10 point scale) or a subjective evaluation (poor, fair, good, etc.). Based on these evaluations, a recommendation for each model's use in the study can be made.

The detailed V&V reports for each model are included in an accreditation report and recommendations for the use of each model are provided to the accreditation authority. The process for the certification of the data is also included in the accreditation report.

Detail Any Shortcomings

In general, anything of significance uncovered during data certification or V&V is reported. The accreditation report details any shortcomings that a model might have. For instance, a model might be well suited to provide MOEs for fixed targets, but incapable of simulating mobile targets. The shortcomings may not necessarily be important to this study, but might help determine the capability of the model in a future effort.

Provide a Recommendation

The accreditation report provides recommendations for the intended uses of each of the models evaluated. Ideally, the report ties each recommended model to the study's methodology. It should contain enough supporting information to answer any question the accreditation authority might have.

When finished, the report is sent to the accreditation authority. In the case of major studies like an AoA, this is typically the Director of Requirements (DR) of the major command conducting the study.

If the authority agrees with the recommendations in the report, the report will be signed and the models can be deemed worthy for use. Thereafter, should anyone start to question the capabilities of the model during a briefing, simply hand them a copy of the accreditation report for their reading pleasure, and proceed with the summary of the analysis.

10 Phantoms of the Opera

Dealing with the Future

In some sense most of the analysis done in the Air Force tries to peer into the future and assess potential future systems in a future world. The analyst has four fundamental questions to answer for any future system: What are the system's physical characteristics and capabilities? How will the system be employed? In what circumstances (scenario, order of battle) will the system be deployed? What will it cost?

The technology community answers the first, the operational community the second. The intelligence community deals with the third and the cost community with the last.

In making these projections we typically have to:

- Model developmental technologies as deployed weapon systems
- Develop operational concepts for these systems (and current systems projected into the future)
- Anticipate the interactions of new and legacy systems
- Choose and possibly augment the scenarios in which they will operate
- Assess their life cycle costs

These are not easy tasks, and they cannot be sidestepped. In this chapter we focus on future systems from the points of view of system hardware (technology) and operational concepts. We also touch on scenarios and cost, although cost considerations are discussed in depth in Chapter 12.

> **"Future. That period of time in which our affairs prosper, our friends are true and our happiness is assured."**
> ***- Ambrose Bierce***

Defining a New System

A new system concept is initially visualized either in terms of a technological solution (hypervelocity penetration, exotic rocket propellants, stealth, etc.) or the system's principal hardware element (e.g., an airplane, a munition, a radar). Only later—usually much later—is consideration given to what a functioning implementation of that technology must look like or how the hardware must operate and interact with other systems.

Thus most new system concepts arrive on the analyst's doorstep in a rather primitive state. So primitive, in fact, that they must be fleshed out by both technologists and warfighters (operators) before they can be subjected to meaningful analysis.

The focus of the technology community is, of course, developmental technology. Their job is to prove that their concept works and can be packaged into a militarily useful system. The warfighter is concerned with how that system can be properly employed.

Unfortunately, these two factions are not accustomed to having long and meaningful conversations with each other. In fact they almost never talk. This leaves the analyst as the go-between, enhancing, integrating and arbitrating their thoughts. The more familiar the analyst is with their workings and concerns, the easier it is to perform this task.

A new system arises from a new application of an existing technology or from the projected application of a developmental technology. In either case the elements of the new system (the hardware) must be engineered to some level of detail to ensure their design feasibility. The engineering design details determine performance and help us deduce the interactions of the new system with other systems. Typical issues of concern for new systems are:

- Performance
- Weight and volume constraints
- Affordability
- Vulnerability and survivability
- Operational concepts
- Interoperability with other systems
- Ease of deployment
- Treaty constraints
- Manpower requirements
- Safety implications
- Maintenance structure
- Ability to be manufactured
- Environmental constraints
- Climatic range of operations
- Development schedule
- Risk (all kinds)

These issues are often difficult to assess even for an existing system. They are much harder to assess for the "paper" systems we are discussing in this chapter.

An Iterative Design Process

Consider for a moment a typical technologist. More than likely, the technologist was educated in the physical sciences or as an engineer and has probably worked on the technology in question for years, beginning with basic research and developing it into something potentially useful. The technologist knows the technology inside and out, but has only a vague concept of how it can be applied. The technologist doesn't understand how or where it will be used, what it must cooperatively interact with, or what threats it will encounter. As the expert, answers to system design questions must begin with the technologist alone.

Logic and Emotion

Technologists are paragons of reason—except when their concept is in danger. In a study of innovative rocket propulsion technologies that I led, the proponents of one of the poorer performing technologies insisted against all logic that their technology be considered to perform at its theoretical upper limit. To keep peace the study team acquiesced, but of course the results were clearly not realistic and were never seriously considered.

The analyst should understand this emotional need to protect a long-term intellectual investment regardless of the facts.

Now consider a typical operational officer, say a pilot. The pilot is geared to flying, fighting and surviving with a real aircraft. The pilot uses that aircraft daily and knows it intimately, has been taught the correct tactics, and has practiced them at every opportunity. Not surprisingly, the pilot does not understand the new technology, nor has much of an idea how best to use it.

As already indicated, your job is to work with both of these capable but narrowly focused sides to create a technologically and operationally sound system concept for evaluation.

RULE ONE: this can only be done through iterating both the system hardware design (technology implementation) and the corresponding operational employment concepts.

Simplistically, the technology community designs, the operational community decides how to employ, and the analytic community analyzes and finds the blemishes before sending the system back for redesign. Of course, the three communities can and should work in parallel in a collaborative and integrated effort, and if you're clever you can make this happen.

In practice the iteration process typically works as follows. For unavoidable reasons, the initial conceptual design of the system that is handed to the analyst for evaluation is seriously flawed. Tunnel vision and poor lines of communication are certainly partly responsible, but in reality system design is inherently difficult. When the system design is analyzed, it will understandably perform poorly. Using insights gained from the analysis, a reworking of the concept and its employment is performed. This results in a significant improvement. Analysis of this improved version will likely uncover fewer, less critical flaws. This process continues through several cycles, until the team is satisfied that they have a viable and competitive system.

The iteration process significantly improves the design of the new system. However, it consumes time and resources that are rarely planned for in the original schedule. This can be a significant reason for not meeting deadlines, and it often leaves the analyst (or the analyst's boss) with a strong desire not to cooperate in the iteration process.

A display of pique is not appropriate. First, non-cooperation leaves the proponents of the new system (justifiably or not) crying foul at the study's end—and there will be listeners. More importantly, the job of the study is to produce the best answer possible, and this means cooperation.

Choosing Future Scenarios

In the old days the intelligence community was reluctant to provide projections of future scenarios and hostile capabilities. That has changed. For example, there are the Defense Planning Guide scenarios that are frequently used in a variety of studies. However, there is no single view of the future and you may find that you need to make additional projections for any of many reasons. This is especially true when dealing with new systems, for new systems generate reactions by potential foes. These reactions are characterized as reactive threats—we do this, they counter with that. If you are going to consider reactive threats—and why wouldn't you?—it is usually best to create them within your study and submit them to the intelligence community for approval. Approval is equivalent to "given what we know, we think your choice is suitable for your analysis. More we won't concede."

Understanding Technology Development

You need to understand the basics of technology development, for it is likely that no one else working on the study will. Understanding technology development is important to assessing system development cost, schedule and risk. Currently, the best yardstick for measuring technology development is the National Aeronautics and Space Administration (NASA) technology readiness level (TRL) scale. This scale is illustrated in Figure 10-1. This figure shows nine stages of technology development beginning with basic research (TRL 1) and ending with a proven system at initial operational capability (IOC, TRL 9). Practical experience with applying this scale has shown that knowledgeable, good-intentioned people will generally agree on the current TRL level for a technology program. Where they do not agree, they will not disagree by more than one level.

Using TRL to Assess Technology Development Schedule

Occasionally the analyst may be faced with the problem of assessing the time from a developmental level to IOC (TRL 9). This is a legitimate question, if hard to answer. One way to approach this problem is by drawing an analogy between the technology in question and a similar technology that has already been incorporated in a system. The answer you get—after allowing for obvious differences between the programs—may not be completely accurate, but it will be better than nothing.

Using TRL to Assess Risk

Technology development risk and cost risk are among the most difficult assessments that you can be asked to make. As with estimating development schedule, you can make an analogy with a previous development program. You can also try to identify the number of development paths available to get from each TRL to the next. The more and more varied the possible paths, the less the risk.

For example, at TRL 3 the technologist must basically prove that the concept works. To get to TRL 4, a breadboard design, there may be only one possible developmental path; if it proves difficult, there are no fallback alternatives. This represents high developmental and cost risk. On the other hand, having multiple routes allows multiple approaches; these may be pursued serially or in parallel as appropriate.

A Simple Analogy

Imagine that there is only a single road you can take to work. If there is an accident that closes that route, you have no alternative. However, if there are three routes to work, it is very unlikely that all three will be blocked. You may be delayed, but not stopped.

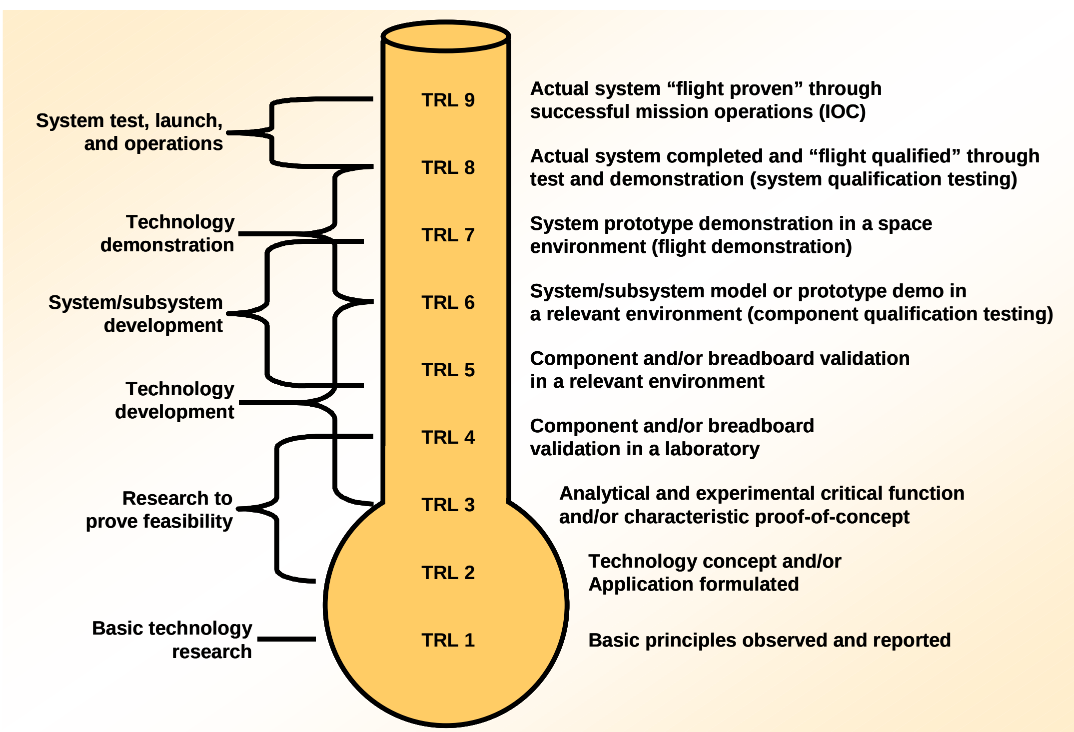

Figure 10-1 : The NASA Technology-Readiness-Level (TRL) Scale

11 We're Only Human!

Modeling Human Behavior

by Larry Looper, Air Force 311th Human Systems Wing

There is a law that governs the behavior of human beings. That law is called the Law of Individual Differences. Stated simply: People are different. There is also a negative corollary to that law which states: People are a lot alike. Given these two conflicting views of human nature, which is correct?

The answer is they both are correct. Generally for a human trait (physical, psychological, or behavioral), the expression of this trait in a population will vary pretty much according to a normal distribution. What this means is that most people will display roughly an average amount of this trait (which we might call being "normal"), but some people will have a great amount and some will have a relatively small amount.

This duality and range of "humanness" are the very things that make us the most interesting of creatures, and at the same time the most difficult to understand and, consequently, the most difficult to represent in mathematical models.

What aspects of human nature are most important to the success or failure of military operations? Is the performance of a military unit simply the sum of the performances of individual members? Is it reasonable (or even necessary) to consider incorporating such information in large-scale theater-level models? How will tomorrow's technology affect the role that humans play on the battlefield?

This chapter first attempts to address some of these questions by providing the analyst with a brief review of what is known about human behavior. This is followed by an equally brief look at the current state of the art in human and military modeling. We conclude by addressing what the future may hold as the role of people in complex technological systems unfolds.

Categorizing Human Behavior

The intent of this chapter is to acquaint you, the analyst, with those aspects that are most relevant to military modeling. This section discusses what we know about human behavior in terms of learning, decision making, team behavior, and cultural and other external factors relevant to military models.

How and what people learn is affected by many factors. Variations in these factors from individual to individual contribute to differences in how military decision makers react to a given situation. In general, people learn from rules or from examples. Many of the courses you had in high school and col-

If It Weren't for the Humans...

A colleague of mine, a management professor in a large university, once declared to me in frustration, "This school would be a great place to work if it weren't for the students!" His frustration arose because he had been delivering his lectures with the same material and in the same manner as always, but he couldn't seem to connect with one particular class. "The students and I are never on the same wavelength," he continued. "I can't understand their questions and they never seemed to understand my answers. What's always worked suddenly doesn't."

(Law of Individual Differences)

lege involved the dissemination of rules. Thus, a common complaint of students is that they don't see the relevance of what they are learning to their situations. They are expressing a need for learning by example.

The problem with examples is that they are, indeed, just examples. They may not be representative of what is (or should) be learned by most students. Even examples in which the student plays an active role (either in a class or in real life) may not be the best example on which to base future behavioral decisions.

Learning is dependent on life experiences and exposure to various learning situations. A graduate of a military academy who's parents were military officers would have a much different set of life experiences than would an ROTC graduate who's parents ran a grocery store. Each would bring (unconsciously, most likely) their backgrounds into any military decision situation, and the resulting decision could be different.

Learning is also dependent upon memory and attention capacities. Memory, whether short or long term, can be populated with information from specific life experiences (e.g., the successful taking of a hill by a platoon of soldiers) or more general experiences not directly linked to specific events (e.g., ability to recognize colors). Attention can be focused (the decision maker ignores irrelevant information), selective (the decision maker pays attention to certain information over other information), or divided (the decision maker attempts to consider many different sources of information at the same time). Each of these aspects is important to the ability of individuals to focus on the task at hand and to do different tasks at the same time, a critical part of most military operations.

The ability to make the right decision at the right time and place is the most critical task of the military decision maker and the most critical to be modeled correctly. People make the decisions, and the bases they use are as varied as the people themselves. Few decisions are made completely without input from other people. It is also critical to understand and model how teams (from two-person teams up to entire divisions and armies) make decisions.

Various models of individual and group decision-making exist. Many of these are "rational person" models, meaning they assume that the individual has all the information necessary to make a decision and that the decision maker will make the optimum decision in that situation. However, given our understanding of the vagaries and variations in humans, this is most often not the case.

This is particularly true in the context of decision making in military operations. The decision maker will probably lack the information needed to be able to enumerate or evaluate all the alternatives. And here, too, memory, attention, and life experiences of the decision maker (especially those most relevant to the decision context) enter into the decision process.

Understanding how teams or groups make decisions is even more vital given the complexity of tomorrow's field of battle and the communication tools that will be available. The design and functioning of the command, control, communication, and intelligence (C3I) gathering and evaluating structure of the team or unit can facilitate certain kinds of responses or make them nearly impossible. Individual and group paradigms about behavior are extremely useful in helping sort through and categorize influxes of information by importance. Such group paradigms or norms may also make it very difficult to identify unusual bits of information or to react to novel situations in any way other than in very predictable past patterns. Nothing is more dangerous for a battlefield commander than to become predictable. Yet, failing to respond in a timely manner while trying to weigh all the factors or attempting to achieve team consensus could be fatal.

In the field of decision making, that eminent scholar and theoretician, Yogi Berra once said, "When you come to a fork in the road, you gotta take it." Yogi was exactly right; decisions must be made. Understand that *not* making a decision is still a decision!

If understanding and modeling individual and team behaviors weren't difficult enough, the analyst must take into account many other types of internal and external factors. Among these are physiological factors such as temperature, humidity, toxicity, noise, and light. The level and nature of the work being performed by the individual or the group and the level of fatigue are also key concerns. Of course, the intelligence level of the individual can influence both the individual and the group's performance. Characteristics such as personality, emotions, attitudes, cultural values, and norms are highly relevant to understanding individual and team behavior; these characteristics are even more difficult to measure.

Models, Models, Models

Although a complete understanding of human behavior may never be possible, there are a number of existing models of individual and group behavior. These models may be categorized as cognitive or non-cognitive.

Cognitive models attempt to capture and replicate the cognitive processes going on the human brain in terms of sensing and perception, memory functioning, decision making, and motor control. Many of these models are at levels of detail far too explicit to be incorporated into large-scale military operations models. They attempt to model the inner workings of human neurological and neuromuscular activity.

Non-cognitive models, on the other hand, attempt to understand, measure, and represent aspects of humanity such as personality, attitudes, and emotions. But the nature of the role such aspects play in military operations is very unclear.

It is said that General George Patton believed he lived a past life as a great warrior. Saddam Hussein was rumored to listen to the advice of only a few close family members. As a youth, General George Armstrong Custer wrote a stirring essay on the defense of the American Indian. How such attitudes and behaviors were related to the decisions these leaders made is unknown. What is clear is that they most certainly were important in either their conscious or their unconscious cognitive processes and the decisions they made in a particular military operation.

American military decision makers use a wide variety of different types of models. Military operations models can range from small individual combat models to large-scale theater-level models. There is a great deal of time and effort put forth in attempting to make certain that the weapon systems in these models—their functioning, performance, and effects—are properly represented. There is also much effort spent in incorporating into the model key C3I activities along with the movement of units.

Often, the effects of external factors (such as weather) are also included. But there is very little time spent incorporating human factors into such models. This is not to say that there is not interest at the highest levels in the Department of Defense that such modeling be undertaken. There is.

In the U.S. Department of Defense, Modeling and Simulation Master Plan, 1995, the Under Secretary of Defense for Acquisition and Technology established an objective to "develop authoritative representations of individual human behavior" and to "develop authoritative representations of the behavior of groups and organizations."

A recent study by the National Research Council's Panel on Modeling Human Behavior and Command Decision Making was completed at the request of the Defense Modeling and Simulation Office. It concluded that existing military operations models "do not consider the current generation of human behavior representations to be reflective of the scope or realism required for the range of applications of interest to the military." The study also found that human behavioral representations were needed at levels from the individual combatant up to divisions, wings, and battle groups, and for friendly as well as opposing forces.

The National Research Council's report identifies 23 current military models and simulations that consider some aspect of human behavior, all of which have relevant human behavior components. There are, however, several hundred models in use by the military services covering a wide range of military disciplines.

Most of these 23 models have rule-based behaviors, usually determined by experts. Some incorporate human behavior implicitly as represented by the decisions of the human players in the simulation (either in live simulations such as exercises or simulations with live players linked by computer networks). However, the Council states that none of these models and/or simulations "provide the fidelity of battle outcomes on a real battlefield, where the number of casualties or weapon system losses depends on real human strengths and frailties and varies significantly from unit to unit based on leadership, stress, consistency of tactical decisions, and effectiveness of training." The Council goes on to say, "this lack of human performance representation in models becomes more significant as the size, scope, and duration of wargaming simulations continue to grow." The Council concludes: "In the future these limitations will become more noticeable as greater reliance is placed on the outcomes of models/simulations to support training and unit readiness, assessments of system performance, and key development and acquisition decisions."

Some models do exist which could be described as individual decision models. Among these are models determining the number of enlisted and officer personnel needed to maintain a current force level and the subsequent number and types of new personnel needed to maintain these levels. Supporting models predict which and how many personnel would reenlist or remain on duty as an officer over a future time horizon. Still other models estimate the number of recruiters and the recruiting advertising budget required to encourage civilians to join the military, estimate the effects of pay, bonuses, and service-specific personnel policies on retention, or estimate the likelihood an individual member will complete specific military training. Often, these models are based on the "rational person" model of optimum individual decision making, but some are simulation models using Monte Carlo techniques. Though narrowly focused, the results of such models provide the analytical bases for structuring the human component of the military forces to meet current and future threats.

The Future Of Modeling Human Behavior

What does the future hold for military modeling and simulation of human behavior? How can the military analyst contribute to the development of better models with the correct representation of human behaviors for the modeling task at hand? To answer the first question, let us take a brief look at three of the many technology trends that are most likely to continue in the future: increasing battlefield and C3I realism, distributed simulations and information warfare.

Computer memory, storage capacity, and speed are growing at nearly geometric rates. One industry that has benefited from this growth is the movie industry. Computer generated images continue to become more realistic as time goes on. Similar advances will also be possible in military operational models and simulations; battlefield and airspace scenes will have incredible visual realism.

Distributed simulation, already a reality in many simulations, will continue to expand. In distributed simulations, participants in simulators interact vocally and visually over a computer network. Participants may be geographically separated and may even be in actual weapon systems. Each participant, individual, or unit can see what others are doing. Such simulations permit more fidelity in the evaluation of C3I activities as well as tactics possible in various operations and terrain.

Lastly, information warfare may well be the type of warfare fought in the future. It may be possible to convince an enemy that their mission has failed and cause them to retreat or surrender simply through manipulation of the enemy's intelligence and situational awareness-gathering systems. That is as positive an outcome as physically defeating the enemy. Of course, the converse is just as true. We must be vigilant that the enemy does not infiltrate our information systems.

Unfortunately, attacks on government computer systems are happening at an increasing rate. Whether by hackers on a joyride through the Internet or by hostile forces, our computer systems, both government and private, are too often open to "backdoor" attacks. Attacks on our electronic inter-bank fund transfer systems, communication systems, and computer networks could be just as devastating as attacks on our military forces. Keeping them secure or "information hardened" is vital to our national defense. The capability to "harden" our troops and our infrastructure against such information attacks, through personnel selection and training (often with operational models) is just as important and a key future concern.

What Can the Analyst Do?

So, what can the analyst do to help ensure that enhancements to military simulations and models properly represent human aspects? The National Research Council's report recommends a series of short and long-range activities. These include:

- Continued development of theories of human performance

- Continued collection of additional human performance data
- Incorporation of such information into the right models
- Ensuring that such models are valid for their purpose and that results are accepted across the user community

The analyst can play a vital role in each of these activities. In the area of theory and new data, the analyst must initiate discussions with psychologists, sociologists, economists, and others who are attempting to further the theoretical bases for human modeling. Attendance by analysts at conferences of professional groups in these disciplines is vital.

To help ensure that human performance data are used in new models, government technical contract monitors—those overseeing the development of new models and simulations—must be kept aware of the need for proper human factor consideration in their models. Here, the analyst must reject contract statements of work that do not include the need for the human issues in a developing model.

In the area of validation and accreditation, the analyst should ensure that evaluation and documentation of such models are accomplished with the participation of human discipline experts.

Conclusion

Our understanding of the thinking, behavior and actions of individuals and groups is still at a rudimentary stage. Military modeling of such behavior is at an even more elementary point of evolution. Yet we must continue to expand our knowledge and our skills in this arena as the Department of Defense relies more and more on the output from models and simulations to make tactical, training, and acquisition decisions.

The very fact that this chapter is included in this handbook is evidence of the importance of including the human factor in tomorrow's models. Analysts must be willing to go beyond modeling the "easy" measures to properly representing in their models the challenging human issues found in complex human-machine systems such as military C3I systems.

The task will not be easy, but the task is absolutely necessary.

12 I Can Get It for You Wholesale

Estimating Costs

by Mary Benze, Office of Aerospace Studies

Cost analysis identifies the cost implications resulting from force structure decisions and has never been more important than it is today. With shrinking DoD and Air Force budgets, affordability is more often than not the deciding factor when new programs and systems are considered. Because military budgets are limited, each decision to follow a particular course of action (for example, acquire a new weapon system) means that the dollars committed for one purpose will not be available to use elsewhere. These considerations point out the criticality of having accurate cost information.

Nothing clarifies program decisions like knowing what the costs are!

Costing Overview

The products of military cost analysis are an integral part of the weapon system acquisition and budgeting process. The entire acquisition process is structured around the cost of an item and the availability of adequate funds at the appropriate time. In an environment that is budget-driven, we must make sure that our inputs to these budgets are founded on credible cost estimates.

To adequately estimate costs of a military system, we must consider the total Life Cycle Cost (LCC) resource requirements. To do this, we identify "cradle-to-grave" costs encompassing research, development, test and evaluation (RDT&E); investment (production); operations and support (O&S); and disposal costs of a system. All are necessary to identify the total program costs.

In the past, we have often based acquisition decisions exclusively on the costs to develop and produce an item, ignoring the costs for fielding and support. This approach more often than not resulted in a misinformed decision, since O&S costs can drive the total costs of a system over its life cycle.

In addition to developing the total LCC for a system, we must consider the risk and uncertainty inherent in any estimate. Since a cost estimate predicts future costs for a system about which very little may be known early in the program, there is a potential for considerable variability in the estimate. Risk analysis accounts for the "known unknowns" in a program and encompasses configuration, technical, schedule, and cost estimating risk. Uncertainty analysis adds dollars to an estimate for the "unknown unknowns" in a program. Once we have considered the cost impact for risk and uncertainty, we can estimate that the total cost will be no greater than a specific amount (for a particular confidence level).

There are a variety of cost tools and techniques available. We can create an estimate using a model such as the Automated Cost Estimating Integrated Tool (ACE-IT) to develop the entire LCC, or we may use specific models designed for particular LCC elements, such as the Cost Oriented Resources Estimating model (CORE) to develop O&S costs. We can use projections of actual costs, extrapolation of catalog pricing, analogies to similar systems, cost factors, or we can develop costs from mathematical formulas that describe relationships between system parameters and cost.

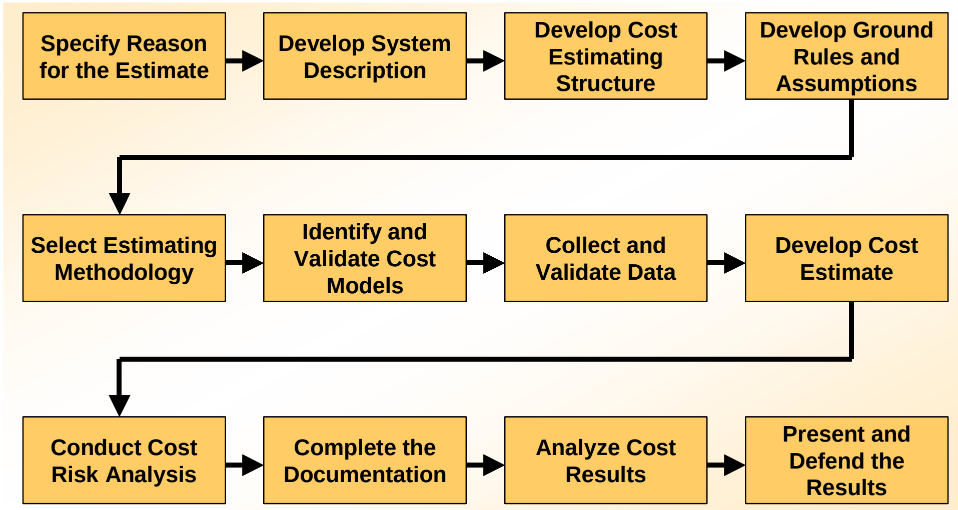

Figure 12-1 : Cost Estimating Process

As military cost analysts, we need to understand the complex technical system issues involved in estimating military weapon costs. A variety of disciplines—business, engineering, economics, mathematics, and operations research—can provide the required expertise.

But military cost analysis is often more art than science. In addition to academic training, cost analysts need other important traits to succeed. Some of these traits are:

- Paying attention to detail, yet being able to maintain a system perspective
- Being flexible, ready to adjust estimates when new information or decisions dictate
- Appreciating the importance of documentation
- Being able to effectively synthesize, communicate, and defend methodology and results to a variety of audiences—top management, peers, and subordinates

Costing Process

We have organized our discussion of military cost analysis around the process outlined in Figure 12-1. Each element of the figure will be described sequentially in this chapter. This process can be tailored to individual circumstances.

Specify the Reason for the Estimate

A cost analysis will have to meet different cost estimating requirements based on the estimate's end use. These different goals will impact the required resources, costing tools, techniques and input data.

For example, the laboratory commander's interest is based on a return on investment (ROI). This involves an analysis of when the investment costs will be recouped based on projected future savings or cost avoidance. A comparison of the "payback" periods can help the commander decide between competing capital investment projects.

In other instances, military cost analysis is used in cost effectiveness studies to help decision makers identify the most cost-effective system or "best value" for the taxpayers' dollar. These analyses examine the operational effectiveness of alternative concepts and the cost of developing, acquiring, fielding,

Technology Advances

Beware of overly optimistic technology forecasts (see Chapter 9).

supporting, and eventually disposing of these systems. The cost and effectiveness results are then presented for each alternative so the concepts can be compared and the most cost-effective solution identified.

A third example of the use of military cost analysis is Activity-Based Costing (ABC). Although initially developed for use in the private sector, ABC has been embraced by DoD as one of the "best practices" of industry and is being applied to many military functions.

ABC examines an organization's activities and products and breaks process and procedures down into the tasks performed within them. By applying costs to the individual tasks performed within each activity, we can then analyze how resource requirements would change if we altered our process or procedures.

ABC links "the cost of doing business" with individual activities and gives managers the ability to better evaluate the cost/benefit of these activities. It also is a powerful tool for conducting "what if" analysis to determine resource requirements if operations are scaled back or expanded.

Develop the System Description

Once we know what type of estimate is required, we then need to understand the system being estimated. Knowing the purpose of the system is a first step in determining the scope and complexity of the estimating task at hand.

The most important input to cost estimating methodology early in a system's life cycle is the detailed description of physical and performance characteristics. This will require extended involvement from the scientific and engineering communities and is often one of the more difficult tasks in estimating system costs, particularly for conceptual systems. Since these early system definitions tend to be only estimates of system characteristics and because they form the foundation for the costing, it is essential that they are presented and documented as the assumed baseline for the estimate.

For major acquisition programs, a Cost Analysis Requirements Document (CARD) describes the complete program, including the system description and salient features and is used as the basis for the program life cycle cost estimate. DOD 5000.4-M, Cost Analysis Guidance and Procedures, contains guidelines for developing a CARD.

The system description does not tell the whole story; there are also technology issues to be considered. Some elements of the system may be based on "off-the-shelf" technology. Others may be developed with existing technology. And still others may require specific technological advances.

It's a common practice to address issues of technology maturity by assuming that the new technology will be available when it's needed. While this assumption allows the estimating task to proceed, it also introduces a potentially high level of risk; this risk must be assessed in the risk analysis and accounted for in the cost analysis. Since minor errors in estimating technology challenges can significantly impact costs, and since system advocates often understate these challenges, we need to pay particular attention to technology issues (see Chapter 9 for a discussion of assessing technology readiness).

Also related to system design is the interrelationship between the system concept—say an aircraft—and supporting functions such as communications or intelligence gathering. These dependencies have cost implications and need to be addressed during the early stages of design to avoid costly redesign resulting from ignored integration or accommodation issues. We must ensure any resulting requirements are fully understood and accounted for in the estimate.

For a life cycle cost analysis, support concepts need to be defined because they will not only impact the O&S elements of the estimate, but will also drive some of the acquisition costs. For example, if full contractor support is anticipated, the development and production costs may not require separate elements for data, training, support equipment, or spares if these are included in the contractor's annual rate. The O&S estimate, as well, may differ from that of an organically supported system. We must be sure to get input from the logistics personnel of the operating command when defining support concepts.

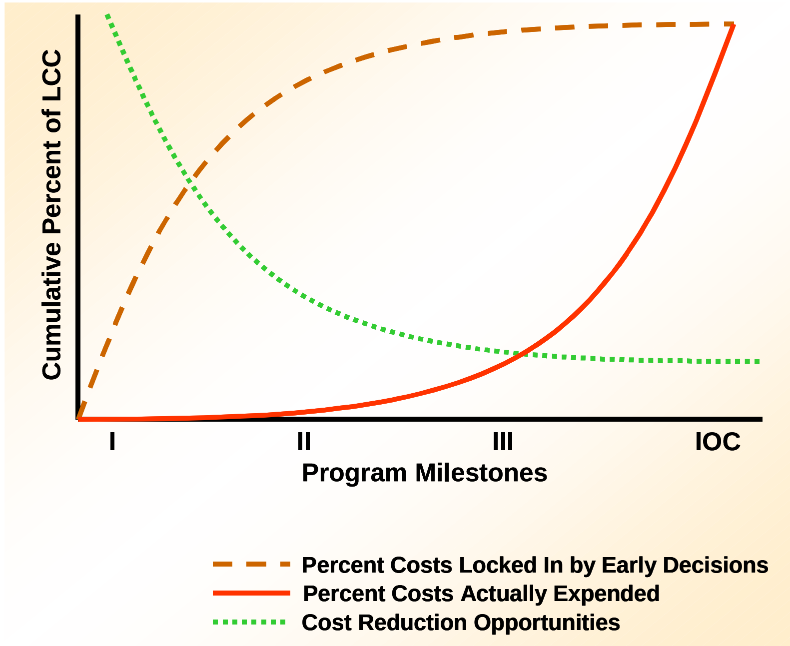

Figure 12-2 : Effect of Early Decisions on LCC

> ### Logistics Input
> **In one instance we completed our LCC estimate, only to find during the MAJCOM review that the logisticians disagreed with our support concept.**

Programmatic information will also be part of the cost estimate input. Development and production schedules and quantities will directly impact the time phasing of the estimate, particularly when then-year dollars (TY$) are applied (TY$ reflect the costs when the expenditures will actually be made and include the appropriate inflation factors). O&S and any military construction costs will also vary based on when the new system is introduced into the inventory and in what quantities it is ordered. For this reason, we must identify the cost impacts associated with development, test, and production schedules.

All the early decisions which must be made for a program are extremely important since they "lock in" significant portions of the ultimate costs for the system. Early cost and performance trade studies, for example, may influence alternative selection and the design and configuration that will ultimately be designed and produced. As Figure 12-2 shows, this occurs at a time when actual expenditures on the program are relatively low and opportunities to reduce cost are diminishing.

Develop Cost Estimating Structure

A Work Breakdown Structure (WBS) is a method of decomposing a system into successively more basic hardware or functional elements. MIL-HDBK-881B provides guidance on preparing a WBS, including specific WBS outlines and definitions for different weapon systems. Table 12-1 shows examples of a generic WBS for an aircraft system.

It's important to define the level of detail and structure of our estimate through the WBS process; it will not only influence what methodologies we can use, but also scope the time and resources required. The WBS process keeps everyone focused "on the same sheet of music." We also need to make sure the WBS level is consistent with the available data.

> **"It's called a "Work Breakdown Structure" because the work remaining will grow until you have a breakdown, unless you enforce some structure on it."**
>
> ***- Anonymous***

Table 12-1 : Sample WBS for an Aircraft System

Aircraft System	**Air Vehicle**	Airframe
		Propulsion
		Air Vehicle Software
		Armament
		Weapons Delivery
		etc.
	Systems Engineering & Program Management	*(no Level 3 breakdown)*
	System Test and Evaluation (T&E)	Development T&E
		Operational T&E
		T&E Support
		Test Facilities
	Training	Equipment
		Services
		Facilities
	Data	Technical Publications
		Engineering Data
		Management Data
		Support Data
	Peculiar Support Equipment	Test & Measurement Equipment
		Support & Handling Equipment
	Common Support Equipment	Test & Measurement Equipment
		Support & Handling Equipment
	Operational/Site Activation	System Assembly, Installation & Checkout
		Contractor Technical Support
		Site Construction
	Industrial Facilities	Construction, Conversion, or Expansion
		Equipment Acquisition or Modernization
		Maintenance (industrial facilities)
	Initial Spares & Repair Parts	*(no Level 3 breakdown)*

Develop Ground Rules and Assumptions

Once we've identified the purpose for the estimate and defined the preliminary cost estimating structure, we need to identify the ground rules and assumptions (GR&A) underlying the analysis. The GR&A will vary based on the specific purpose of the analysis. In the case of cost effectiveness analyses, we must ensure that the costing and effectiveness GR&A are consistent with each other. Some common areas that need to be addressed include:

- Cost basis of the estimate (specified BY$/TY$)
- Specific inflation indices used
- Treatment of sunk (already expended) costs
- Schedule issues including major milestones and significant events
- Basing, MILCON, and logistics concepts
- Personnel requirements

One ground rule that always applies to life cycle cost determinations is that costs are based on peacetime operations. Since system effectiveness is based on wartime analysis, integrating these two elements into a meaningful metric can be a significant challenge. It is essential that we ensure costs used in cost effectiveness studies are based on expected costs to acquire and support the system over its entire life cycle, not on wartime attrition for a hypothesized scenario.

It's especially important to make sure our GR&A meet the realism check. Manpower can be an area where unrealistic assumptions surface. In one Army study, the analyst was attempting to use an approach that compared alternative systems on an equal effectiveness basis. One system required an extensive build-up of manpower to support an increased force structure, but in a manpower-constrained environment that wasn't realistic. In another study, the system being analyzed also required additional personnel, but the analysts didn't include these costs because they knew adding more personnel wasn't acceptable. Of course, the new equipment was useless without the personnel to operate it.

> **"It's better to be approximately right than precisely wrong."**
>
> ***- Anonymous***

> ### Be Realistic
>
> **Make sure personnel assumptions are realistic—particularly in a manpower-constrained environment.**

Select Estimating Methodology

Once we identify the cost estimating ground rules and assumptions and any additional study assumptions and constraints, we're ready to select appropriate costing methodologies. Again, the methodology chosen will depend on the scope of the estimate and purpose for the final product. The extent of system definition, level of detail required, data availability, and schedule also have to be considered. Costing methodologies generally fall into the following categories: parametrics, analogy, and engineering build-up.

Parametric estimates are normally applied during a program's early stages when program and technical parameters are only grossly defined. Parametrics capture cost at a top level and require less detailed inputs than other approaches. This methodology often involves the use of cost estimating relationships (CERs)—mathematical expressions that predict cost based on a relationship between cost and another system variable.

Factors are also used to estimate areas such as training, data, peculiar support equipment, systems engineering, or program management (among others) when detailed analysis of these elements is not possible due to time or resource constraints or when data are limited or not available. Analysts typically estimate these costs by applying a percentage (factor) to a higher-level cost, such as prime mission equipment (PME). For example, in one study System Engineering/Program Management (SEPM) was estimated as 15% of PME where PME was $540,000. Therefore, SEPM was estimated at $81,000 ($540,000 x .15).

Analogy costing takes advantage of similarities between systems. This is possible due to the evolutionary nature of most weapon system programs. Analogy costing derives the new estimate by adjusting the actual costs of a similar system to account for complexity, technical, or physical differences between the two systems. This technique is normally used early in the program acquisition cycle when there are insufficient cost data but an increased understanding of the technical system design. Engineering insight is essential to ensure that we choose the right system for an analogy and that we develop the appropriate complexity factors.

The availability of actual cost data on the historical system is also an issue; without the data, we cannot estimate the new system. Estimates principally relying on parametric or analogy methodologies are typically rough order of magnitude (ROM) estimates, and we should portray the cost results at a corresponding granularity level—don't present a ROM cost at seven significant digits!

The engineering build-up method of developing a cost estimate is often referred to as a "grass roots" or "detailed" estimate and is performed at the functional level of the WBS. We use this method during a program's production phase when the system configuration is stable and actual cost data are available. The engineering build-up approach, by virtue of the level of detail involved, is a very time and resource consuming process.

There are various other cost estimating methods like catalog pricing for commercial-off-the-shelf items, use of expert opinion, or the use of engineering standards. For more detailed information on these techniques and others, refer to the AFSC Cost Estimating Handbook (AFSC is gone, but the handbook lives on).

In reality, we probably wouldn't use a single method across the entire estimate, but would employ a combination of several methodologies. It's important to apply other techniques as a validity check of major cost drivers. An example of a cost driver is PME, which is often used as a base for factor estimating. We also need to consider the unique requirements of each estimating task and any applicable assumptions and constraints when selecting methodologies.

There are many costing methodologies to consider. The element being estimated, the maturity of the technology or system design, the phase of the acquisition cycle, data availability, and

Unclear on the Concept

For years I worked with a cost analyst who provided cost estimates to the penny, literally. An estimate like $40,956,123.87 was the norm. The concept of significant digits was as foreign to him as the Aramaic language. In one instance he asked a fellow analyst to provide the formula used to generate a particular curve. He asked in all seriousness because "he couldn't read the curve with enough accuracy." This was for a curve that was likely in error by at least 20%-30%. I'd like to think he was the last of a breed. I'm not that optimistic.

sition cycle, data availability, and time and manpower constraints are factors to consider when choosing the appropriate costing technique.

As alluded to at the start of this chapter, program affordability must also be considered. Affordability has been defined as the degree to which the LCC of a program is consistent with DoD's long-range investment and force structure plans. Assessing affordability fosters greater program stability and avoids the cost growth resulting from program cuts or cancellations. DoD 5000.2-R stipulates that affordability be assessed at each milestone decision point beginning with program initiation (Milestone I). Programs must be consistent with the DoD Strategic Plan and based on realistic projections of likely funding available in the FYDP and the years beyond.

Model Results

In one program there were significant differences in software costs between the program office estimate and the independent estimate. It turned out a majority of the variance was based on which software estimating model was used.

Identify and Validate the Cost Models and Data

Along with selecting the estimating methodology, we must identify cost models. There are a variety of cost models to choose from, and it is important that the model be appropriate for the specific costing task.

The cost analysis community should generally accept the model. We can contact the Air Force Cost Analysis Agency (AFCAA) for assistance in identifying appropriate models. Their web page lists a variety of tools, models, databases, and studies. While AFCAA's web site describes models that have been successfully used in the past, it is ultimately the cost analyst's responsibility to determine the model's applicability.

Model Calibration and Validation

Model calibration and validation increase our confidence in an estimating tool. Calibration adjusts a model developed from data that does not necessarily represent the system we're estimating. Validation uses input data with known results to evaluate the model's predictive capability. Calibration is required before a cost model can be validated.

Collect and Validate the Data

Developing the database is as important as selecting the appropriate cost models. Collecting and processing historical cost data is an early and key step in developing a cost estimate.

Primary cost data, by definition, is found at the original source and, like other areas of research, is generally preferable to secondary data. The main sources for primary data are contractor reports or actual on-site data collection at the contractor's facility or appropriate government organizations.

There are, however, situations that dictate the use of secondary data. Some examples:

- It may be redundant to duplicate time-consuming efforts if a well-documented secondary source exists for the data
- Primary data are not easily accessible
- Sufficient time is not available
- Data are needed for a top-level test of reasonableness only

We need to evaluate the primary versus secondary data issue early in the cost estimating process. We can use secondary data if time, use, and availability make it the smart thing to do.

Historical cost data are the basis of the estimating process, but because they have some inherent limitations, we must have a thorough understanding of the data to ensure a credible estimate. Identifying limitations early in the data research phase can help avoid spending valuable time on data that are not applicable.

Develop the Cost Estimates

We are actually developing the cost estimates during each of the previous steps. At this point, we have a system description and a cost estimating structure. We have developed general and specific ground rules and assumptions and selected appropriate costing methodologies for each WBS element. We have also selected cost models and calibrated or validated them, if necessary. We've also collected cost data and validated it, normalizing it to the appropriate base year.

It is now that all our previous work comes together, as we input the data, execute the models, and apply the costing techniques to each WBS element. Once we've generated the estimate in base year dollars, our next step is to time phase the costs so that the estimate can be converted to TY$. In some cases, we may be required to perform net present value analysis (discounting) to evaluate the time value of money for different cost streams. We already identified applicable DoD inflation indices and discount rates in the ground rules and assumptions, and it's important to ensure that these are indeed the rates used. If for some reason we use different indices, we'll need to adjust the ground rules and assumptions so they are consistent with what we actually did.

Conduct Risk Analysis

Because a cost estimate is a prediction of the future, there is a significant chance that actual costs may differ from the costs we developed in the estimate. Risk and uncertainty analyses address this reality.

Risk and uncertainty are often used interchangeably, but they have distinct meanings. Risk refers to the "known unknowns" in an estimate or analysis; uncertainty addresses the "unknown unknowns." In general, cost estimating is considered to operate in the realm of uncertainty rather than risk, although in reality most estimates are a composite of both risk and uncertainty. In our discussion, we will use the term "risk" generically to cover both types of "unknowns."

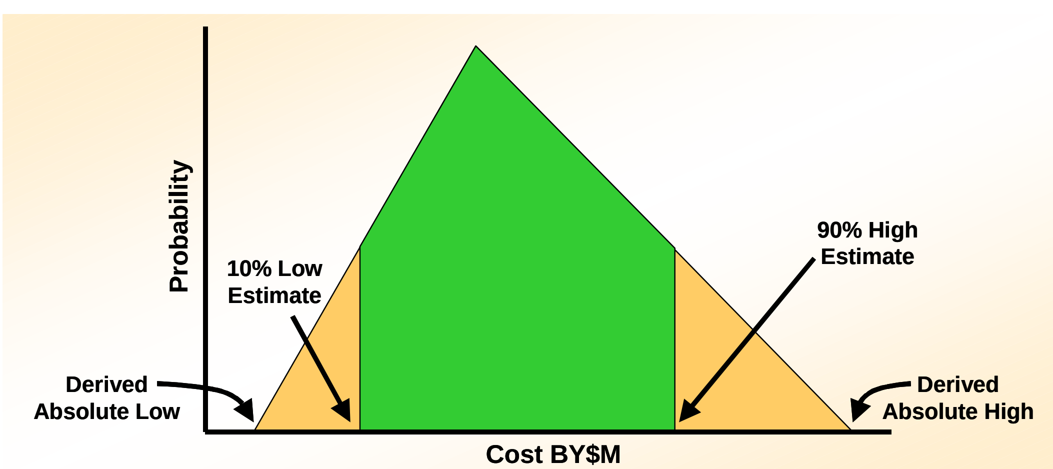

Figure 12-3 : Notional Triangular Probability Distribution of Cost

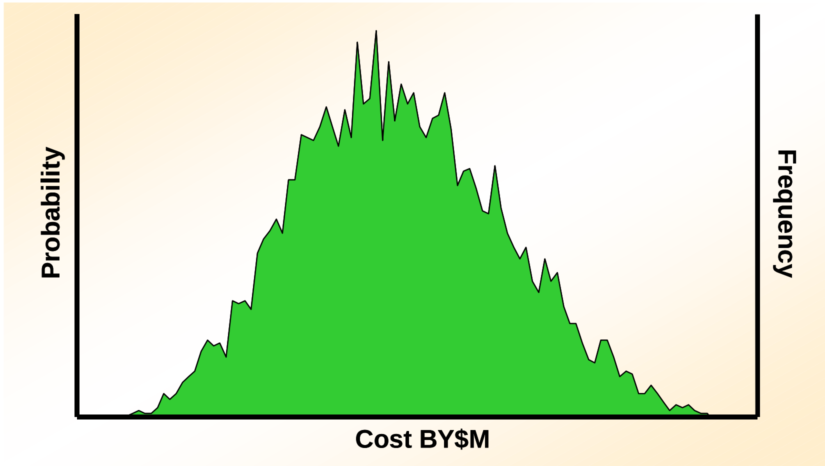

Figure 12-4 : Notional Frequency Distribution of Total Cost

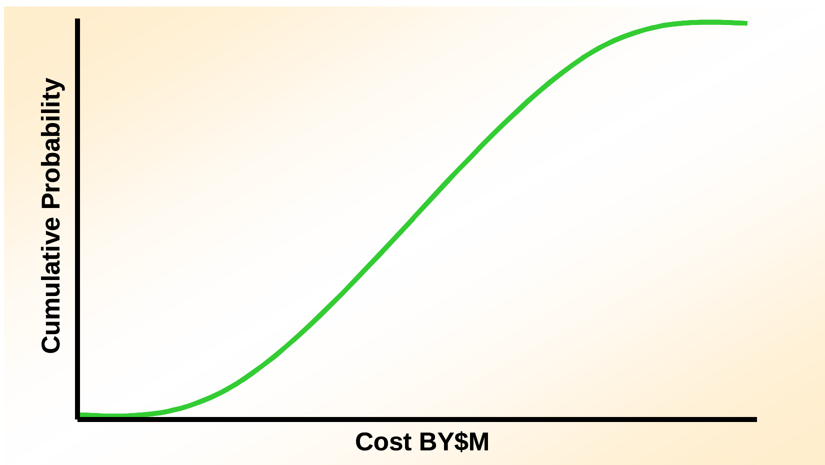

Figure 12-5 : Notional Cumulative Frequency Distribution

Risk stems from three primary sources: configuration changes, technical and schedule problems, and cost estimating error. Technical and schedule risk and cost estimating error can be accounted for in the risk analysis, but major configuration changes may require a new estimate rather than trying to compensate by applying a risk approach.

While the point estimate generated in the cost analysis is considered to be the "most likely" estimate of cost, it provides no information about risk. A confidence interval, on the other hand, provides a range within which the actual cost should fall for a specified confidence level. When both the point estimate and the confidence interval are taken together, they provide the decision maker with valuable information.

Several approaches are available to treat risk in an estimate, ranging from very subjective techniques to complex statistical approaches. One of the more qualitative techniques is Subjective Evaluator Judgment. This approach estimates risk by applying a percentage increase to the point estimate.

Another approach specifies the lowest and highest possible values in addition to its most likely value for each system element cost and sums to total system costs. This approach tends to misrepresent system risk since it is highly unlikely that all elements will be at the highest or lowest value at the same time.

A third approach quantifies risk by expressing cost as a probability distribution around an expected value. Triangular distributions composed of high, low and most likely values are most commonly used (Figure 12-3). Such a distribution is estimated for each cost element. Since it is unlikely that we will be able to estimate the absolute low and high costs, these points are considered to be the 10 percent and 90 percent values, respectively (the analyst's estimate of the most and least optimistic costs for the elements being estimated). The absolute low and high values are then derived.

Each triangular distribution is then treated as a population from which random samples are drawn using Monte Carlo simulation. Different random numbers are chosen for each cost element and the values of all cost elements are summed to arrive at a total cost. The entire process is repeated perhaps 100 to 1000 times. This simulation results in a normal distribution of random total costs and can be portrayed as in Figure 12-4.

A cumulative probability distribution (see Figure 12-5) is then used to identify the total cost including risk for a given probability level. For example, costs estimated at the 50th percentile indicate the total LCC (including risk) has an even chance of being greater than the estimated cost. Costs are often estimated at a second probability level as an excursion.

Confidence levels should be chosen prior to conducting the simulation based on how much risk we are willing to bear, not after we have the cost results and wish to limit the total cost to a specific value.

Where Did You Get That Number?

We were working on a program with an estimate that consisted of an Excel spreadsheet with no written documentation. Luckily, we were able to do a spreadsheet analysis of every cell to "create" the documentation. Not the easiest or most efficient way to document!

Complete the Cost Documentation

Documentation is one of the most essential elements of the cost estimating and analysis process. If an independent, qualified cost analyst can't replicate the estimate using the documentation provided with the estimate, it is clearly deficient and difficult to defend.

Often the documentation task is deferred until the estimate is completed. This is a mistake. While formal documentation may not be completed until the estimate is done, the documentation should be created along with the estimate. The ground rules and assumptions that form the basis of the estimate, the specific methodologies used, and also the rationale behind decisions affecting the cost estimates need to be clearly documented. Human memory can be a fleeting thing, and without written documentation of why and how certain approaches were taken it may be difficult if not impossible to recreate the data and processes. Many cost models facilitate documenting estimates "on the fly" by allowing methodology and WBS definitions to be input as the estimate is being generated.

AFI 65-508, Chapter 3, provides guidance on cost estimate documentation. Attachment 5 of the same instruction contains a documentation checklist to ensure all the required areas are covered.

Analyze the Results

At this point we have a fully documented cost estimate, including allowances for risk. While we may consider the estimate complete, we need to do additional work to generate a cost analysis. We need to examine the results to reveal cost drivers, or elements that drive other costs (which may be cost or non-cost parameters). As part of the analysis process, we measure how sensitive system costs are to these parameters by varying the cost-driving parameter(s) and observing the resultant cost effects. This sensitivity analysis will help us identify the trade space between system requirements and cost, and gives valuable information for the Cost As an Independent Variable (CAIV) process to help identify the "best value" solution.

In addition to performing sensitivity analysis of the cost drivers, we may find additional areas to investigate by examining different elements of the cost results (for example, unusually high or low costs). Often this provides insights not otherwise revealed. This part of the process is often overlooked and marks the difference between a cost analyst and a cost estimator. If we fail to analyze the estimate, the job is only half done.

> **The "So What" Factor**
>
> **A good cost analyst doesn't just produce numbers—a good analyst interprets them!**

Present and Defend the Results

The final step in any cost analysis is to present and defend the results to the decision maker. It's important to remember the purpose of the analysis is to provide information, not to make the decision. However, in providing that information, we must communicate clearly what the cost analysis represents, explain and defend the methodology we used, and defend the estimate with a good rationale for the ground rules and assumptions that underlie the estimate. Chapter 13 of this handbook discusses briefing and presentation techniques.

> **"When a technical analyst makes a presentation everyone but the analyst believes the results. When a cost analyst makes a presentation no one but the analyst believes the results."**
>
> ***- Anonymous***

Cost Summary

We've discussed why good cost analysis is critical to various types of analysis. Costing is a complex process that requires special skills and traits. We've identified some of these and outlined a process that a cost analyst would typically follow to generate a complete and robust cost analysis. In addition, we've discussed methodologies, tools, and techniques available to the analyst and tried to identify some of the pitfalls based on our experience in the field.

Although demands for competent and complete cost analyses are increasing, experienced cost analysts are a diminishing resource. Many experienced analysts are retiring from government service; others are being diverted to the pressing demands of budget formulation and execution. We are increasingly dependent on contractor support to conduct cost analysis, and expertise in this area varies from contractor to contractor. It is our hope that the information we've provided here will help bridge the gap.

13 Just the Facts, Ma'am

Briefing Your Results

Briefings are the principal means of communicating ideas within our community and the community we serve. A briefing is an exchange of both visual and spoken information. Slides provide the images; the briefer supplies the words. Happily or not, from time to time all analysts must prepare briefings and sally forth to present them. Given the importance of these presentations, one would think that briefers and their organizations would insist on presenting clear, concise, and entertaining briefings tailored to their audiences. All too often it doesn't happen.

Some briefing problems are unavoidable. Even the most conscientious analyst is beset by the need to meet deadlines and by conflicting priorities. But worse than these realities are the avoidable failures: the failure to anticipate the needs of the audience, the failure to properly structure the briefing, the failure to follow good briefing practices. How can we prevent these preventable failures? In this chapter we will focus on a goal of making every briefing a positive experience for all.

Deciding What to Present

There are two interrelated questions that need to be asked before a single briefing slide is ever made:

- Who am I briefing? This is normally the highest-ranking member of the audience.
- What do I want to accomplish with the briefing…is this an information briefing or a decision briefing?

The answers will determine briefing content and structure. When you brief your analytic peers they will be interested in the results and precisely how you got them. Brief decision makers higher up the food chain and they will (generally) be interested in top-level results and being convinced that the results were obtained using good data and analytic practices. Should these be the same briefing? Get real.

So how do we use the answers to these questions to decide what is appropriate? The first step is to empathize with your audience. Try to anticipate why they need your information. Do they want to critique your MS&A? Do they need to approve the validity of your results prior to briefing a wider audience? Are they going to make a decision based on your results? If so, what type of decision—reorganize, buy new materiel, change operational procedures?

With your new insight, you can determine the content and structure of your briefing. For example, imagine an audience of decision makers determining their degree of support for a proposal. As part of your briefing, you think it appropriate that they receive information on the input data to your analysis. Do you provide the data? Or do you provide the source of the data? Without specific information to the contrary, it's a no-brainer. Present the sources. They speak to the legitimacy of the data, which is often the decision maker's most pressing data issue.

In a like manner, every potential slide in the briefing can be examined for appropriateness. Of course each decision you make is a judgment, but if you don't ask yourself the question and receive feedback on your decision, it's unlikely your judgment will improve.

Briefing Goals and Techniques

Be Brief

Have you ever seen a briefing that wasn't brief? Of course you have. And long briefings can be appropriate—when they tell the audience only the essentials they need. Your goal is to make your briefing as long, and only as long, as necessary to convey the needed information. Your audiences won't criticize you for doing a good job too quickly. If you have misjudged what is essential, your backup slides are your insurance against your briefing being too brief.

Tell a Story

Briefings frequently don't tell a story. They start by promising one thing, present another, and summarize who knows what.

Imagine a joke with these flaws: Did you hear the one about the farmer's daughter and the traveling salesman? The traveling salesman walks into a bar with a talking dog and the farmer's daughter said "we don't serve dogs in this bar."

Farfetched? You're still a tenderfoot.

The first step in telling a story is to think of the briefing not as individual slides, but as a whole composed of three elements: a beginning to set the stage, a middle to develop arguments and/or present facts, and a summary to pull it all together. With these elements in mind, you can consider some less concrete aspects of story telling.

Flow

The flow of a briefing pertains to its seamlessness. A briefing that flows well has an apparent directed movement—metaphorically the briefing sweeps you along. You can identify a well-flowing briefing by the audience's anticipation of succeeding slides. This anticipation simplifies the briefer's job because the audience is looking for what's coming. Techniques for improving the flow of a briefing are conceptually simple:

- Build the briefing around a unifying theme and explicitly present the theme as often as necessary. This can be as simple as a slide that illustrates your progress through the briefing.
- Telegraph upcoming ideas and topics to prepare the audience in advance.
- Remind your audience how your current topic relates to what they have already seen in order to reinforce the continuity of your presentation.

> **"The most beautiful thing we can experience is the mysterious. It is the source of all true art and science."**
> ***- A. Einstein***

Mystery and Suspense

All good stories are mystery stories. Often it is possible to structure your briefing to create suspense, laying out facts to support an initially unknown conclusion. This is valuable because it keeps the audience focused on your briefing. Look for ways to create suspense for your audience, but be careful not to misdirect them in the process. We may be human, but we are all as curious as cats.

Be Understood

In preparing a briefing you must make assumptions about the level of knowledge of your audience—how prepared they are to absorb your story. These assumptions are dependent on knowing (once again) whom you're briefing, their background, and their desire for detail. This usually requires some research. Obvious sources are official biographies, others who have briefed the same audience, executive officers, etc. Often—but by no means always— the higher your briefing goes, the less likely it is that your audience will desire the details of your work.

In harsh terms, if your audience doesn't understand your briefing, you have wasted everyone's time. At the other extreme, too much detail will put insomniacs to sleep. What's a poor analyst to do?

A good option is to be slightly conservative in your assumptions—don't give the audience quite as much credit as they might deserve.

During the course of the presentation you will get clues about the accuracy of your decision. One sign of understanding is an audience that asks intelligent questions. Another is the General prompting you to keep moving—through your briefing, that is. Of course, there can be negative indicators as well: requests to repeat yourself, questions from left field, puzzled looks. In either case you should attempt to make the necessary adjustments to your spiel on the fly. You may elect to provide more or less detailed remarks. You also might bypass slides or resort to your cache of backup slides. However, understand that these techniques will not save a poorly structured briefing.

Be Believed

You're sure your audience will understand your briefing. Will they believe your message? Again the key is anticipation—anticipation of what you need to present to convince your audience that you knew what to do, and furthermore, that you did it. Ideally, you want to be able to say, believe, and defend the following:

- Here are our assumptions and this why they are the right ones
- Here are our scenarios and this is why they are the right ones
- Here are the models we used and this is why they are the right ones
- Here are the data and data sources we used and this is why they are the right ones
- Here is how we did the analysis and this why it was the right way

The word "right" means that overall your selections were best for the job at hand. Your choices may have been constrained by schedule, funding, degree of support from participating organizations, or the expertise of the analysts. But if you are prepared to make and convincingly defend these assertions, an open-minded audience should believe you.

There are other factors at play as well. Patience in answering questions and a willingness to admit uncertainties when they exist help to build credibility. It is also a plus to have previously established credibility with your audience through previous presentations.

Fifth Principle of OR

If it were easy, someone would have done it already.

Don't Oversell

Nothing destroys rapport with an audience faster than overselling your work. There are two ways of doing this: 1) have a glib answer for every question and 2) exaggerate the accuracy of your results.

Glib answers indicate insincerity and a lack of understanding of the complexity of the problem. The unvarnished truth is that complex problems rarely if ever have easy answers. Marketers most often ignore this.

Exaggerating the accuracy of results by overstating the number of significant digits in an answer indicates a lack of understanding for the limitations of the supporting analysis. We have all seen inputs with one significant digit of accuracy produce outputs with multiple "significant" digits. Thus we're told that System A at 0.49 is better than System B at 0.46, when in fact the analysis should show them indistinguishable at 0.5.

Classic Advice

Tell 'em what you're going to tell 'em; tell 'em; tell 'em what you told 'em. Simple as it is, it works.

Hammer Your Important Points Home

Typically an audience will leave a briefing with at most one or two of your ideas tucked away in their heads. It is your responsibility to ensure that they are the right one or two ideas. There are various techniques for accomplishing this. Among the most successful are emphasis, repetition, graphics, analogies, and anecdotes. Don't bury the critically important points amid lesser ones.

Don't Invite the Wrong Questions

Every briefer likes questions. It is a sign that the audience is interested. However, you don't want to your briefing slides to invite the wrong questions. Almost invariably these "invitees" take you and your audience into areas that dilute your message and lengthen your presentation. You will know you have a problem when your first response to a question is "Why was *that* question asked?"

Unfortunately, the chances are very good that you invited the question by omitting something crucial on the slide or by using an undefined or ambiguous word, phrase, bullet, etc. Try to ensure that the question doesn't get asked. At every step of your preparation, ask yourself if you are clearly saying or displaying what you intend. And pay attention to your reviewers' comments. This is a type of problem that is easier to spot in other peoples' briefings than your own.

Appearance of the Briefing

The appearance of a briefing is important. A poor choice of fonts or font size, incorrect punctuation, insufficient white space, lack of parallelism in lists, and visual distractions can all reduce the effectiveness of your presentation. There are simple techniques for minimizing these problems.

Font and Font Size

The body text of this document appears in 12 point Times New Roman font. The Times New Roman font is a serif font, which means—for practical purposes—that it has small, unneeded strokes on many of the letters. Serif fonts are common in printed materials and they pose no problem on the printed page. However, on briefing slides they are noticeably harder to read than a sans-serif font like Arial. For this reason, Arial makes a good font choice for briefing slides.

Obviously, font size can also affect the comprehension of material. For textual information appearing in bulleted lists, it is good practice to use at a minimum a 20-point or 24-point font. For very special situations (i.e., almost never), 16-point or even 14-point fonts can be used. However, realize that even under the best of conditions (a dark briefing room and a sharply focused, bright image) such small fonts will increase the time needed to comprehend information. This information may also be unreadable in all except the front few rows.

Lists

Punctuation and Capitalization

Terminal punctuation is normally not used for list items. Initial words of items should be capitalized. Within items capitalize only what would normally be capitalized, e.g., proper names and acronyms. THE USE OF ALL CAPITAL LETTERS GENERALLY MAKES TEXT HARDER TO READ.

Parallelism

One of the most common problems found on briefing slides is a lack of parallel grammatical structure in list items. Thus, the first item of a list might lead with a noun and be followed by a verb, effectively forming a clause. The second item might lead with a verb and be followed by noun functioning as a direct object. The third item might a single adjective. This type of inconsistency slows down comprehension of the information: it forces the audience to mentally shift gears for every item.

Here's an example of a list lacking in parallelism:

- Analysis is an attempt to predict the future
- Decision makers find analysis useful
- Analysis should precede a decision, not follow it
- Can't be rushed
- Analysts mix art and science

In this example, there is no consistency of subject—analysis, decision makers, analysts—and the basic subject, analysis, is used once as a direct object and is omitted in the last bullet.

As a contrast, an edited version of the same list with consistency of structure is shown here:

Analysis:
- Is an attempt to predict the future
- Is useful to the decision maker
- Should precede a decision, not follow it
- Can't be rushed
- Is part art, part science

It is easy to discern the improvement in ease of understanding. Avoiding this problem is a big deal to a conscientious briefer.

Text

Most briefing charts consist of text, usually in a bulleted format. The biggest issue to be faced in preparing such slides is that of deciding how completely to express ideas—how many words to use.

The rule: fewer is better.

Brevity improves comprehension by reducing extraneous material, by increasing white space, and by allowing the audience to fill in the holes.

For example, suppose a slide is titled "Early Rocket Developments." Instead of the bullet: "Germany developed and used both the V-1 and V-2 during World War II," we could reasonably use "V-1, V-2 (WWII)" or even "V-1, V-2." The audience can be expected to know it was Germany and that these weapons were WWII vintage. In this example, the audience would have to read and process the long version, while in effect they would be able skip the shortest version without missing a beat.

There can be exceptions to the "fewer is better" rule when the added words are not likely to be extraneous. The most important exception occurs when the slides must be self-explanatory when distributed in hardcopy.

Graphs and Tables

Graphs and tables are used to portray quantitative results. When they clearly and quickly convey their information they are part of the solution. When they do not they are part of the problem. Clearly we want them to be part of the solution, and in this section we present simple guidelines for making that happen.

Graphs

Graphs come in many varieties: line, bar, stacked bar, pie, scatter, and 3-D surface just to name a few. Following the "rules" below will improve any graphical briefing chart.

- Design every graph expecting that it will be distributed in black-and-white hardcopy. Without your words to go with it, information not explicitly visible will be assumed—probably incorrectly.
- Design graphs to minimize extraneous information. Extraneous information must be processed before it can be excluded, thereby acting as a distraction.
- Place the caption below the graph.
- Select the ranges of scales to maximize use of the graph's area. An exception occurs when similar graphs will be compared; in such cases identical graph size and scales should be used.
- Label axes in a large, easily readable font. Yes, Arial. And label the ordinate (y-axis) with horizontal text at the top of the axis.
- Label each scale in a similarly easy-to-read font.
- Use appropriate scale units. Again, know your audience. Different people may expect different units. Metric should be first choice. This is Air Force policy (followed more in the breach than the observance).
- For each axis include the zero or other suitable reference point on the graph. Failure to include a relevant reference point can distort audience perception of the presented data.
- Try to avoid the use of logarithmic scales. Many people are not comfortable with log scales. However, do not avoid them when absolutely necessary.
- Label curves and points directly with text, rather than using a legend. Use a different color and style for each line or type of point plotted. It is time-consuming for the audience to have to constantly refer back and forth between a legend and the data.

True Story

In displaying the data from a complex analysis we worked more than two weeks refining the format of the all-important tables that summarized the results. The time was well spent. We successfully used the tables in briefings to every level of audience, including the Deputy Undersecretary of Defense for Space. The same tables appear in black and white versions in the final report.

Tables

Here are a few rules for tables:

- Design every table expecting that it will be distributed in black-and-white hardcopy. Without your words to go with it, information not explicitly visible will be assumed—probably incorrectly.
- Design tables to minimize extraneous information. Extraneous information must be processed before the audience can exclude it, thereby acting as a distraction.
- Place the title above the table.
- Simple tables with a maximum of three columns and eight rows should not have column and row dividers (an arbitrary, but reasonable choice).
- Larger tables should have row and column dividers.
- Use an easily read font. Arial is again preferred.
- Use as large a font as is practical, realizing that white space around text enhances its readability.
- Use shading, coloring and other accents to improve audience comprehension of results.

Visual Techniques

The advent of the personal computer and presentation software has given every analyst the ability to create slides that once would have required an illustrator working in an elaborate graphics shop. Used thoughtfully, available graphics enhancements help make your points. Overusing them reduces their impact and may give the impression that the medium is the message. Thus when it comes to using visual techniques, less (use) is usually more (impact). After all, what is important if everything is highlighted?

Accents

Italics, bold fonts, underlining, shading, color and arrows are some of the options available for adding emphasis to a briefing slide. Some of these options are useful for accenting words; others can be used to highlight graphical components.

Using Background Color

The use of a background color on slides is seen by one faction as a classy adjunct to the slides. There is a counter faction that regards the use of background color as a distraction to be overcome by the audience. In practice, a background color is probably omitted more often than not. Not having a background color eases the job of making copies of the slides on a copy machine.

Using Clip Art

There is little question that the use of appropriate clip art helps a briefer make points. It also adds entertainment to a briefing. What must be avoided is clip art that makes the audience question its relevancy. This is a distraction.

Animated Slides

Animated slides can enhance understanding by focusing an audience's attention on the animated items. The animation can take many forms, including sound, but it is perhaps best used to present information that is sequential in nature. A drawback to animation is that it can be invoked only in computer presentations. It can also easily be used to excess, when—you guessed it—it becomes a distraction.

Practice Makes Perfect

We had performed most of the analysis, but the job of briefing it to the august review group went to the fellow heading up the analysis team. His only dry run came the afternoon before he was to brief. Impossibly inept, he couldn't get the facts straight and he hemmed and hawed at every opportunity. Disaster loomed. The next morning we were ready to hide when he took the floor. Amazingly, he pulled off a near-perfect performance that was accurate and entertaining. Bad rehearsal, good opening.

Using Color to Convey Information

Color, in addition to being a great accent, can often be used effectively to convey quantitative or qualitative information. Typical is the use of the colors green, red and yellow—the colors of a stoplight—to indicate "good," "bad," and "marginal" in a table. Keep in mind that some people are colorblind and cannot distinguish between red and green. This suggests use of a letter, or perhaps a check mark, as a secondary indicator, which is also useful when the slide appears as black-and-white hardcopy. The use of color is one visual technique that may be underused.

Review of the Briefing

A briefing represents you, your organization, and possibly a much larger community. It needs a thorough review. The definition of thorough will vary with your experience at preparing briefings, the intended audience, and other factors that will change with circumstances. Given these caveats, a reasonable review scenario for important briefings can be postulated.

The earliest reviews should be the ones you and your team members perform as the briefing is built. The next reviews will come from coworkers who are not on your team. You usually select these reviewers, and their reviews are typically done with paper or e-mail copies of the briefing. Following this, there may be a walkthrough or dry run during which the slides are projected and critiqued individually and as a package. Attendance at this stage will depend on organizational factors, but your immediate boss is a good candidate to be present. Finally, if everything has gone well a full-blown dry run will follow. From here you may go through a series of coordination briefings both within and outside your organization. Reactions to these briefings can generate everything from minor revisions to burying the briefing (if not the briefer) where the sun don't shine. With the reviews successfully finished, you can finally brief the ultimate customer(s).

The review process can be contentious, but you and your reviewers share the common goal of creating a solid briefing that meets expectations. The issues discussed most often are ones of content, organization, and format. Also, there are sure to be substantive questions raised about your slides and words. Take these questions to heart and try to eliminate the reason they were raised. Often a minor improvement to a slide or your explanation of a slide is adequate. For questions that you cannot answer, find the answer and share it with the asker. Finally, your briefing habits and style may also be critiqued if they are perceived to be detracting from the briefing. Buck up, we've all been there.

True Story

At one briefing I gave I had to endure two hours of acid comments about a study I had led. It was an ugly experience, but nothing short of a different conclusion would have pleased the audience. My vindication came in the form of warm receptions from the audiences who counted. Moral: When your audience doesn't like your message, nothing helps.

Presentation Techniques

The proof of the pudding is in the eating; the proof of the briefing is in its presentation. A good presentation doesn't happen by accident. It happens because of good preparation and good delivery. There are both things to do and to avoid.

Motherhood

Your mother knew most of the secrets to giving a good presentation. In case you have forgotten:

- Practice your presentation before going public, preferably where you will brief
- Be confident (practice builds confidence)
- Be polite
- Stand up straight
- Speak distinctly
- Make eye contact with the audience
- Keep your hands out of your pockets
- Take all questions seriously (if you can't immediately answer a question, offer to follow up later with a reply)
- Graciously acknowledge any shortcomings identified in your presentation
- Don't pretend to have all the answers (no one will believe you)
- Always bring extra electronic and hard copies

Briefing Individual Charts

Every chart in a briefing should appear for a good reason and consequently deserves to be understood by the audience. This means that every chart should be on the screen long enough to be understood. It is your job as briefer to minimize that time by briefing the material on the chart—what else are you there for? Below are suggested techniques for handling each of the three fundamental types of briefing charts: text, graph and table. The goal is to call attention to the important material on each chart in a logical sequence.

Text

Begin with a one-sentence overview of the purpose of the chart (consistent with the title of the slide), then brief the slide from top to bottom. For each bullet you explicitly call out, either paraphrase the material or read it explicitly and then amplify on its meaning, source, significance, etc. Both techniques add information and speed assimilation of the presented idea.

Graph

Take time to orient the audience when briefing a graph. Begin by stating what the graph shows: "This graph plots the cost of sturdleys as a function of time." Next, explicitly discuss the units and scales represented on the axes. If it is not already obvious, identify the source of the displayed data. If there are multiple curves on the graph, discuss what each represents. Explain why the data behave as they do: "Sturdley costs decrease with time because…"

Table

As with a graph, the same logical presentation of the features of a table should be developed. (Is there a pattern here?) What does the table show? "The table shows sturdley usage for each MAJCOM every third year since 1980." What appears in the first column? What do the other columns represent? Do an example. "In this column we see that the Air Force used 95 sturdleys in 1986."

Is all this necessary? Is your audience too dumb to see all these things themselves? Probably not, but comprehension takes time and your well-chosen words are value added. A logical presentation of facts simply guides your audience more quickly to where they would eventually arrive on their own. Obviously, when a series of similar slides appear in a briefing, repeated detailed orientations are not necessary.

The End

The last few slides of a briefing are critical to its success. It is time to summarize the briefing and in many cases present possible courses of action. The first rule of summaries: Everything appearing in a summary must have been presented previously in the briefing. It is absolutely unacceptable to introduce material in the summary; it seems foolish to say this, but it is done with great regularity. Do not do it!

14 Document? I'd Rather Eat Nails!

Handing Your Work to the Future

If analysts have one great flaw, it is their reluctance to properly document their work. A million dollars gets spent on a study and its legacy is often a set of briefing charts (or annotated briefing charts if we're lucky). To what end? Within weeks, details of a completed study are lost if there is no documentation: sources of data, ground rules and assumptions, engineering data slip away. Within a year, except for studies with major impact, the results are lost—or become suspect because of age or conflicts with newer studies. Even the most influential study, if undocumented, has a useful life typically determined by the memory or longevity in office of a decision maker. What is the problem? Why are we so reluctant to hand our work on to the future?

Physicists do it. Chemists do it. Mathematicians and engineers do it. Why are we so poor at providing comprehensive documentation of our research? Not the research into methods and techniques—our academics do well at this— but the important details of the analytical studies we undertake. Are we ashamed of our work? Is it not useful to know how we dealt with issues that will come up again and again? Is it not worthwhile to be able to look back and be able to understand why this study reached one conclusion and that study reached another?

Answer 1: We ran out of time and money. OK, now make a case to your management explaining why that's not a good excuse. Management more than anyone should understand the need to document (and if pigs could fly…).

Answer 2: It was a quick and dirty study and wasn't worth documenting. Maybe it was quick and dirty. The trouble is that today's quick and dirty study often points to tomorrow's major issues. Also you would be surprised at what decision makers sometimes bring up from earlier briefings.

Answer 3: We don't want anyone to know what we did (not advertised in public). Sounds like you didn't understand the first thirteen chapters. Go back to Chapter 1,do not pass "GO," do not collect $200.

Answer 4: I don't like to document (also not advertised in public). This is not a matter for personal preference. Consider a new career.

> **Sixth Principle of OR**
>
> **A study not documented is a study not done.**

Document What?

There are two types of study documentation: internal and external. Internal documentation consists of the nitty-gritty details of what you have done as an analyst during the study. Internal documentation is normally kept within the study team. Aspects of it were discussed briefly in Chapter 8. External documentation is the published documentation available to anyone with "a need to know." It is the documentation that circulates outside the confines of the study team.

How Much is Enough?

Internal Documentation

Most internal documentation is informal. It consists of personal notebooks, memos and e-mails, working papers, computer runs, computer programs, informal computer program documentation, phone lists and logs, interim briefings, etc. In short, it is the detailed record of what you did and thought during the course of the study. Generally, this documentation is filed away "as is" at the end of a study, perhaps organized by date or subject. Some of it may have enduring value. Working papers, phone lists, and interim briefings are possible examples. However, the useful life of most of this material is usually short, and only rarely will you be called upon to consult it. However, rarely is not synonymous with never, and it is therefore a good idea to keep all internal documentation for at least two years.

External Documentation

External documentation takes many forms: memoranda, letters, briefings, annotated briefings, reports, and journal articles. The issue is not what the documentation is called, but its sufficiency for future uses.

If a study is, in fact, a quick and dirty effort, it needs substantially less documentation than a yearlong multi-million-dollar effort. But it still must be adequate to allow critical questions to be answered over its useful life.

Matching the Useful Life of Analysis to the Form of Documentation

The useful life of an analysis can be thought of as the time over which its information is considered worth consulting. For example, the useful life of the explanation of a new analysis technique is long—years or decades. On the other hand, the useful life of that quick and dirty study we've mentioned will be short—perhaps as little as a few months. Given knowledge of the useful life of the information we are documenting, we can select a format for the documentation that has a corresponding useful life. Thus a journal article—which has an almost "infinite" useful life—is appropriate for documenting a new analysis technique, while a memorandum or annotated briefing better matches the quick and dirty study.

Letters and Memoranda

Letters and memoranda by their nature are short and not widely circulated. These attributes give them a very limited useful life. They are suited to documenting only the simplest of analyses, typified by the one or two week quick look.

Briefings

A briefing without annotations generally has a very short useful life. You can easily demonstrate this by attempting to brief an unfamiliar set of slides. You may understand the major points of the briefing, but you will not be able to address any aspect of the study not specifically appearing on the charts (and many that do). This makes it impossible to recommend using unannotated briefing charts for any final documentation.

A well-annotated briefing is a different matter. By adding slides with appropriate explanations, an annotated briefing can document any level of detail. Given that an annotated briefing is easier to produce than a formal report, it makes a good alternative for documenting studies of intermediate size and complexity. The key to this format is an adequate level of detail in the slides themselves, accompanied by well-written annotations. A poorly annotated briefing is just an unannotated briefing that uses twice as much paper.

Formal Reports

When you want to tell a coherent story, address all aspects of the study, and provide significant details not suited to other formats, the written report is the solution. A report can also serve to document techniques and methodologies that might not be suitable for journal articles (see below). A written report can have a useful life of decades. Reports should be submitted as a matter of course to the Defense Technical Information Center (DTIC). DTIC archives thousands of reports every year, cataloging them in a searchable database and providing copies to qualified DTIC users.

Journal Articles

Journals are the lifeblood of academia, and as such are hoarded by libraries for future generations. Thus, a journal article has a life comparable to that of the libraries that hold them. A journal article is generally much shorter than a report, but it has the advantage of being refereed (reviewed in detail) by disinterested third parties. All these factors make the article great for documenting academic OR contributions, but not well suited for documenting OR analyses. The techniques and methodologies of academic contributions are fundamental OR building blocks and can remain of interest indefinitely.

"Self-Documentation"

There seems to be the hint of a trend toward analysis software that allows the analyst to document on the fly. Introduce a variable and the software provides the opportunity to define it. Enter a formula and the software lets you describe it. Write lines of code and comment as you go. If this trend continues, we may see a revolution in how analysis is documented—especially internal documentation.

Don't Shade Your Eyes, Plagiarize (from yourself)*

It is not only acceptable to steal from yourself, it is the thing to do. When documenting a study, use every idea, word, picture, and table you can lift from your previous briefings, reports, and working papers. There is no glory in recreating or unnecessarily embellishing what is already satisfactory.

**With apologies to Tom Lehrer*

Document What?

OK, you've bitten the bullet. You've bought into the idea that you need to document your study (but you still don't like it). What do you need to document? What topics do you need to address? How deep is deep enough? The best way to approach this is to be guided by the question: What needs to be said so that five years from now (or ten or twenty) a reader of the report can answer reasonable questions about how the study was done?

As a guide for our discussion, we will use the AoA final report outline taken from the AoA Handbook:

Executive Summary
1. Introduction
   1.1. Background
   1.2. Purpose
   1.3. Scope
2. Acquisition Issues
   2.1. Mission Need
   2.2. Scenarios
   2.3. Threats
   2.4. Environment
   2.5. Constraints and Assumptions
3. Alternatives
   3.1. Description of Alternatives
   3.2. Nonviable Alternatives
   3.3. Operations Concepts
4. Determination of Effectiveness Measures

4.1. Mission Tasks
4.2. Measures of Effectiveness
4.3. Measures of Performance
5. Effectiveness Analysis
5.1. Effectiveness Methodology
5.2. Models, Simulations, and Data
5.3. Effectiveness Sensitivity Analysis
5.4. Effectiveness Results
6. Cost Analysis
6.1. Life Cycle Cost Methodology
6.2. Models and Data
6.3. Cost Risk Methodology
6.4. Life Cycle Cost Result
7. Cost-Effectiveness Comparisons
7.1. Cost-Effectiveness Methodology and Presentations
7.2. Cost-Effectiveness Criteria for Screening Alternatives
7.3. Cost-Effectiveness Results
7.4. Selection of Preferred Alternative(s)
8. Organization and Management
8.1. Study Team/Organization
8.2. AoA Review Process
8.3. Schedule
A. Acronyms
B. References
C. Other Appendices as Necessary

While this outline is tailored to AoAs, it serves well enough as a generic example of final study documentation. And what applies to the final documentation is generally equally applicable to the study plan.

Executive Summary

Any lengthy study documentation should be preceded by an executive summary. An executive summary should be only as long as necessary. As a rule of thumb, it should not exceed three pages. It should be a summary of the study as a whole, carefully written to contain only the most critical information. Anyone reading such a summary should come away with a clear understanding of why the study was done, what the study did, how it was constrained, and any important conclusions. An executive summary is best written by an accomplished writer totally familiar with the study. Because it is the only thing most readers will read, it should be reviewed repeatedly to insure that it presents a succinct and accurate picture.

> **An Experiment**
>
> **Pick up an issue of Scientific American and select an article in an unfamiliar area. Skip the first four paragraphs and start reading. Odds are you'll have serious trouble putting what you read into a meaningful context. That's why we write introductions.**

Introduction

Every stand-alone study document needs an introduction to set the stage for what follows. What is the big picture? What is the motivation for the analysis? How does the analysis relate to the motivation? How does this analysis fit within the big picture? In short, the introduction contains the study's history and its goals, essentials needed if the reader is to appreciate the details to follow. Without this, the reader will be lost.

We might think that the need for an introduction is obvious. Unfortunately, most inexperienced authors assume that everyone shares their background knowledge. Please believe me, they do not!

Issues

While the Introduction basically describes the pre-study environment, the Issues section discusses the study planning and execution environments. Issues discussed in this section generally have study-wide influence. They define the area of the study and its boundaries. A representative list of issues is given below. Some issues may be expanded upon in other sections of the report.

- Who authorized the study and expects the results?
- What is the study's chain of command?
- What schedule must the study meet?
- What resources will support the study?
- What are the study's constraints and assumptions?
- In what place and time is the study set?
- What script will events in the study follow?
- Are there threats to be considered? What are they?
- How will the physical environment be handled?
- Are interactions with other organizations or systems to be considered?
- Is costing to be done? What kind?

Alternatives

In the very first paragraph of Chapter 1, we stated "OR's goal is to provide information for making decisions." Decisions imply alternatives, and in truth all studies look at alternatives. The Alternatives section of the documentation allows the alternatives to be identified and described. If the alternatives are simple, they can be completely described in this section. If they are more complex, the descriptions should present an overview that includes (as appropriate) basic concepts of operation. In these latter instances, the details of the alternatives are usually presented in an appendix.

The omnipresent alternative is to do nothing based on the results of the study. The other alternative(s) is (are) to do something. That "something" may be as unimaginative as initiating another study or as mundane as changing the way paperclips are purchased. It also can be as significant as deciding to pursue a technology development program or initiating a change in national strategy. However, no matter what the alternatives are, they need to be explicitly acknowledged. They are, in effect, study constraints and they have affected how the study was structured and proceeded.

Effectiveness Measures

Every decision is made in the context of a value system and based on available information. It is the job of analysis to provide information that is relevant, accurate, consistent, timely, and complete. To ensure that these goals are served, alternatives are measured one against another within a framework of effectiveness measures. These measures were discussed in detail in Chapter 6. They consist of mission tasks, measures of effectiveness, and measures of performance. These measures need to be completely defined and discussed in the documentation. It is necessary that the reader be as convinced as you that the right measures were selected.

Effectiveness Analysis

As the saying goes, this is where the rubber meets the road. The effectiveness analysis documentation is likely to constitute the major portion of a report (and receive the greatest scrutiny). For example, consider the numbers in Table 14-1. They show the page breakout by topic of the final report of a major study with which I was associated. Almost one half of the pages of the report dealt with the effectiveness methodology and results.

The effectiveness methodology consists of the effectiveness measures and the methods of their calculation. In discussing the methods of calculation, there are many areas that may be touched upon as appropriate:

- Analysis ground rules and assumptions
- Related studies and their results
- Descriptions of tasks to be modeled
- Alternative methods of calculating effectiveness measures
- Rationale for selecting specific methods of calculation
- Theoretical/mathematical details of calculations
- Discussion of models and the sources and values of inputs
- Overall analysis flow and schedule
- Selection of viable cases to be considered
- Computational run matrices

This list is not exhaustive, but it is representative. It also is not a checklist. What is needed is needed, and what is not is not. The idea is not to fill squares, but to tell the story.

The second significant aspect of the effectiveness analysis is the presentation of results. Presenting results is more an art than a science. It is best learned from experience. The goal is to be concise while clearly conveying the essence of the results. If necessary, an appendix can be used for more detailed results.

As a caution, do not fall in love with slick visual presentations unless they also pass the conciseness and clarity tests. Save the slick technique for when it is the best technique.

Table 14-1 : Breakout by Section of a Sample Final Report

Report* Section	Number of Pages	Percent of Pages
Introduction	7	3
Scenarios	14	6
Alternatives	14	6
Effectiveness Methodology	**47**	**22**
Effectiveness Analysis	**52**	**24**
Cost Analysis	26	12
Cost-Effectiveness	9	4
Summary	4	2
Appendices	45	21
Total	218	100

**Space Propulsion and Power: Operational Effectiveness and Cost Study,* Christopher A. Feuchter and Alexander V. Giczy, Office of Aerospace Studies OAS-TR-96-1, October 1996.

Cost Analysis

Whenever the cost of alternatives is considered, there is cost analysis. Significantly, cost analysts have proclaimed higher standards for documenting their work than any other analytical contingent. This may well stem from the endless stream of brickbats tossed at them over the years for frequently being less than successful in their estimates. Whatever their reason, they deserve to be commended. They have it right!

In practical terms, the principle of reproducibility usually translates to providing a comprehensive summary of the cost analysis in the primary documentation and a separate appendix to catalog the cost details. Following this pattern, the primary documentation should contain a summary of the cost methodology (including cost ground rules and assumptions), a discussion of sources of data, a work breakdown structure, and a summary of major results (including a discussion of cost risk and identification of cost drivers).

Cost Effectiveness Comparisons

Cost-effectiveness is the integration of the cost and effectiveness analysis results into a single story. Cost-effectiveness is extremely important in most analyses and must be adequately documented. In most instances this discussion is short and to the point. It typically consists of 1) this is how we have chosen to portray cost-effectiveness, 2) here are the cost-effectiveness results, and 3) this is our interpretation of those results. The cost-effectiveness discussion is always contained in the main study report.

Organization and Management

Documenting the organization of a study team and its management is quick and easy. It's simply a question of telling it like it is. So-and-so headed the study. These are the subgroups we formed, this is what they did and how they interact, here are the individuals and organizations represented on the panels (a formal acknowledgement page may be used instead). The results will be reported to this one, that one and the other one. So-and-so is supplying the funding.

It shouldn't be any more difficult than that. This documentation serves several purposes. Besides spreading any subsequent blame, it acknowledges the participants' efforts and serves as evidence that the right organizations were involved.

The Principle of Reproducibility

The cost community has a standard for the documentation of cost analyses: in principle a competent cost analyst should be able to reproduce the entire cost analysis based on what is contained in the documentation. The principle may not always be followed, but it's a laudable goal. We should wish that all analysts were equally high-minded.

Appendix: Acronyms

TLAs and More...

ABC	Activity-Based Costing
ACE-IT	Automated Cost Estimating Integrated Tool
ADM	Acquisition Decision Memorandum
AFCAA	Air Force Cost Analysis Agency
AoA	Analysis of Alternatives
BY$	Base Year Dollars
C3I	Command, Control, Communications, and Intelligence
CAIV	Cost as an Independent Variable
CARD	Cost Analysis Requirement Description
CER	Cost Estimating Relationship
CORE	Cost Oriented Resources Estimating
CSEL	Combat Survivor Evader Locator
DIA	Defense Intelligence Agency
DOE	Design of Experiments
DR	Director of Requirements
DTIC	Defense Technical Information Center
DVWG	Data Validation Working Group
EW	Electronic Warfare
FOC	Full Operational Capability
GAO	Government Accounting Office
GRA	Ground Rules and Assumptions
GUI	Graphical User Interface
IG	Inspector General
IOC	Initial Operational Capability
ISR	Intelligence, Surveillance, Reconnaissance
JPALS	Joint Precision Approach and Landing System
LCC	Life Cycle Cost
MNS	Mission Need Statement
MOE	Measure of Effectiveness
MOP	Measure of Performance
MS&A	Modeling, Simulation and Analysis
MSFD	Multi-Spectral Force Deployment

NAIC	National Air Intelligence Center
NASA	National Aeronautics and Space Administration
NTIC	National Technical Information Service
ORD	Operational Requirements Document
O&S	Operations and Support
PMD	Program Management Directive
PME	Prime Mission Equipment
QDR	Quadrennial Defense Review
RDT&E	Research, Development, Test & Evaluation
ROI	Return on Investment
ROM	Rough Order of Magnitude
SEPM	System Engineering/Program Management
SME	Subject Matter Expert
TEMP	Test and Evaluation Master Plan
TLA	Three-Letter Acronym
TOR	Terms of Reference
TRL	Technology Readiness Level
TY$	Then Year Dollars
V&V	Verification and Validation
VV&A	Verification, Validation and Accreditation
WBS	Work Breakdown Structure