9-14-2006

Theory of Effectiveness Measurement

Richard K. Bullock

THEORY OF EFFECTIVENESS MEASUREMENT

DISSERTATION

Richard K. Bullock, Major, USAF

AFIT / DS / ENS / 06-01

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT / DS / ENS / 06-01

THEORY OF EFFECTIVENESS MEASUREMENT

DISSERTATION

Presented to the Faculty
Graduate School of Engineering and Management
Air Force Institute of Technology
Air University
Air Education and Training Command
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

Richard K. Bullock, B.S., M.S.
Major, USAF

September 2006

AFIT / DS / ENS / 06-01

THEORY OF EFFECTIVENESS MEASUREMENT

DISSERTATION

Richard K. Bullock, B.S., M.S.

Major, USAF

Approved:

Dr. Richard F. Deckro Dissertation Advisor	11 Aug 2006 Date
Dr. Robert F. Mills Research Committee Member	11 Aug 2006 Date
Lt Col Jeffery D. Weir, Ph.D. Research Committee Member	11 Aug 2006 Date
Dr. Michael J. Havrilla Dean's Representative	11 Aug 2006 Date

Accepted:

Dr. Marlin U. Thomas Dean, Graduate School of Engineering and Management	17 Aug 2006 Date

AFIT / DS / ENS / 06-01

ABSTRACT

Effectiveness measures provide decision makers feedback on the impact of deliberate actions and affect critical issues such as allocation of scarce resources, as well as whether to maintain or change existing strategy. Currently, however, there is no formal foundation for formulating effectiveness measures. This research presents a new framework for effectiveness measurement from both a theoretical and practical view. First, accepted effects-based principles, as well as fundamental measurement concepts are combined into a general, domain independent, effectiveness measurement methodology. This is accomplished by defining effectiveness measurement as the difference, or conceptual distance from a given system state to some reference system state (e.g. desired end-state). Then, by developing system attribute measures such that they yield a system state-space that can be characterized as a metric space, differences in system states relative to the reference state can be gauged over time, yielding a generalized, axiomatic definition of effectiveness measurement. The effectiveness measurement framework is then extended to mitigate the influence of measurement error and uncertainty by employing Kalman filtering techniques. Finally, the pragmatic nature of the approach is illustrated by measuring the effectiveness of a notional, security force response strategy in a scenario involving a terrorist attack on a United States Air Force base.

ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Dr. Richard F. Deckro, for taking me on as a student, as well as thank him for his guidance and sharing his wisdom during this research effort. I also owe thanks to Dr. Jeffery D. Weir and Dr. Robert F. Mills for their critical feedback.

In addition, I would like to express my appreciation to the Operational Sciences Department and Dr. Gilbert L. Peterson in the Department of Electrical & Computer Engineering. Their courses provided relevant and timely knowledge to tackle this research effort as well as prepare for future challenges.

Further, I would like to thank Dr. Jacqueline R. Henningsen for her encouragement in pursuing an Operations Research Ph.D. Important thanks also go to the USAF for giving me the opportunity to attend AFIT; and AFIT itself for providing unparalleled facilities and staff.

Most of all, I would like to thank my parents and my brother for their continued love and support.

Richard K. Bullock

TABLE OF CONTENTS

FIGURES

Page

TABLES

Page

SYMBOLS

$\boldsymbol{X}$	bold indicates a set
$=$	equality
x_i	i^{th} element of a set
$\cdot$	placeholder for an element or parameter
$\langle \cdot \rangle$, $(\cdot)$	set boundaries; parameter boundaries
$f(\cdot)$	function or mapping, f, indicating required parameter(s) $\cdot$
f:	function or mapping, f
$\in$	'element of'
$\mathbb{R}$	set of real numbers
$\mathbb{R}^+$	set of positive real numbers
$\mathbb{R}^n$	n-dimensional Euclidean space
$\rightarrow$	'maps to'
$\Leftrightarrow$	logical equivalence; 'if and only if'
$>$	'greater than'
$\mathbb{Z}$	set of integers
$\mathbb{Z}^+$	set of positive integers
$+$	addition
k^n	'k raised to the n^{th} power'
Σ	summation
Π	product
$\boldsymbol{A}$	empirical system
$\boldsymbol{a}$	empirical sub-system
$\boldsymbol{x_A}$	model of empirical system $\boldsymbol{A}$
x_i	formal representation of i^{th} empirical sub-system
α_i	i^{th} node or attribute
$\acute{\alpha}$	measure of α
S_i	system state, i
$\boldsymbol{S}$	all possible system states
S_e	end-state
t	time
$\boldsymbol{y}$	resources
$\boldsymbol{C}$	capabilities
$\approx$	approximately
E	effect; system state change
$\frac{d}{dt}$	change with respect to time
$\boldsymbol{I}$	influences
Δ	delta; 'change'
$<_{E,\, Se}$	'is less effective than, with respect to S_e'
$\times$	Cartesian product
$\ni$	'such that'
$\forall$	'for all'

$\Rightarrow$	logical implication; 'it follows'
$\vee$	logical OR
$\square$	'that which was to be demonstrated'
$\mathcal{A}$	collection of subsets
$\varnothing$	null set
$\subset$	'proper subset of'
$\sim$	complement
$\bigcup_{i=1}^{n} A_i$	finite union
$\ldots$	series continues
σ	sigma
$\bigcup_{i=1}^{\infty} A_i$	infinite union
$\cap$	intersection
$\cup$	union
μ	non-negative set function; measure
δ	metric; measure that gauges distances between entities
$\geq$	'greater than or equal'
$\leq$	'less than or equals'
$\cdot_{\boldsymbol{S}}$	a relation on set $\boldsymbol{S}$ (e.g. $\leq_{\boldsymbol{S}}$, $<_{\boldsymbol{S}}$, $=_{\boldsymbol{S}}$,…)
$\wedge$	logical AND
$\neq$	not equal
w_{x_i}	sub-system weighting
w_{α_j}	attribute (node) weighting
$S_{i_{\alpha_i}}$	α_i dimension of state S_i
$\exists$	'there exists'
$\lvert \cdot \rvert$	absolute value
$-$	subtraction
$\lVert \cdot \rVert$	Euclidean norm (i.e. $(\sum_{i=1}^{n}(x_i - y_i)^2)^{1/2}$)
$\sqrt{\cdot}$	square root
$\hat{\mathbf{x}}_k$	estimate at k^{th} period
$\mathbf{x}^*$	all historical system measurements
$\mathbf{x}_k^*$	system measurement at k^{th} period
$\mathbf{P}(a \mid b)$	'probability of a given b'
$\boldsymbol{F}$	system dynamics matrix
$\boldsymbol{G}$	system model
$\boldsymbol{u}$	control vector
$\boldsymbol{w}$	process noise
T_s	measurement periodicity
$\boldsymbol{\Phi}$	fundamental matrix
$\mathcal{L}^{-1}$	inverse Laplace transform
$\boldsymbol{I}$	identity matrix
$\boldsymbol{H}$	measurement matrix

$\boldsymbol{v}$	measurement noise
$\boldsymbol{K}$	Kalman gains
$\boldsymbol{M}_k$	error covariance matrix before measurements
$\boldsymbol{P}_k$	error covariance matrix after the measurements
$\mathbf{E}[\,\cdot\,]$	expected value

THEORY OF EFFECTIVENESS MEASUREMENT

I. INTRODUCTION

THE PROBLEM OF MEASURING EFFECTIVENESS

One accurate measurement is worth a thousand expert opinions.
– ADMIRAL GRACE HOPPER, 1906 – 1992

Measurement is an integral part of modern life. We measure our surroundings, ourselves, and the passage of time. Measurement is needed to characterize the universe and everything in it (Potter, 2000:7). Some have even suggested our advancement as a civilization is a direct consequence of our ability to measure (Sydenham, 2003:3). Despite its seemingly overwhelming importance, measurement is generally regarded with a 'just look and see' attitude; the complexities surrounding measurement often avoid critical analysis (Margenau, 1959:164). This is largely due to the concept of measurement being closely aligned with the physical sciences where measurement is relatively more deterministic. Other disciplines do not enjoy this level of objectivity. Fields in the social and behavioral sciences examine events, processes, and other complex phenomenon that are difficult to understand, let alone measure (Geisler, 2000:35). Another endeavor where measurement is difficult is military operations (Roche, 1991:165). Military operations are characterized by a dynamic and unpredictable environment (Clausewitz, 1976:119). In this complex arena, one would like to measure the outcome of deliberate actions and specifically be able to measure them relative to a desired end-state. One theory on how to achieve such desired end-states in military operations is called Effects-based Operations (EBO).

Effects-based Operations are activities designed to achieve specific outcomes versus activities focused on particular targets or tasks (Deptula, 2001a:53; Lazarus, 2005:23). EBO offers the potential to effectively and efficiently attain objectives across a wide spectrum of complex environments (Henningsen, 2003:3). Based on its potential and supported by results since the 1991 Gulf War, joint doctrine and service doctrine, particularly Air Force doctrine, has undergone change to reflect the EBO concepts. Despite the tremendous promise of EBO, a key challenge is assessment or measuring the outcomes of military activities relative to the desired end-state (Glenn, 2002; Murray, 2001; Bowman, 2002:24). History has shown theory is of limited value if not supported by an empirically feasible measurement method (Scott, 1958:113; Zuse, 1998:84). The challenge in military operations is the system of interest is often ill-defined, exhibiting dynamic, non-deterministic relationships.

Although EBO is typically used in a military context, the problem of measuring the influence, or effectiveness, of actions in a complex, dynamical situation is certainly not unique to the military (Da Rocha, 2005:31). Any situation where there is not a direct and intuitive way to measure progress towards a desired outcome (e.g. economics and law/policy making) relies on actions to shape the situation's environment in order to bring about the desired end-state. However, feedback is required to ensure the actions taken are moving the situation in a favorable direction. This feedback is in the form of effectiveness measures. Effectiveness measures provide the critical link between strategy and execution, essentially translating strategy into reality (Melnyk, 2004:209). Effectiveness measures amount to 'cognitive shortcuts' in the face of an overwhelming complex reality (Gartner, 1997:43). Effectiveness measures influence how decision makers assess the impact of deliberate actions and affect critical issues such as resource allocation as well as whether to maintain or change existing strategy (Gartner, 1997:1). Currently, however, there is no formal foundation or framework for formulating these effectiveness measurements. Lack of a foundation and framework can lead to erroneous measures of effectiveness as exemplified in the following discussion between General George Patton and General Orlando Ward during WWII (Perret, 1991:156):

> "How many officers did you lose today?" asked Patton. "We were fortunate," Ward replied. "We didn't lose any officers." "Goddamit, Ward, that's not fortunate! That's bad for the morale of the enlisted men. I want you to get more officers killed." A brief pause followed before Ward said, "You're not serious, are you?" "Yes, goddamit, I'm serious! I want you to put some officers out as observers," said Patton. "Keep them well up front until a couple get killed. It's good for enlisted morale."

RESEARCH OVERVIEW

This research presents a new framework for effectiveness measurement from both a theoretical and practical view. The research begins by examining the foundational aspects of measurement in a generic sense. The examination includes a brief history of measurement to help establish a context for the many views of measurement as well as establish a basis for the presentation of Measurement Theory. Attention then turns to application of measurement and the concepts surrounding measurement systems to establish a basis for the problems encountered in applied measurement. Moving from measurement in general, to measurement of military effectiveness, an overview of military EBO is provided and compared to a formalized and disciplined framework for decision making, which is followed by a detailed look at 'effects'. A brief survey of concepts and military effectiveness modeling approaches is then provided. Combining these general measurement concepts, as well as effect specific concepts, a general, domain independent, effectiveness measurement methodology is established. This is accomplished by defining effectiveness measurement as the difference, or conceptual distance from a given system state to some reference system state (e.g. desired end-state). Then, by developing system attribute measures such that they yield a system state-space that can be characterized as a metric space, differences in system states relative to the reference state can be gauged over time, yielding a generalized, axiomatic definition of effectiveness measurement.

As noted, military operations, as well as other activities where measurement is critical, are conducted in environments that can be characterized as ill-defined and exhibiting dynamic, non-deterministic relationships. Measurements in these environments can contain error yielding uncertainty concerning the true state of the system resulting from deliberate actions. To address this problem with regard to effectiveness measurement, various probabilistic reasoning approaches are explored. The effectiveness measurement framework is then extended to mitigate the influence of measurement error and uncertainty by employing Kalman filtering techniques.

The effectiveness measurement methodology, along with the probabilistic reasoning technique, forms the basis for the research key result which is a Theory of Effectiveness Measurement establishing the necessary and sufficient conditions for such activities. Measurement itself, however, is an applied task. Thus, to demonstrate the pragmatic nature of the proposed approach, the effectiveness measurement framework is illustrated by measuring the effectiveness of a notional, security force response strategy in a scenario involving a terrorist attack on a United States Air Force base.

THEORY OF EFFECTIVENESS MEASUREMENT

II. BACKGROUND

PREVIOUS WORK

The following sections of this chapter outline key measurement concepts as they relate to effectiveness measurement. The initial three sections are intended to be generic in nature and applicable to any endeavor requiring measurement, covering fundamental notions about measurement, followed by a summary of the representational view of measurement and concepts relating to applied measurement. The initial three sections will identify the general elements required for an effectiveness measurement framework. Then, moving from the general to the specific, an introduction to Effects-based Operations (EBO) and effects, as well as a brief survey of approaches for modeling effects, is provided. These effects related sections will identify the specific elements to make a measurement framework unique for effectiveness measurement.

MEASUREMENT FUNDAMENTALS

> *Not everything that can be counted counts,*
> *and not everything that counts can be counted.*
> – EINSTEIN, 1879 – 1955

Measurement is the objective representation of objects, processes, and phenomenon (Finkelstein, 1984:25). Measurement captures information about these systems through their attributes (also known as characteristics, features, or properties). These attributes can be either directly or indirectly observable (Cropley, 1998:238). Additionally, a system embodies a set of elements where relationships exist between the elements (Feuchter, 2000:12; Artley, 2001b:3). Thus, a system $\boldsymbol{X}$ is defined by the attributes x_i chosen to represent it:

$$\boldsymbol{X} = \langle\, x_1, x_2, \cdots, x_i \,\rangle \qquad (1)$$

Although objective, an important distinction is that measurement is also an abstraction. This challenging aspect of measurement makes it imperative to have formalized frameworks and theories for measurement in order to clarify concepts and ideas about measurement within a particular domain. Measurement is an abstraction because measurement does not directly represent the system but only addresses the attributes selected to represent it (Pfanzagl, 1971:16). In this light, measurement can be thought of as the process of assigning symbols to the attributes of a system such that the assigned symbols reflect the underlying nature of the attributes (Caws, 1959:5). This nature is defined by relations evident when attribute measurements are compared (Pfanzagl, 1971:16).

The assigned symbols can take on any form as long as the set of symbols reflect or can take on the same underlying structure as the attributes being measured (i.e. homomorphic). Typically, the symbols assigned are numerals, where numerals are the material representation of the abstract concept of numbers (Campbell, 1957:295). The assignment of numerals then allows the formal language of mathematics to be applied, enabling further insight into the system's behavior (Torgerson, 1958:1) or more specifically the system's change in behavior, which is central to effectiveness measurement.

All measurement is carried out within a context (Morse, 2003:2). This context is shaped by a purpose, existing knowledge, capabilities, and resources; all of which influence the measurement process (Brakel, 1984:50). Within this context, measurement begins by identifying the system of interest and the attributes to be used in defining the system as depicted in Figure 1. Attribute selection is crucial since the validity of a system measurement is influenced by the number of attributes used in the measurement (Potter, 2000:16). Although fewer attributes will simplify the measurement process, too few can result in poor and/or misleading insights about the system (Sink, 1985:68). Attributes are usually measured independent of one another (Pfanzagl, 1971:15) but hierarchies of attributes can be developed where the attribute or concept under assessment (Mari, 1996:128) can be a complex attribute made up of basic attributes that can not be further sub-divided (Wang, 2003:1321).

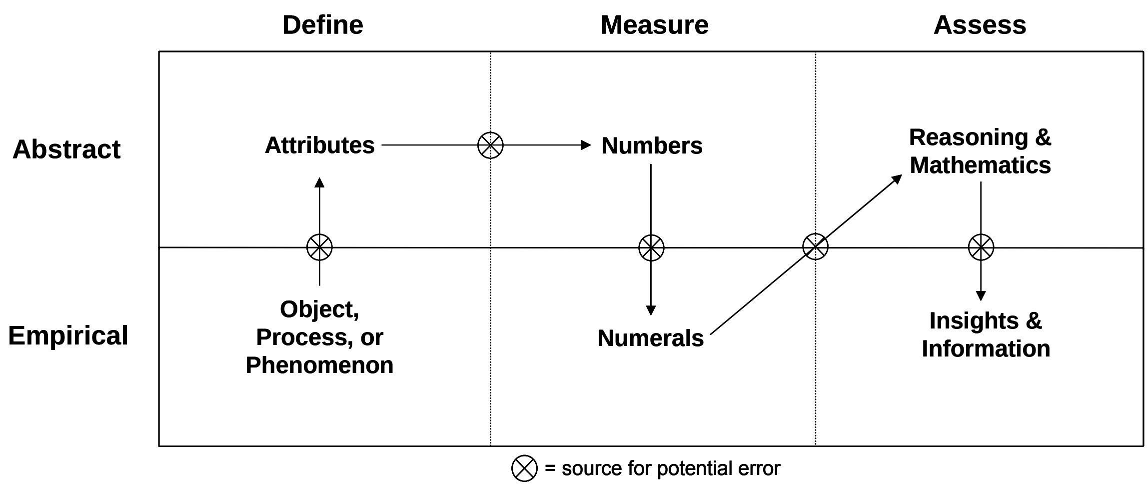


Figure 1. Stages of Measurement

Once the attributes are identified, observations or data collection, on the system attributes can take place. Many terms are often used to describe the result of an observation such as measurement, indicator, or metric. To clarify, a measurement is the raw symbol derived from the observation while an indicator, or index, is a measure for a complex attribute. Further, the term metric has a precise mathematical definition: a distance between two entities where relations between them are non-negative, symmetric, and transitive (Apostol, 1974:60). However, in measurement practice, a metric generally represents a system of measurement composed of the system attributes, the units of measurement, and unit reference standards (Geisler, 2000:75).

Clearly, attributes affect the validity of a measure. Validity characterizes how well a measure reflects the system attributes it was supposed to represent. Another characteristic of a measure is reliability. Reliability, or precision, addresses the consistency or repeatability of the measurement process. A final characteristic of a measure is amplitude, which is how well a measure represents abstract or higher order constructs and complex attributes (Geisler, 2000:40).

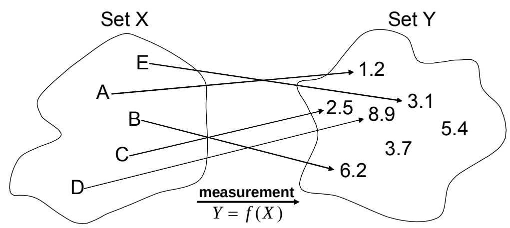


Figure 2. Measurement Scale

Measurements can be made through the human senses or made through use of a measurement instrument, which is an apparatus or construct used for measurement (Geisler, 2000:36). Instruments can be simple like a 'tape measure' or complex, such as a mathematical model. Regardless of form, the instrument must be based on a scale having the same underlying relationships as the system attribute being measured (Mitchell, 2003: 304). A scale (Figure 2) is a predefined mapping from one domain to another, representing empirical system relationships (Sarle, 1995). Because of this, measurement is closely tied to definition (Caws, 1959:3) and the family of mappings for attributes of a system can be considered a mathematical model of the system, since the embedding of the empirical relationships (Scott, 1958:116) requires an understanding of the empirical domain in order to map it into the target, or formal domain. Further, the mappings can encompass uncertainty through use of fuzzy scales to represent the degree to which an attribute is considered present (Benoit, 2003).

Referring back to Figure 1, scales can be a source of error since a measure will always contain any error inherent in the construction of the scale (Potter, 2000:11). In addition to scale construction error, each observation itself is a random variable with an underlying distribution (Potter, 2000:3). A key issue in system measurement, depicted in Figure 1, is the possible sources for error in the process from selection of system attributes to system assessment insights. This error creates divergences between the perceived state of a system and the true state. These divergences can yield misleading insights about the effectiveness of deliberate actions on a system and thus, must be addressed in any framework for effectiveness measurement.

There are three primary sources of measurement error: random, systemic, and observational. Random error is non-deterministic variation from any source impacting the system including the system itself. Systemic error derives from construction of the measure or definition of the measurement process and comes in the form of measurement bias. Finally, observational error is the oversight of key system attributes requiring measurement or using the wrong measures for identified system attributes.

Error in measurement is well established within the physical sciences (Campbell, 1957:437) and will be part of the measurement process even when the system is well-defined (Krantz, 1971:27). Error is an inescapable feature of measurement (Mitchell, 2003:301; Finkelstein, 2003:45) and is a key focus of Metrology, the science of measurement. Measurement error can be partially addressed with Statistical Theory; however, it should be noted, the field of mathematical statistics concerns making inferences from data, while Measurement Theory, discussed below, addresses the link between the data and the real-world. From this point of view, one needs *both* to make inferences about empirical systems (Sarle, 1995:64).

In many contexts, there is a 'Catch-22' with regard to system measurement. In order to properly measure a system, one needs to know something about it; however, the very reason one may want to measure a system is to gain an understanding of it (Geisler, 2000:35). Often for complex objects, processes, and phenomenon with intricate networks of connections, the attributes that best define a system may be unknown, inaccessible, or only visible as an outcome. Measurement of these systems requires use of a proxy or indirect measuring method (Potter, 2000:3) where a proxy measure is essentially a model or approximation of the system attribute of interest. Quantification is the process of developing these indirect measures (Mitchell, 2003:302) or in other words, the process of converting empirical relationships into logical operations. Although there is no universal approach for deriving these proxies, the process typically involves reducing complex aspects of a system into understandable, measurable components.

By one definition, measurement is the assignment of numerals to a system according to a rule (Stevens, 1959:25). However, not all assignment techniques are useful and some techniques have constraints on how the results can be assessed. Although there is not a standard of measurement for complex objects, processes, and phenomenon (Bulmer, 2001), a set of axioms for approaching measurement of these systems can help avoid deriving erroneous insights. Such a set of axioms is embodied in Measurement Theory.

MEASUREMENT THEORY

To measure is to know.
– LORD KELVIN, 1824 – 1907

Formalisms regarding measurement are evident in Ancient Greek culture dating back to the 4th century B.C., but the initial foundations for an axiomatic approach to measurement did not emerge until the late 1800s (Finkelstein, 1984:25). Much of this early work concerned the physical sciences, however. It was not until the mid-1900s, as efforts to measure abstract concepts such as utility and aspects associated with psychology appeared, that a more robust set of principles regarding measurement evolved (Narens, 1986:169). Interestingly, methods of measurement in classical, or Newtonian, physics have evolved without theoretical foundation while the 'softer' sciences required a more robust framework because of the abstract nature of the systems of interest (Finkelstein, 1984:29). This robust framework is contained in Measurement Theory.

Measurement Theory is

> a branch of applied mathematics that attempts to describe, categorize, and evaluate the quality of measurements, improve the usefulness, accuracy, and meaningfulness of measurements, and propose methods for developing new and better measurement instruments. (Allen, 1979:2)

Although there are several viewpoints regarding measurement (Cyranski, 1979:283; Schwager, 1991:618; Niederée, 1992: 237), the most widely accepted form is 'representational' (Finkelstein, 1984:26). The representational view is built upon three theorems: Representation, Uniqueness, and Meaningfulness (Luce, 1984:39). For a system to be measurable, it must be possible to map a formal domain to an empirical domain. The collection of axioms supporting such a representation is called a theory and generally consists of the necessary and sufficient conditions for measurement within a particular domain (Schwager, 1991:619). The purpose of a theory of measurement for a particular domain is to provide structure for a set of empirical observations describing the relations within a system (Finkelstein, 2003:41). This structure can then be used to measure the system of interest for purposes of assessment (Scott, 1958:113). The representational view asserts the symbols assigned to the system represent perceived relations between its attributes (Suppes, 1963:4). Thus, the representational view is based on relational systems.

A relational system is a set of elements where relationships exist among the elements (Pfanzagl, 1971:18). A relational system can be mathematically stated as:

$$\boldsymbol{X} = \langle x_i, \boldsymbol{R} \rangle \tag{2}$$

where x_i represents elements in $\boldsymbol{X}$ and $\boldsymbol{R}$ symbolizes the set of relations between those elements. Real-world relational systems are referred to as empirical relational systems. As an example of another type of relational system, let $\boldsymbol{Y} = \langle y_i, \boldsymbol{A} \rangle$ where $y_i \in \mathbb{R}$, $\mathbb{R}$ is the set of real numbers, and $\boldsymbol{A}$ represents the algebraic operations on $\mathbb{R}$. $\boldsymbol{Y}$ is known as a numerical relational system (Finkelstein, 1984:26). Measurement, m, then can be formally defined as:

$$m: \boldsymbol{X} \rightarrow \boldsymbol{Y} \tag{3}$$

where m is the one-to-one mapping of elements in $\boldsymbol{X}$ to elements in $\boldsymbol{Y}$ in a manner such that $\boldsymbol{R} \Leftrightarrow \boldsymbol{A}$ (Roberts, 1979:52).

This is a fundamental aspect of Measurement Theory and is known as the Representation Theorem and amounts to justifying the assignment of symbols in $\boldsymbol{Y}$. It is accomplished by proving portions of $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the same structure (Suppes, 1963:4) or that relations in the formal domain preserve the relations in the empirical domain (Finkelstein, 2003:43). This theorem implies m is a structure preserving mapping between domains (i.e. homomorphism) (Apostol, 1974:84).

Another component of Measurement Theory is uniqueness. Uniqueness concerns the mathematical characterization of the family of allowable transformations. The Uniqueness Theorem requires any transformations of the mappings $m \in \boldsymbol{M}$ from $\boldsymbol{X} \rightarrow \boldsymbol{Y}$ to maintain the representation conditions (Suppes, 1963:19). In other words, only admissible transformations are allowed (Finkelstein, 2003:43). A great source of difficulty in developing a theory of measurement is not only discovering relations which have an exact and reasonable numerical interpretation, as well as a practical empirical interpretation, but proving under which conditions the relations hold (Scott, 1958:113). However, if these conditions do hold, a scale of measurement $\boldsymbol{S}$ can be defined as:

$$\boldsymbol{S} = \langle\, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{M} \,\rangle \tag{4}$$

Table 1. Scale Types (Narens, 1986:168)

Scale	Admissible Transformations	Examples
Absolute	$x \rightarrow x$	"John is twice as tall as Bill"
Discrete Ratio	$x \rightarrow k^n$, constant $k > 0$, $n \in \mathbb{Z}$	length in lines of code
Ratio	$x \rightarrow rx$, $r \in \mathbb{R}^+$	age, speed, Kelvin temperature
Discrete Interval	$x \rightarrow k^n x + s$, constant $k > 0$, $n \in \mathbb{Z}$, $s \in \mathbb{R}$	murder rate (based on population proportion)
Log Discrete Interval	$x \rightarrow s x^{k^n}$, constant $k > 0$, $n \in \mathbb{Z}$, $s \in \mathbb{R}$	$\dfrac{\text{murder rate per 100,000}}{\text{police force per 100,000}}$
Interval	$x \rightarrow rx + s$, $r \in \mathbb{R}^+$, $s \in \mathbb{R}$	temperature (Fahrenheit or Celsius), calendar dates
Log Interval	$x \rightarrow s x^r$, $r, s \in \mathbb{R}^+$	density (mass/volume), fuel efficiency in mpg
Ordinal	$x \rightarrow f(x)$, f monotonic	beauty, hardness
Nominal	$x \rightarrow f(x)$, $f \in$ 1-to-1 functions	names, numbering on athletic uniforms

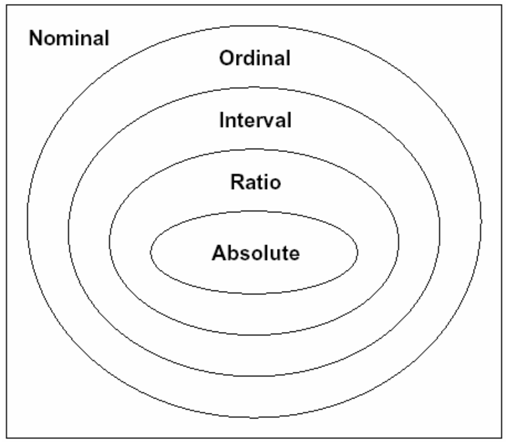


Figure 3. Scale Hierarchy of Commonly Used Measures (Ford, 1993:9)

Despite the generalized notation, only a few scale types exist (Stevens, 1946:677). These are listed in Table 1. These types are sometimes referred to as levels of measurement since each distinguishes the number and types of information contained within the relations of the formal domain. The most common scale types are the Nominal, Ordinal, Interval, Ratio, and Absolute scales (Sarle, 1995:63). A nominal scale only contains equivalence meaning. The ordinal type has both equivalence and rank order meaning. Interval measures have these two meanings as well but also have meaning in the intervals between the values. Ratio measurement further adds meaning in the ratios of values. Finally, absolute scales measure ratios with no units attached, but are also often interpreted as measurement by counting. These scale types are hierarchically related, with the absolute scale type being at the top as shown in Figure 3. Thus, a higher level scale type can always be converted to a lower level scale but not vice versa (Ford, 1993:9). As noted, scale type provides an indication of how much information the assigned symbol contains about the system attribute (Torgerson, 1958:21) but also provides guidance on the transformations allowed to maintain the information (Luce, 1984:39).

A final tenet of Measurement Theory concerns meaningfulness. A measure is meaningful, if and only if, the resultant is invariant for admissible transformations meeting the uniqueness condition (Suppes, 1963:66). Meaningfulness is specific to scale type and can yield misleading or erroneous results when truth or falsity depends on the scale type used (Burke, 2003; Roberts, 1984).

As can be seen, the Representation, Uniqueness, and Meaningfulness Theorems have a hierarchical relationship. The construct starts with proof of the formal representation of the system. Then, uniqueness addresses the class of transformations that maintain the representation. Stated differently, a representation theorem shows how to embed a qualitative structure homomorphically into some family of numerical structures and the corresponding uniqueness theorem describes the different ways that the embedding is possible. Finally, meaningfulness deals with the invariance of a specific symbolic (numerical) statement across admissible transformations.

Many attributes can be measured directly. These are termed extensive attributes, or fundamental measures (Narens, 1985:78). Other measurements may be based on assumed relations or by arbitrary definition (Torgerson, 1958:22). However, as already noted, not all attributes are easily measured. For these intensive attributes (Suppes, 1963:15), indirect measures may not be empirically significant. These proxies are also referred to as weakly defined measures. Systems with such attributes are characterized by ill-defined representation, uncertainty about relational aspects within the system, and have little theory supporting the underlying nature of the system. For attributes of these systems, measurement often precedes definition working in an exploratory, recursive process where measurement leads to definition and definition leads to refined measures (Finkelstein, 2003:45).

One approach to addressing measurement of these ill-defined systems is through conjoint measurement (Luce, 1964:1). Conjoint measurement, or multidimensional scaling (Torgerson, 1958:248), assumes additive or multiplicative decomposability of qualitative structures and combines several indirect or derived measures to increase empirical significance (Narens, 1985:182), where a derived measure is a measure based on other measures (Pfanzagl, 1971:31). Decomposability implies multi-attribute mapping functions, with corresponding scales which preserve empirical ordering, exist (Krantz, 1971:317). Conjoint measurement is common in developing utility functions and development follows a similar procedure (Keeney, 1993:91). Further, the mathematics for working with these constructs is well established (Narens, 1976:197). Although conjoint measurement was initially developed to address weakly ordered attributes, the framework results in a structure for the simultaneous measurement of all attributes (Finkelstein, 1984:28). It should be noted, these structures are sometimes referred to as product structures, where dependent system variables are explained by a number of system stimuli (Roberts, 1979:198).

As already suggested, all measurement is carried out within a context. This implies some purpose for conducting measurement. This purpose can be for system description, monitoring, and/or forecasting. With the theoretical foundations for measurement laid out, the next section examines the application of measurement.

APPLICATION OF MEASUREMENT

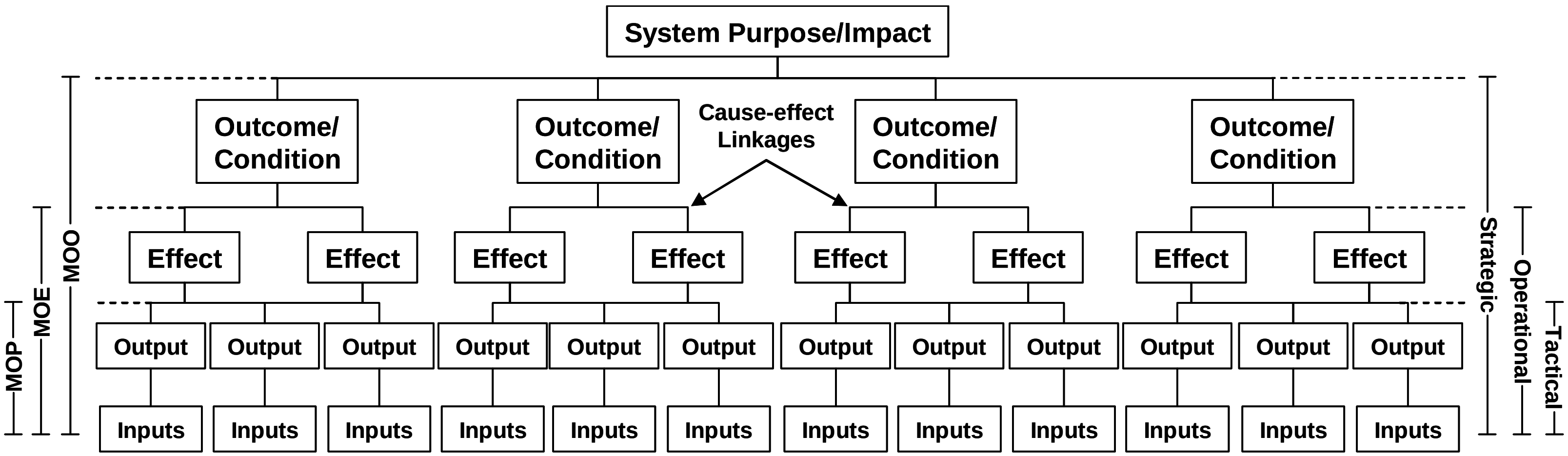
> *Count what is countable, measure what is measurable, and what is not measurable, make measurable...*
>
> – GALILEO, 1564 – 1642

As noted, measurement is a routine, everyday process and a necessity in most fields of endeavor (Rumsey, 1990:19). Measurement is fundamental to understanding, controlling, and forecasting (Wilbur, 1995:1; Antony, 1998:7). Whether conducted explicitly or implicitly, measurement is the mechanism for extracting information from empirical observation. However, obtaining this insight is dependent on having feasible implementation methods, as well as reliable models, for approaching the task of measurement (Sink, 1991:25).

Measurement is applied to a system within a specific context (Morse, 2003:2). The measurement context defines the need for conducting system measurement. This can be for exploratory purposes such as characterizing a new system, but commonly involves resource commitment decisions. Regardless of context, a key aspect for measurement of a system is its environment. Within its environment, a system has some purpose or normative behavior. The behavior of most real world systems is the result of a complex set of interactions and real world systems typically have a complex, abstract purpose. Measurement translates this complex behavior or abstract purpose into a set of 'vital signs' indicating variations in system behavior or gauging fulfillment of system purpose (Kaplan, 1996:75; Melnyk, 2004:209; Ittner, 1998:205) and, most importantly, measures indicate when a system has fulfilled its purpose or is acting in accordance with its normative behavior (Sproles, 1997:16). Further, depending on the measures used, measurement can yield information on when and why a system is deviating from its normal or desired behavior (Kaplan, 1996:84). In order to achieve maximum benefit, however, measurement *must* be an explicit and objective activity. This is accomplished through measurement planning (Antony, 1998:14). If proper planning is not conducted, measurement can become unreliable, untimely, and be more of a burden than a benefit (Antony, 1998:17; USAF, 2003:40).

Measurement activities are often executed as an afterthought and evolve without oversight (Melnyk, 2004:210) leading to ineffectual measures and wasted resources (Hamner, 1993:1-4). One way to prevent this is by developing a measurement plan (Sink, 1985:77). A measurement plan addresses the information to be derived from the measurement activity (Park, 1996:1) and how the system will be measured to include how measures will be determined and how measurements will be collected, as well as the allocation of resources for measurement activities to include training and tools (Eccles, 1991:133). The plan contains all information required to conduct system measurement within a specific context (Neely, 1997:1138) and is sometimes referred to as the measurement protocol (Kitchenham, 1995:937). Additionally, the measurement plan may be integrated with other plans concerning the system such as a strategic plan. Further, the measurement plan should be a 'living document' implying it not only serves to guide the measurement process, but should be used to document, or be an 'audit trail', for how the system measurement process was executed (Sproles, 1996:37).

Before measurement planning can begin, however, a framework for conceptualizing measures is needed. Measure frameworks ensure measurements are traceable back to the original purpose for taking the measurements in the first place.



Inputs – any controllable or uncontrollable factor that enters the system
Outputs – system transformation of the inputs
Effect – changes resulting from the outputs
Outcome – the conditions created by system effects
Purpose/Impact – reason for system existence or expected system behavior

Measure of Outcome (MOO) – gauges conditions created by system effects	Strategic – directly concerns the system purpose or normative impact
Measure of Effectiveness (MOE) – measure changes resulting from outputs	Operational – intermediate events required to achieve the system purpose
Measure of Performance (MOP) – measure of system transformation of inputs	Tactical – short-term activities necessary to attain operational level outcomes

Figure 4. System of Measures

Differentiating between the different frameworks is crucial for effectiveness measurement. These frameworks are commonly classified as either vertical or horizontal. The vertical, or hierarchical, structure is associated with measures that can be directly linked to the system purpose or normative behavior. The horizontal structure, or process framework (De Toni, 2001:50), on the other hand, is normally aligned with system processes, where a process is a set of actions or functions yielding some result (Artley, 2001a:15). Additionally, the vertical structure is often linked with fundamental system objectives, where a fundamental objective is the overall desired or expected system end-state. Alternatively, the horizontal structure is usually linked with means objectives, where a means objective is an enabler for a fundamental objective (Keeney, 1992:66). Typically, measures in the vertical construct are associated with system effectiveness and measures in the horizontal construct concern system efficiency. However, these structures are not exclusive of each other. They can exist at the same time for a system and further, a single measure can exist simultaneously in both constructs (Keeney, 1992:89).

Measures of effectiveness and measures of efficiency provide different insights about a system. A measure of effectiveness (MOE) concerns how well a system tracks against its purpose or normative behavior (Sproles, 1997:17). However, a measure of efficiency, which is also known as a measure of performance (MOP), describes how well a system utilizes resources (Sink, 1985:42). In other words, a MOE determines if the right things are being done and a MOP determines if things are being done right (Sproles, 1997:31). This subtlety is crucial since these measures are developed from differing viewpoints. A MOE can be considered invariant to means of achievement (Lebas, 2002:73; Sproles, 2000:54) while a MOP characterizes system capability or the attributes of a system under a specified set of conditions and is thus, system dependent (Sproles, 1997:16; Sproles, 2000:57). The key distinction, however, is a MOP alone does not provide indication of progress towards a system's purpose or indication of normative behavior. Beyond measures of effectiveness, measures of outcome (MOO) gauge indirect conditions created by system effects (DSMC, 1994), as depicted in Figure 4.

For example, suppose a transshipment warehouse desires a low wait time for items awaiting transit. If we let the desired effect be measured by amount of wait time, one choice for a MOE is average item wait time. As alternatives for transit, trucks, trains, and planes can be used. Regardless of which alternative is used, the MOE will not change. However, each mode of transit will have a different performance measure or MOP (e.g. truck loads, box cars, and plane loads). Additionally, in this hypothetical scenario, because of lower wait times, items are getting to customers faster, resulting in repeat business, as well as new business, having the outcome (MOO) of increased profitability.

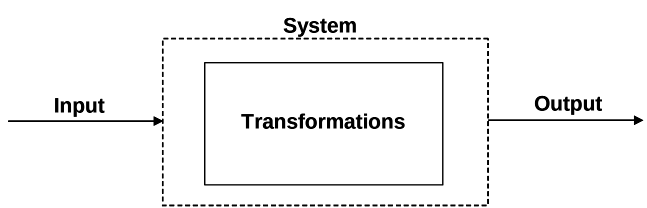


Figure 5. Input-Output Model (Sink, 1985:3)

Another useful construct for conceptualizing a system is an input-output model (Figure 5). Inputs can be any controllable or uncontrollable factor. These inputs enter the system and are 'transformed' into outputs. The outputs result in various effects contributing to conditions in the system's environment which leads to attainment of the system's purpose or normative behavior. The input-output concept is invariant regardless of perspective, with the only change being the type and size of the system and its associated transformations. The key task in development of the model is operationalizing the relationship between the input and output (Sink, 1985:4) where 'operationalize' is the act of quantification or defining an attribute by the way it is measured. The input-output model provides a means for system feedback or quantifying the impact of an input, which is fundamental to understanding and control of any system (Kaydos, 1999:1; Neely, 1997:1132).

A critical element of the input-output construct is defining system boundaries. The boundaries of a system are where elements of the system interact with elements outside the system. Everything outside this boundary is considered the system's environment. The system environment can be described as those factors external to the system that will influence the system over the period of measurement (Artley, 2001a:9). Identifying the boundaries is crucial since they influence the scope of measurement (Sink, 1985:27). Further, making accurate inferences from measurements requires an understanding of the circumstances surrounding the system when the measurements were taken (Wilbur, 1995:17). This contextual information provides insight into why a system behaved the way it did; identifying pressures working with and against the system.

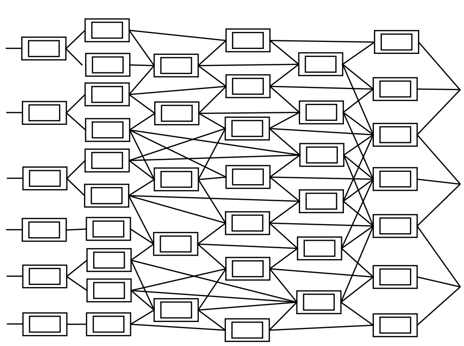


Figure 6. System of Systems

A conceptually helpful extension of this construct is visualizing a network of linked input-output systems, where outputs of one system are the inputs of others (Figure 6). In fact, every system can be seen as part of another larger system (Ackoff, 1971:663). Thus, the combining of systems yields a larger system with its own inputs, outputs, effects, outcomes, purpose/behavior, and boundaries. However, within this larger system, each sub-system still has its own input, output, effect, outcome, purpose/behavior, and boundary (Sproles, 1996:34).

This system-of-systems view allows for conceptualizing the overall system at different levels to include strategic, operational, and tactical (Figure 4). The strategic level directly concerns the system purpose or normative behavior. The operational level focuses on intermediate events required to achieve the system purpose or normative behavior. Finally, the tactical level addresses short-term activities necessary to attain operational level outcomes (Artley, 2001b:12). An interesting analogy for this system-of-systems construct that could also be used to identify significant sub-system inter-linkages is a neural network. Neural networks are made up of perceptrons, which are simple input-output systems. Collections of these perceptrons, as a neural network, can be used to model and explain highly non-linear systems (Mitchell, 1997:81). However, even with simple linear systems, there are numerous challenges confounding the measurement process.

The key to successful measurement is ensuring the right measures are being used to gauge the system purpose or normative behavior (Brown, 1996:3; Leonard, 2004:2). The goal is to understand which inputs or environmental conditions lead to which outcomes (Morse, 2003:38). The key challenge, however, is what we would *like* to measure and what we *can* measure are usually not the same thing (Meyer, 2002:17). Additionally, most endeavors are very situation dependent, ruling out 'one size fits all' sets of measures (Antony, 1998:9; Balkcom, 1997:28; Roche, 1991:191). It is generally accepted, however, the vertical framework should be used for effectiveness measures where all measures are derivative of the system strategic purpose or normative behavior (Brown, 1996:162). Thus, even operational and tactical level measures should flow from the strategic level (Campi, 1993:8-4.3).

The crux of the problem in understanding which inputs lead to which outcomes is identifying and articulating the cause-effect linkages between the strategic, operational, and tactical levels as well as the impact of inputs and environmental factors on each of these levels (Kaplan, 1996:76; Sink, 1985:86). The difficulty in establishing these linkages is usually understated (Hamner, 1993:2-7). The cause-effect relationship can be difficult to discern because the output of one system may be the input of another system and some of the systems may be hidden or inaccessible (Leonard, 2004:35). Additionally, there may be a dynamic delay between a system input and when the impact of that input is seen. Further, for systems in dynamic environments, the cause-effect relationships can change over time (Kaplan, 1996:84) or the system may even adapt to being measured (Neely, 1997:1132; Meyer, 2002:79).

Basic approaches, such as cause-effect mapping, can assist in identifying and explaining some of the linkages. However, the use of historical measurements and statistical techniques are normally required to understand more complex systems (Kaydos, 1999:115; Evans, 2004:219). If these linkages can be identified and understood, a model representing the logical framework of interdependencies between elements within a system and between the system and its environment can be developed. A model based on this representation can then be used for purposes of system forecasting (Feuchter, 2000:12; Kircher, 1959:66).

Despite the challenges of uncovering system relationships, applied measurement concerns the outward behavior of systems versus their internal dynamics. Thus, an effectiveness measurement framework should consists of system measures explaining this behavior. The primary goal in developing system measures is to create a set of measures yielding the most insight while imposing the least amount of burden (Antony, 1998:8).

Approaches to developing measures vary; however, there appears to be wide agreement the starting point is defining the system's strategic purpose or normative behavior as well as associated fulfillment criteria (Sink, 1985:86; Hamner, 1993:2-9; Brown, 1996:11; Antony, 1998:9). These strategic level definitions can be abstract and difficult to quantify for real world systems. Thus, subsequent steps involve reducing the strategic level concepts into conditions or outcomes supporting the system purpose or normative behavior (Hamner, 1993:2-9). An extension of this step sometimes employed is determining the relative importance, or weighting, of multiple, and possibly conflicting, conditions or outcomes (Hamner, 1993:2-12). These can then be further reduced to effects that would bring about the outcomes or conditions (Brown, 1996:11). Next, system outputs that would achieve the effects can be identified. Finally, inputs required to create the outputs are defined (Sink, 1985:86) as shown in Figure 4. The basic concept is to work backward through the cause-effect relationships, iteratively decomposing abstract concepts to a point where they are so narrowly defined a measure suggests itself (Sink, 1985:86). Hopefully, this approach yields a direct, natural measure, or a measure with a universal interpretation that directly measures system purpose or normative behavior. If it does not, a constructed measure must be used.

A constructed measure is defined for a specific context and has two forms. The first is a subjective or categorically defined scale. The second form is an aggregation of several natural measures to form an index. However, if no natural measures are readily apparent and a constructed measure can not be derived, a proxy or indirect measure reflecting attainment of an objective associated with the strategic objective can be used (Keeney, 1992:101).

Table 2. Measure Types (Kirkwood, 1997:24)

	Natural	Constructed
Direct	- Commonly understood measures directly linked to strategic objective - Example: Profit	- Measures directly linked to the strategic objective but developed for a specific purpose - Example: Gymnastics scoring
Proxy	- In general use measures focused on an objective correlated with the strategic objective - Example: GNP (economic well being)	- Measures developed for a specific purpose focused on an objective correlated to the strategic objective - Example: Student grades (intelligence)

The relationships between these measure types are summarized in Table 2. Regardless of the type of measure, the above reductionist process assumes linear decomposition, implying the sum of the constituent parts is representative of the overall system behavior, however, it does not propose how quantification of a complex system can be made tractable (Beckerman, 2000:97). The reductionist philosophy is based on the premise that elements of one kind are combinations of elements of a simpler kind (Sproles, 1996:34) and is central to developing an effectiveness measurement framework.

However, this decomposition process may not be applicable to all systems. Some systems may be better suited to a Systems Thinking, or holistic approach, where the focus is on the interactions between the elements in a system versus the elements themselves, implying the sum of the parts is greater than the whole (Beckerman, 2000:98). Instead of breaking the system into smaller and smaller parts, as in reductionism, the Systems Thinking approach takes an expansionist view by incorporating more and more of the system element interactions. In other words, Systems Thinking moves a system boundary incrementally further out to incorporate more interactions. However, this can result in more complex system models. One methodology to leverage the strengths of both of these views is to start with the reductionist approach and then build back up with the Systems Thinking approach (Beckerman, 2000:99).

Regardless of the modeling approach, large, complex systems can result in numerous measures, each providing only a narrow view of the system. Having numerous narrow views can make it difficult to assess the overall system status. Although the lower level measurements provide the most unambiguous insight about system attributes (Jordan, 2001:17), to get strategic, system level insights, measurements must be combined to summarize this lower level data (Antony, 1998:13; Brown, 1996:4). The problem, however, is the measurements are usually not in the same units. Aggregation involves normalization, standardization, or other means to make these dissimilar measurements commensurable so they can be mathematically combined. Combining dissimilar measurements to get an overall system measurement, however, requires an understanding of the scale types being used in order to ensure the aggregated measurement is meaningful and preserves the original scale level information (Antony, 1998:13).

The normalization process only yields dimensionless measures. Thus, a means of aggregation is needed. There are a number of ways to achieve this. The most obvious is:

$$M = \sum w_i m_i \tag{5}$$

where aggregated measure M is derived by summation of i measures (m) each multiplied by a predetermined weighting (w) or influence on the aggregated measure. If the relationship between the measures is known to be non-linear, a multiplicative aggregated measure can be used:

$$M = \prod w_i m_i \tag{6}$$

Finally, for well understood systems, a high order polynomial may yield an aggregated measure more closely capturing the system's underlying nature (Pinker, 1995:10):

$$M=\sum w_i m_i+\sum w_i m_i^2+\cdots+\sum w_i m_i^n+\sum w_{ij} m_i m_j+\sum w_{ijk} m_i m_j m_k+\cdots+\sum w_{ij\cdots n} m_i m_j\cdots m_n \tag{7}$$

Despite unique measures being required for most systems, and even for the same system in different environments, good measures share some common characteristics. These properties can be categorized as strategically-linked, timely, objective, economical, complete, and measurable.

- Strategically-linked – Effectiveness measures should be traceable to the system strategic purpose or behavior (Kaplan, 1991:73). Additionally, strategically-linked implies the measure is responsive to change and provides an indication of how much change can be attributed to a system input (Neely, 1997:1137). However, other measures, such as process measures, are important for determining *why* a system is behaving the way it is (Brown, 1996:44; Meyer, 1994:97).

- Timely – Measures should be collected and processed in a timeframe that is needed to be relevant within the context (Kaplan, 1991:73; Harbour, 1997:8). This property is at the heart of the trade-off between timeliness and measurement accuracy.
- Objective – This category has two dimensions. 1) Collection: Measures should be easy to understand, be the same regardless of the assessor (accuracy), and be the same under similar circumstances (repeatability) (Finkelstein, 2003:41). Objectivity also implies credibility which concerns measure 'face-value' or whether the measure logically represents what it is supposed to represent. It should be noted, an objective measure can be qualitative but subjective measures should be avoided (Kaydos, 1999:19) since these types of measures are difficult to verify. Subjective measures are commonly associated with questionnaires and interviews (Wilbur, 1995:20). 2) Interpretation: Measures, once obtained, should have an unambiguous interpretation (Antony, 1998:9) and more importantly, distinguish between desired and undesired consequences (Meyer, 2002:79).
- Economical – Collection and processing of measurement data should provide benefits that off-set the burden of measurement activities (Kaplan, 1991:73). Part of an economical measurement system is ensuring the measures are unique and do not contain redundant information (Artley, 2001b:39).

- Complete – Measures should address all areas of concern in enough detail to discern reasons for differences in actual and expected system results (Kaydos, 1999:48). Completeness does not require identifying every relevant system attribute, however; a spanning set of measures associated with the system's purpose or behavior should be attained. Additionally, measures should be limited to those vital for assessing system strategic purpose/behavior and reasons for deviations (Hamner, 1993:2-6; Harbour, 1997:9). Too many measures can result in 'measurement disintegration' (Balkcom, 1997:29) as well as become an economic burden. Completeness can be characterized by breadth and depth where breadth addresses how many of the system attributes are being measured and depth refers to the unit of analysis or 'granularity'. Completeness is also closely related to the concept of balanced measures (Kaplan, 1991). Unfortunately, there is no comprehensive method for developing a complete set of measures. However, achieving completeness typically requires both critical and creative thinking in an iterative process involving negotiation and compromise among those interested in and knowledgeable about the system (Sproles, 2002:258).
- Measurable – Measures should hold for the representation, uniqueness, and meaningfulness conditions. Additionally, measurable implies within a given context if the measure can be feasibly obtained with available resources. This is commonly referred to as being operational (Keeney, 1992:82). Further, measurable implies the collected measures are accurate and can be verified

(Artley, 2001b:39). This is crucial since any system insights gleaned are only as good as the measurements taken (Jordan, 2001:15).

Beyond these specific properties, measures can be categorized based on the type of system they represent. These types include task, process, and object measures. For example, task measures compare a plan versus actual performance. Process measures, on the other hand, are typically used to monitor productivity against a predefined standard, benchmark, or goal. Finally, object measures address specific attributes of a system such as physical properties or functions. Additionally, measures can be grouped by dimension. Single dimensional measures represent fundamental attributes of a system. It follows, multidimensional measures are simply mathematical (linear or multiplicative) combinations of single dimensional measures (Artley, 2001b:3).

The purpose of measurement is to provide meaningful information in support of the context (Antony, 1998:18; Jordan, 2001:3). Measurement alone, however, will not provide this information (Leonard, 2004:14). Measurement, although a crucial element, is only a part of the process of system assessment (Wilbur, 1995:16). Assessment is a systematic process of monitoring a system (Blanchard, 1991:14). Assessment converts raw measurement data into information and knowledge yielding insight (Artley, 2001a:41). Assessment can be categorized either as enumerative or analytical (Evans, 2004:219). Enumerative studies or evaluations are descriptive in nature, describe why a system behaved the way it did, and commonly only provide hindsight (Evans, 2004:222).

Analytical studies, on the other hand, can provide foresight and attempt to understand how a system will behave in the future under certain conditions (Meyer, 2002:49). Analysis is primarily based on historical measurements which can be problematic since past data may not necessarily be a predictor of future system behavior (Meyer, 1997:33). Thus, analysis insights based on historical measurements assume system relations are stable (Lebas, 1995:26). Further, to objectively state an input had a significant system impact requires use of statistical techniques (Evans, 2004:219) yielding a confidence statement.

Assessment, as well as identification of system causal linkages, can be further aided through use of tools from the field of Artificial Intelligence such as Support Vector Machines, Neural Networks, and Decision Tree Learning (Mitchell, 1997; Cristianini, 2000). With any statistical technique, however, there is the possibility of making an incorrect inference. These mistakes are termed Type I and Type II errors. A Type I error, or false-negative, is where a hypothesis is rejected when it is true and a Type II error, or false-positive, occurs when a hypothesis is not rejected when it is false.

Finally, an important, but often underemphasized aspect of system measurement is communication, or the design of information (Tufte, 1997:9). As noted, measurement is carried out within a context. This context could be exploratory or for a resource commitment decision. Modern word-processing, spreadsheet, and database software provide flexible means to generate information displays to support the context. Further, some analytical software packages may come with 'canned' output reports. However, depending on the context, some methods are better for communicating information than others (Tufte, 1997:27). Additionally, the target audience and intended use of the information must be taken into consideration (Jordan, 2001:41). Effective communication of insights via words, numbers, and pictures generally requires creativity. Although there are no universal rules for every situation, the goal for an information display should be to present the maximum amount of information possible while ensuring unambiguous understanding of the insights and their implications for the target audience (Tufte, 1983:105; Jordan, 2001:43). The key point is regardless of how impeccable the measurement plan and implementation, and regardless of how rigorous the assessment, if the insights cannot be effectively communicated, then the measurement context was not effectively supported (Tufte, 1997:9).

The previous three sections provided a general survey of measurement concepts without an application focus. The following sections concentrate on measurement for military campaign assessment and specifically measurement in support of Effects-based Operations (EBO), which will establish specific concepts required for an effectiveness measurement framework.

EFFECTS-BASED OPERATIONS

> *We must make the important measurable, not the measurable important.*
>
> – ROBERT MCNAMARA, 1916 –

Although commonly thought of as an operating concept, Effects-based Operations (EBO) is a theory for the employment of capabilities in dynamic and uncertain environments in a manner to best attain objectives (Williams, 2002:1). EBO provides a conceptual framework for determining the integration and application of capabilities to achieve specific effects, and if correctly applied, influencing an environment of interest yielding desired outcomes (Timmerman, 2003:1). Key tenets of this theory, in the military realm, are a focus on end outcomes, reduced emphasis on weapon systems, and de-emphasis on destruction as a sole means of achieving effects (Henningsen, 2003:2). Another misconception of the theory is EBO requires advanced technology and perfect information (Williams, 2002:11). In fact, concepts concerning EBO are evident in the writings of Sun Tzu and Clausewitz (Ho, 2003:ii). The relatively recent resurgence of effects-based concepts was not so much a re-discovery, but an effort to institutionalize these ideas. This charge was championed by Major General David A. Deptula (Lowe, 2004:2). He suggested air attack with precision weapons as the best means for implementing Effects-based Operations (Deptula, 2001b:25). His emphasis on air power, however, alienated many of those outside the Air Force (Williams, 2002:22). That said, EBO is not solely an Air Force approach. Further, Effects-based Operations is not just a military approach. The concepts of EBO have close parallels to techniques from the discipline of Decision Analysis for deriving better decisions to achieve objectives.

EBO is supported by three pillars: Planning, Employment, and Assessment (USJFC, 2003a:B-3). The major paradigm shift for EBO compared to traditional military approaches is in the planning phase, with the focus on the end-state and the effort to establish the 'objective-to-effects-to-node-to-action' linkages. Like other approaches, EBO is reliant on the efficient employment of capabilities; however, with EBO there is an increased emphasis on non-lethal means. Finally, EBO assessment requires determining if the intended effects were achieved and if they are shaping the desired outcomes.

Military EBO planning starts with the desired outcome being articulated by senior, civilian decision makers supported by input from military leaders. Next, the Joint Force Commander develops supporting in-theater objectives and the outcomes characterizing the end-state, as well as the effects needed to shape those outcomes (Williams, 2002:4). At this point, Measures of Outcome and Measures of Effectiveness, along with their associated success criteria, are established (USAF, 2003:7). After appropriate effects and measures have been selected, courses of action (COAs) can be developed and analyzed, where a COA, or strategy, delineates the who, what, where, why, when, and how (to include with what resources) (McCrabb, 2002:135). The measures are key to the approach since they tie the three pillars together and are used to determine if the intended effects are being achieved and if the strategy and course of action needs adjustment (Smith, 2002:355). Since the planning process starts with the end goal and does not apply weapon system or target solutions during COA development, inherent to EBO is the application of operational art, allowing the strategist to be flexible and innovative.

Key to the EBO approach is understanding the decision context. This is achieved through the 'operational net assessment'. The decision context includes all factors of the strategic, operational, and tactical environment, especially those outside the military realm such as culture, religion, and economics (Meilinger, 1999:55). Another important part of this decision context is understanding who all the participants are, their objectives, and the value each attaches to their objective. The operational net assessment emphasizes the fact that information is a critical enabler for EBO.

The enduring theme of EBO is always keeping the end-state in sight. This type of approach is certainly not unique to military operations. Numerous endeavors require a strategic view (Da Rocha, 2005:31). One methodology from the field of Decision Analysis, Value Focused Thinking, epitomizes this concept:

> You begin with the fundamental objectives that indicate what you really care about in the problem. Then you follow simple logical reasoning processes to identify the mechanisms by which the fundamental objectives can be achieved. Finally, for each mechanism, you create alternatives or classes or alternatives by asking what control you have over that mechanism. (Keeney, 1992:14)

The above quote suggests EBO is based on a robust and formal framework for strategic thinking, yielding strategies for creating effects to influence behavior (Smith, 2002:108) and an optimum way to approach a wide array of situations (Mann, 2002:43). The crux of the challenge in successfully implementing EBO, however, is understanding the nature of effects.

EFFECTS

There is measure in all things.
– HORACE, 65 – 8 B.C.

History has shown warfare is often focused on destruction of the enemy's military forces (McCrabb, 2001:3). History, however, has also shown efficient prosecution of war focuses on strategic ends which are typically not the enemy's military forces (USAF, 2003:7). The end focus is usually some desired end-state where attacks on an enemy's military forces are a means to achieve the end-state, but not necessarily the only means; use of military force is only one instrument of power. Other means of influence include diplomatic, economic, and informational actions for example. Regardless of the means, efficient prosecution of war must be focused on using only actions, and their supporting actions, necessary to shape the end-state, which is accomplished through creation of effects.

A metaphor often used for discussing concepts related to war, which is also useful in discussing effects, is comparing an enemy to a system (Warden, 1995), where a system is a set of related elements that collectively has some purpose or impact (Bouthonnier, 1984:48). The interconnected elements of the enemy typically consists of a directive function for leadership and governing with a strategy, or adaptable plan, for addressing the operational environment; essential resources allowing the enemy to exist, such as money or even a supportive populous; key supporting infrastructure allowing the enemy to translate strategy into action; and some means to carry out strategy, such as armed forces. Effects are generally aimed at affecting one or more of these elements (USAF, 2003:10).

Table 3. Effect Attributes

Attribute	Types
Order	Direct (First-order) Indirect (Higher-order)
Timing	Parallel Sequential
Impact	Cascading Cumulative
Intent	Intended Un-intended (Collateral)
Result	Positive Negative
Persistence	Permanent Non-temporal
Domain	Physical Functional Systemic Psychological
Level	Tactical Operational Strategic

An effect is a state change in a system brought about by an input to the system (Smith, 2002:111; Gallagher, 2004:9; Lowe, 2004:4). An effect can be categorized in a number of different ways (Table 3). The first attribute of an effect is order. A first order effect, or direct effect, is the result of actions with no intervening mechanism between a deliberate action and its corresponding state change. Higher-order effects, or indirect effects, on the other hand, are effects created via intermediate effects, or mechanisms, which can be traced back to the original action that brought them about (Lowe, 2004:5).

Effects can also be classified by timing. Parallel effects are effects planned to occur at or near the same time while sequential effects occur one after another in series.

Another attribute of effects is impact: cascading or cumulative. Cascading effects ripple through a system, degrading or affecting other associated elements of the systems. Cumulative effects, on the other hand, are the aggregation of many smaller direct and indirect effects. Effects can also be described by intent. Intended effects were expected to happen while unintended effects, or collateral effects, were not expected.

Result is another way to discuss effects. Correspondingly, effects can have either a positive or negative influence on friendly operations (USJFC, 2003b:17). Effects can be classified by persistence as well. For instance, an effect may be permanent or its impact may decay over time. Domain is another important aspect of effects. In the physical domain, effects are 'local' and created by direct impact, through physical alteration of an object. In the functional domain, effects represent an impact on the capability, in part of a system, to operate properly (Mann, 2002:37). Systemic effects, however, concern system wide impacts. Finally, psychological effects are aimed at influencing the emotions, motivations, or reasoning of individuals and groups (Mann, 2002:38). Alternatively, an effect's domain can be classified as either physical, informational, or cognitive, where the physical domain is where physical actions take place, the information domain is where actions are detected and reported to higher authority, and the cognitive domain is where decisions as to how to respond at various levels are made (Smith, 2002:161; USAF, 2005:3). Finally, effects can exist at the tactical, operational, and strategic levels of war (USAF, 2003:8).

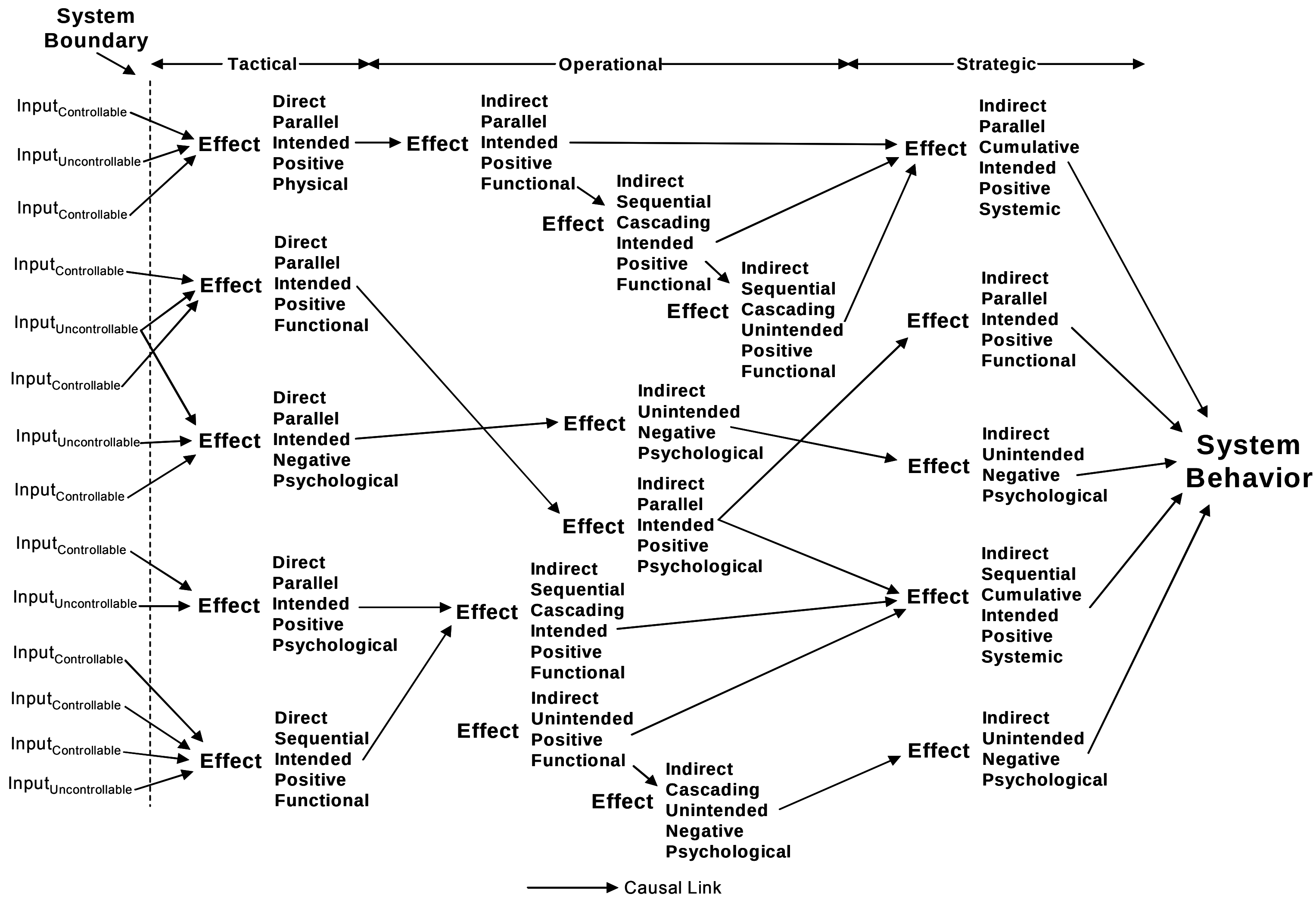


Figure 7. Effects and Causal Links

The system concept can also be used to describe friendly forces. Thus, a conflict can be viewed as a collision of the system describing friendly forces and the enemy's system. Using this as a basis, a construct for thinking about and implementing effects can be developed (Figure 4). As noted earlier, a key point, especially with regard to measuring results of actions, is delineating system boundaries. The boundaries of a system are where elements of the system interact with elements outside the system. Everything outside this boundary is considered a system's environment.

Effects result from inputs to a system. These inputs can be uncontrollable environmental factors or they can be driven from within a system, such as when a country seeks out and obtains monetary aid. Inputs can also be driven externally, as when an adversary attacks. In this sense, the adversary's system is using its own inputs (resources) and transforming them into actions. This output then becomes an input to the system being attacked as depicted in Figure 4. The transformation of inputs to outputs is a measure of efficiency and is generally referred to as a Measure of Performance (MOP). The adversary's outputs, or inputs to the system being attacked, create an effect, or state change. The state change is gauged using a Measure of Effectiveness (MOE). Further, the culmination of effects creates some condition or outcome which can be measured by a Measure of Outcome (MOO). These outcomes shape the system's behavior. Often no distinction is made between MOOs and MOEs with MOOs being assumed as strategic level MOEs.

Figure 4 significantly understates the complexity of the cause-effect chain from a deliberate input to a corresponding system state change. The influence diagram in Figure 7 gives a better sense of the complexities involved. However, in reality, identifying and definitively articulating the cause-effect linkages between the strategic, operational, and tactical levels, as well as the impact of controlled and uncontrolled inputs, is extremely difficult, if not impossible (Kaplan, 1996:76; Sink, 1985:86).

The first aspect of the problem is the abstract nature of the system and the desired change. Essentially, military action is aimed at changing the collective will of a group, where 'will' has no physical form (Meilinger, 1999:50). The cause-effect relationships are difficult to discern because the system being attacked is actually a system-of-systems. The sub-systems are all interconnected, with the output of one sub-system being the input to one or more other sub-systems. Additionally, these sub-systems may be ill-defined, hidden, unknown, and/or inaccessible (Leonard, 2004:35). Further, there may be a dynamic delay between a system input and when the impact of that input is detectable (USAF, 2003:8). Finally, the cause-effect relationships can change as the system adapts to its new, effects shaped environment (Kaplan, 1996:84).

The problem of identifying these cause-effect chains is one of the major objections to EBO. The idea that a group of numerical indicators can determine strategic progress towards victory will always be in question (Murry, 2001:134). However, those championing EBO have recognized this as a problem and in response have put effort into identifying these links as the first step of effects-based planning (USJFC, 2003b:18). The process of uncovering these 'effect-node-action-resource' links is called the operational net assessment (ONA). ONA integrates people, processes, and tools to build shared knowledge of opposing forces, the environment, and friendly forces. The focus of ONA is to understand key relationships, dependencies, and vulnerabilities within and across political, military, economic, social, information and economic systems. The resultant analysis provides insight on ways to influence an adversary which can then be used to develop alternatives for decision makers on how to achieve desired outcomes (USJFC, 2003b:4). Despite the emphasis on uncovering these relationships, it is still a very challenging endeavor. However, a number of approaches have been developed to model effects.

MODELING EFFECTS

I can calculate the motions of heavenly bodies, but not the madness of people.

– ISAAC NEWTON, 1642 – 1727

As already suggested, before a system can be measured, it is first necessary to know something about the system. However, the very reason for measuring the system may be to obtain an understanding of it. This line of reasoning suggests, if one wants to measure a system, one first has to know something about it, and if one has enough knowledge to measure the system then, one can, at least to some degree, model it. The reverse should certainly be true: If one has modeled a system, implied is that the system is understood (at least to the level of modeling), and if the system is understood, insights on how to measure the system should be evident. In fact, one should be able to use the model itself as a measuring instrument for the system of interest. Based on this logic, what follows is a review of some of the current approaches to modeling military effectiveness. Although no single approach captures all the concepts surrounding military effects, collectively they represent the required, known core elements.

In general, existing approaches to modeling military effects can be grouped into three broad categories including Non-linear Sciences, Influence Networks, and Value-based Models.

- *Non-linear Sciences* – The non-linear sciences encompass non-traditional analysis techniques from fields such as Complexity Theory and Chaos Theory. These techniques are especially well suited for exploring the non-linear dynamics of systems which arise from repeated interaction and feedback (Khalil, 2002). Warfare is often characterized in such a manner (Schmitt, 1999:5). Thus, it is no surprise there have been many efforts to use the non-linear sciences to model military actions and their resulting effects. The basic concept behind the non-linear sciences approach is that cause-effect relationships are modeled implicitly and inputs to a system bring about a change or 'emergent behavior' resulting from the collective consequences of the inputs, where 'emergent behavior' is a non-linear science term for strategic effect (Bullock, 2000:63).

  The US Marine Corp began experimenting with Complexity Theory and Chaos Theory, which is typically implemented as a complex adaptive system using agent-based modeling, because existing models did not capture the way Marines fight with respect to maneuver warfare (Ilachinski, 1997). The US Air Force also explored using agent-based models to capture airpower strategic effects (Bullock, 2000). Although the efforts produced promising results (Hill, 2003:17), the non-linear sciences typically use a 'bottom-up' orientation, requiring every element in the system to be modeled and in turn making the non-linear sciences approach very data intensive and time consuming. This reality would make the non-linear science approach difficult to use for deliberate and crisis planning. In addition, the implicit cause-effect mechanisms can be difficult to validate (Champagne, 2003:12). Because of these obstacles, non-linear science approaches are often considered to be in the realm of fundamental science and exploratory analysis (Henningsen, 2003:89).

- *Influence Networks* – While the non-linear sciences model cause-effect relationships implicitly, influence networks model these mechanisms explicitly. Although an influence network is a specific type of tool, here it is also used to describe a family of techniques that include Bayesian Networks, System Dynamics, and Input-Output models. In general, these modeling approaches are composed of a network of nodes and arcs where the arcs characterize the relationships, or flows, between elements in the system represented by the nodes. These types of approaches have the flexibility of not only being able to model physical networks, such as a communications network, but can address abstract processes and situations as well, such as a social network. The exception to this are Input-Output models which typically focus on 'commodity flows' on which the system elements are dependent (Snodgrass, 2000:7; Snodgrass, 2004).

There are numerous benefits to using influence network approaches. Tools implementing these methods, such as the Situational Influence Analysis Model or SIAM (Rosen, 1996), tend to be graphical in nature and thus, have a low leaning curve and are very intuitive for the user. Additionally, the critical thinking required to identify the elements in the system being modeled and the relationships between the elements has a significant benefit beyond insights provided by the model output. Thus, influence network approaches would greatly benefit ONA efforts. Further, influence networks have been successfully integrated with various 'legacy' models (Snodgrass, 2000; DeGregorio, 2004).

Another important benefit of influence network approaches is they can be used as a tool for planning as well as strategy monitoring during plan execution. A model built during planning to evaluate courses of action can be transparently used to monitor plan progress. As probabilistic future events come to fruition and become known, they can be incorporated into the model as 'evidence' providing immediate feedback on changes in the probability of success in achieving a desired end-state (Levis, 2001:17).

Despite these benefits, influence network approaches have some drawbacks. In general, influence networks have unidirectional flow and do not incorporate feedback. This makes it difficult to encompass the dynamic interplay characterizing a clash between adversaries. Further, many influence network implementations do not include time as an input parameter, which is clearly a crucial element in modeling conflict. However, recent efforts have included time to capture the persistence of an effect due to certain actions, providing insight on the impact of timing and the synchronization of actions on outcomes, as well as yielding insight on the probability of success as a function of time (Levis, 2001:12). Additionally, the discipline of System Dynamics, which is focused on developing models of dynamical systems, by design includes feedback as well as time parameters (Forrester, 2003; Byrnes, 2001).

- *Value-based* – Value-based approaches attempt to characterize what is important within a decision context and then describe those elements in a mathematical formula. Specifically, the motivation is to determine what is important to ones own forces and what is important to the adversary; then protect what is important to you and put what is important to the adversary at risk. This is the basic concept proposed by Thomas Schelling in Arms and Influence (1966), despite his emphasis on strategic bombing of the populous as the means of influence.

Value-based models have shown promise in forming the foundation of cognitive models of an enemy (Davis, 2001:76; Whittemore, 1999). Typically, however, value-based approaches focus on what is 'valuable' from a military capability standpoint (Doyle, 1997). Technically, these approaches specify tasks, objectives, and/or values, prioritize them through weighting, and then quantify and normalize them on a scale from zero to one. Success and threshold levels are also identified. With each of the elements weighted so the sum of the weights equals one, the elements can be combined into a single mathematical formula providing a decision maker an indication of overall accomplishment (Larimer, 2004).

Warfare can be described as a clash of highly interconnected system-of-systems where 'soft factors' driven by the 'human element' are pervasive. While most would say this is an accurate description of warfare, the description is certainly not unique to warfare. Other disciplines, such as economics and political science, face the same type of conflict modeling challenges faced in the military realm. Many of the modeling approaches in these fields are applicable to combat.

Arguably, one of the most seminal works in political science attempting to characterize conflict is The War Trap (Bueno de Mesquita, 1981; Bueno de Mesquita, 1985). The War Trap presents a mathematically robust, decision-theoretic based, general theory of war focused on conflict initiation and escalation. Although its 'expected-utility theory of war' is focused on what causes war, the formulation provides insight on how systemically derived statements about conflict and their relationship to empirical evidence can lead to generalizations about complex phenomenon.

The 'expected-utility theory of war' model purports to include rational, war-or-peace decision making with variable orientations towards risk and uncertainty as well as adjustments for national power and capabilities. The goal of the model is to discriminate between those who might expect gain from war and those who would expect to suffer a net loss if they started a war. Fundamentally, the model is based on the following factors: 1) the relative strength of the attacker and the defender, 2) the value the attacker places on changing the defender's policies relative to the possible changes in policies the attacker may be forced to accept if it loses, 3) and the relative strength and interests of all other states that might intervene in the war.

While the aim of the 'expected-utility theory of war' model was to develop a theoretically sound explanation for conflict decision making, it was missing a key element: strategic interaction (Maoz, 1985:88). Expected-utility, and decision theory techniques in general, do not account for the impact a decision will have on other decision makers and do not factor in the decisions of others for the decision at hand. This deficiency was obviously recognized, as a game-theoretic version of the theory appeared in War and Reason (Bueno de Mesquita, 1992).

Game Theory is a framework for thinking about strategic interaction and helps formulate an optimal strategy by forecasting the outcome of strategic situations (Beebe, 1957:1). The idea of a general theory of games was introduced by John von Neumann and Oskar Morgenstern in 1944, in their book Theory of Games and Economic Behavior.

They describe a game as a competitive situation among two or more decision makers, or groups with a common objective, conducted under a prescribed set of rules and known outcomes (von Neumann, 1944:49). The objective of Game Theory is to determine the best strategy for a given decision maker under the assumption the other decision makers are rational, or consistently make decisions in alignment with some well-defined objective, and will make intelligent countermoves, where intelligent implies all decision makers have the same information and are capable of inferring the same insights from that information (von Neumann, 1944:51).

Clearly, strategic interaction is a crucial component when analyzing international conflict or economic situations. Although War and Reason and Theory of Games and Economic Behavior are focused on conflict at a strategic level, Game Theory has proven to be useful for characterizing interaction at the operational and tactical levels as well (Hamilton, 2004:3). Despite a rich history in military modeling, Game Theory is noticeably absent in EBO modeling approaches. Although the focus of this research is on measuring effects versus modeling them, the concepts behind Game Theory are important in understanding the military measurement context (Gartner, 1997:5). A more detailed review of Game Theory can be found in Appendix A.

CURRENT STATUS

...things are to you such as they appear to you
and to me such as they appear to me...

– PROTAGORAS, 485 – 421 B.C.

Currently there is no explicit, theoretical foundation for measuring effectiveness. Additionally, attempts at just defining effects concepts have focused on action verbs which violate the requirement for effects to be invariant to means of achievement

(Gallagher, 2004:9). Given that these measures provide feedback on strategic direction and thus, significantly influence irrevocable decisions concerning allocation of scarce resources, a Theory of Effectiveness Measurement is needed. The purpose of such a theory would not be to replicate reality in a specific domain, but to provide a coherent, organized approach to understanding complex, real events in general. Such a theory would be based on theorems, axioms and assumptions providing a basis for simplifying and organizing reality by delineating the precise conditions and domain where the theory holds, and the ramifications when the conditions are violated. Such axioms and theorems would help the analyst discriminate critical phenomenon from incidental phenomenon, providing a basis for simplifying a complex reality without distorting its essential characteristics (Bueno de Mesquita, 1981:10; Gartner, 1997:9). Clearly, there are no definitive measures which can be prescribed for every objective across every application area (Fenton, 1994:200; Park, 1996:1). Because of this, effectiveness measurement concepts need to be defined in general along with the mathematical properties that characterize these concepts, regardless of the specific attributes to which the concepts are applied.

Key elements supporting a Theory of Effectiveness Measurement include precise definition of concepts, theorems and properties concerning the concepts, and a formalized notation for discussing the concepts in terms of mathematics. Under propositional logic, such an axiomatic-based theory would ensure proof of logically true propositions. However, logical proof does not necessarily guarantee anything of interest will be revealed. A logically true, but empirically trivial or irrelevant theory is of little operational value (Wacker, 2004: 631). With respect to war, too many seemingly valid measures may provide a confusing and competing indication of strategic performance. Additionally, interpretation of measures can be problematic even when the inherent noise accompanying factual information is discounted (Gartner, 1997:8). Therefore, this research includes an empirically feasible framework demonstrating the benefits of the theory, all of which will be discussed in more detail in the sections to follow.

III. RESEARCH METHODOLOGY

OBJECTIVE & TASKS

The objective of this research was to develop a theoretically-based, but empirically feasible approach to measuring effectiveness. Theoretically-based implies mathematically rigorous and a connection to existing, established theories. Empirically feasible, on the other hand, implies robustness, intuitiveness, and practicality. To achieve these somewhat conflicting sub-objectives required the following new contributions.

1) *Scope Problem* – This task involved establishing a foundation for approaching the problem of measuring effectiveness. The task required developing a conceptual construct and bounding the problem in such a way as to ensure precision when mathematical operators are applied. However, the framework needed to be flexible enough to accommodate a wide array of domains and measurement endeavors. This task was accomplished by integrating the concepts of effects and EBO into the representational view of measurement.

2) *Define Concepts* – Effects and EBO have been areas of critical interest in the DoD since the 1991 Gulf War. Because of this, numerous efforts originating within the DoD and external to it, including international efforts, have sought to develop a widely accepted effects lexicon. Unfortunately, the goal has yet to be met and a precise, operational definition of effects for EBO is still being debated (Gallagher, 2004:9). This task involved synthesizing key tenets from the existing, although disjoint, effects literature.

3) *Develop Notation* – The purpose for developing notation was to establish a formal language in order to discuss the qualitative concepts of effects in quantitative, mathematical terms. Since the theory resulting from this research is generic and not tied to any specific domain or measurement effort, this step was critical since mathematics allows the potential for truth to be established independent of reality (Zuse, 1998:7). Additionally, mathematical notation was a critical enabler for accomplishing the next step of establishing the theory (Wacker, 2004: 632).

4) *Establish Theory* – The purpose of effectiveness measurement is to obtain objective information for use in strategic decision-making. However, one cannot be assured of objective information from effectiveness measurements unless they are based on a firm theoretical foundation (Zuse, 1998:9). This final task, building off the previous three, established such a foundation. Because of the desire for the theory to be domain independent, an axiomatic approach was used. The axioms represent basic assumptions about reality. Clearly, such a rule-based framework will not hold under all circumstances. However, the advantage of an axiomatic approach is that the conditions under which the theory holds can be clearly delineated (Zuse, 1998:10).

The above four tasks represent the core contributions of this research. In summary, task one establishes a formal context for thinking about effectiveness measurement. Task two develops unique terminology for effectiveness measurement. Task three, then devises notation, or in-other-words, the syntax of effectiveness measurement. Finally, task four, through a framework of axioms, creates a mechanism for selecting, interpreting, and comparing effectiveness measurements, or essentially, the semantics of effectiveness measurement.

The above four tasks yield a deterministic, effectiveness measurement framework. 'Deterministic' implies perfect information. As noted in a previous section, however, uncertainty and error in measurement is inescapable. Additionally, some would suggest the crux of the problem in measuring effectiveness is uncertainty (Murray, 2001; Glenn, 2002; Bowman, 2002). Thus, a probabilistic framework for reasoning about this error and uncertainty is needed.

The uncertainty exists at many levels. Since for most domains of interest, key attributes will not likely have a direct, natural measure, proxy measures will have to be used. This is tantamount to developing a model of the attribute. Thus, the first aspect of uncertainty concerns whether the model spans the attribute, or in-other-words, if the model is collectively exhaustive. Another, perhaps more fundamental issue of uncertainty, involves whether the right measures are being used to represent a system attribute. A final aspect of uncertainty involves the measurements themselves. Each measurement, or observation, is essentially a draw from some distribution; however, numerous draws from the distribution may be costly, time prohibitive, or just not possible. In fact, many circumstances may only allow for one observation (e.g. satellite image). Thus, the uncertainty is in the form of not knowing where on the distribution the obtained observation lies (i.e. is it at the mean or an outlier).

There are a number of established approaches in other fields for dealing with these types of uncertainty including Kalman filters, Bayesian techniques, and the Theory of Evidence to name a few. This research task, addressing probabilistic reasoning, explored these approaches within the context of measuring effectiveness, establishing the benefits and downside to each, in an effort to determine which technique best supported the deterministic framework.

All the above tasks, resulting in the deterministic and probabilistic frameworks for measuring effectiveness, complete the Theory of Effectiveness Measurement. However, a key goal of the research was to ensure the resulting effectiveness measurement methodology was pragmatic. Thus, to meet this final research objective, the frameworks were demonstrated in a military scenario. This entailed systematizing the theory into a series of steps for application to effectiveness measurement problems. Additionally, this involved demonstrating the consequences of violating the conditions set forth in the axioms of the theory. A key impediment to accomplishing this final task was availability of data. Most available data on historical military battles is attrition-based and not effect-based oriented. Thus, a notional scenario was developed in a combat simulation model called Point of Attack. Output data from the scenario was then used as a basis for demonstrating the effectiveness measurement frameworks.

THEORY OF EFFECTIVENESS MEASUREMENT

IV. RESEARCH FINDINGS

DETERMINISTIC FRAMEWORK

The first step in developing a Theory of Effectiveness Measurement is establishing a philosophical view of effects. While the purpose for creating effects is commonly understood, there is less consensus on the conceptual meaning of an effect as evidenced by the number of effect attribute combinations (Table 3). Current effects literature is dominated by a verb-centric philosophy, implying an effect is a consequence, or result, of a *particular* action. However, significant confusion arises from this approach due to different interpretations and the imprecise meanings of words (Gallagher, 2004:9). A more precise paradigm is to simply view an effect as a change, or more specifically a system state-change (USAF, 2003:8, USJFC, 2003:17).

For example, let an empirical `SYSTEM` of interest, $\boldsymbol{A}$, with `ELEMENT`s, $\boldsymbol{a}$, be represented as $\boldsymbol{A} = \langle a_1,\ldots, a_n \rangle$ where $a_i \in \boldsymbol{a}$, for $i = 1$ to n, are the elements, or `SUBSYSTEM`s, germane to the measurement context. For a world actor, United States Joint Forces Command defines these elements as political, military, economic, social, infrastructure, and information sub-systems, or PMESII (USJFC, 2003b:16). Additionally, 'of interest' implies there is a clearly defined, desired behavior, or `END-STATE`, for $\boldsymbol{A}$, and if the current behavior differs from the desired behavior, some action will be taken.

Further, let $\boldsymbol{x_A} = \langle x_1, \ldots, x_n \rangle$ be the formal representation of the empirical system, or the `MODEL`, where $x_i \in \boldsymbol{x}$ are formal representations of $a_i \in \boldsymbol{a}$. Alternatively, the formal representation could be a function of the elements, $\boldsymbol{x_A} = f(\ x_1, \ldots, x_n\ )$. Additionally, for $i = 1$ to n, let $x_i = \langle \alpha_1, \ldots, \alpha_m \rangle$, where $\alpha_j \in \boldsymbol{\alpha}$, for $j = 1$ to m, are the relevant `ATTRIBUTE`s (or `NODE`s) characterizing element x_i, out of all possible attributes, $\boldsymbol{\alpha}$. These attributes are identified during the Operational Net Assessment, along with `LINK`s, or the relationships between attributes (McCrabb, 2001:28), and `MECHANISM`s which explain the causal *and* temporal aspects of system wide changes (Gill, 1996:175). Finally, both the elements, x_i, and the attributes, α_j, can be reduced to facilitate quantification yielding $x_i = f(\ x_{i1}, \ldots, x_{in}\ )$ and $\alpha_j = f(\ \alpha_{j1}, \ldots, \alpha_{jm}\ )$.

A `MEASUREMENT`, or observation, is a particular manifestation or instantiation of an attribute (McCrabb, 2001:28). System attributes provide a true gauge of the system status. With respect to system measurement, attributes can be broadly categorized by awareness and measurability. Thus, attributes can be known and measurable, unknown and measurable, known and un-measurable, or unknown and un-measurable. If an attribute is known and measurable, the measurement task is relatively straightforward since the attribute will likely have a natural and direct measure (e.g. money, time). Most attributes of interest, however, cannot be directly measured and require an indirect, or proxy `MEASURE`, $\acute{\alpha}$, where $\acute{\alpha} \approx \alpha$. Further, several proxy measures may be required to assess a particular attribute yielding $\alpha_j \approx f(\ \acute{\alpha}_{j1}, \ldots, \acute{\alpha}_{jp}\ )$, where p is the number of measures used to characterize α_j. Additionally, a measure, $\acute{\alpha}_{j1}$, could be composed of lower level measures (i.e. $\acute{\alpha}_{j1} = f(\ \acute{\alpha}_{j11}, \ldots, \acute{\alpha}_{j1q}\ )$), where q is the number of measures used to characterize the higher level measure, $\acute{\alpha}_{j1}$. The lowest level measures can be considered 'atomic' measures, since they cannot be further reduced. Finally, a system STATE, S_t, is a particular instantiation of all atomic measures and thus, an instantiation of all system attributes (or state variables) at a particular point in time, t (Lowe, 2004:4).

Anything not encompassed in $\boldsymbol{x}$ is considered to be the system's ENVIRONMENT where system INPUTS originate. Inputs can be deliberate or can be uncontrollable environmental factors. Deliberate inputs, or control variables, are derivative of RESOURCES, $\boldsymbol{y}$. Like atomic attributes, or attributes that cannot be reduced into more basic attributes, resources are primitives, or basic inputs, and consist of essentials such as information, money, people, and equipment. When choreographed and orchestrated, the resources become a means of influence (Mann, 2002:30), or a CAPABILITY, C. Formally, $\boldsymbol{C} = f(\ \boldsymbol{y}\ )$, assuming the capability to plan and bring together resources is also a resource. It should be noted, capability, as it is used here, implies more than material capabilities, but encompasses the ability to exercise influence, as well as the ability to resist the influence attempts of others (Geller, 1998:57).

It follows, an EFFECT, E, is a system state change, or a change in one or more of the system state variables. Additionally, time, t, is a fundamental parameter in measuring effectiveness since inputs do not yield instantaneous results, but propagate, culminate, and dissipate in a system over time (McCrabb, 2001:10). Further, these system changes are brought about by the inputs (Lowe, 2004:4). As noted, inputs can be controllable and uncontrollable so, system INFLUENCE can be stated as $\boldsymbol{I} = f(\ \boldsymbol{C},\ \text{Inputs}_{\text{Uncontrollable}}\ )$, yielding $E = f(\ \boldsymbol{I}\ ,\ t\ )$.

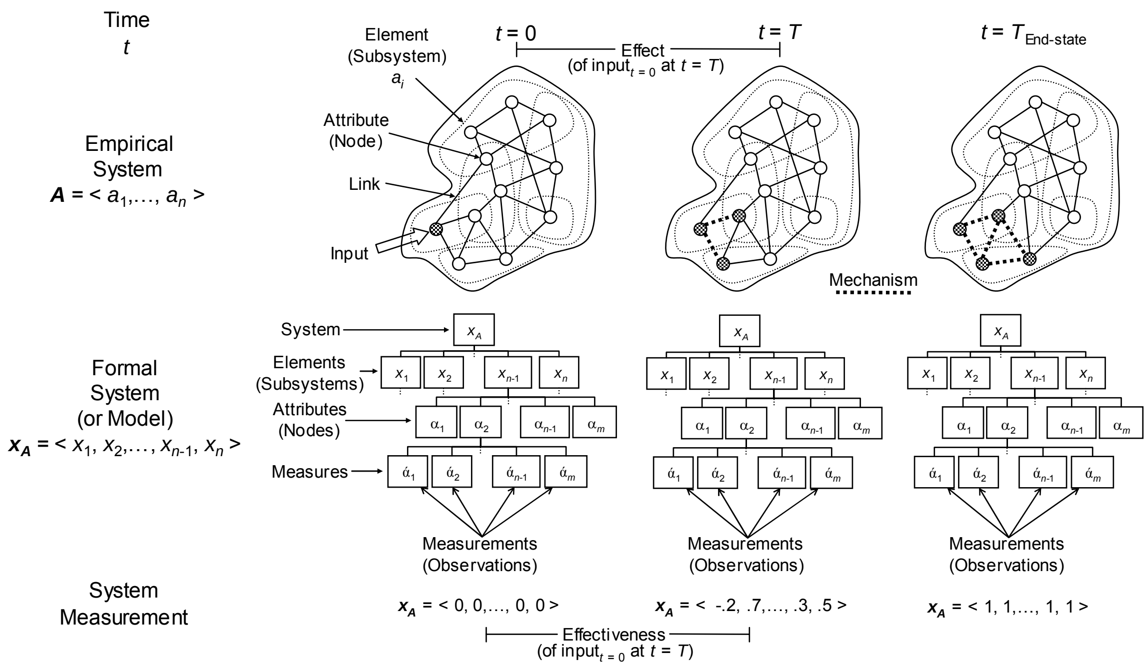


Figure 8. Concept of Effectiveness Measurement

Table 4. Fundamental Definitions

DEFINITION 1: A SYSTEM is a set of elements where relationships exist between the elements and the SYSTEM has a purpose or normative behavior.

DEFINITION 2: A system ELEMENT, or SUBSYSTEM, is a system providing functionality or support to a parent system.

DEFINITION 3: A MODEL is a formal image of an empirical structure.

DEFINITION 4: An ATTRIBUTE, or NODE, is a characteristic, feature, or property of a system that is directly or indirectly observable.

DEFINITION 5: A MEASURE is a model of an attribute.

DEFINITION 6: A MEASUREMENT, or observation, is a particular manifestation, or instantiation, of an attribute.

DEFINITION 7: A system STATE is a particular instantiation of all system attributes, or state variables, at a particular point in time.

DEFINITION 8: An EFFECT is a system state change.

DEFINITION 9: EFFECTIVENESS gauges the magnitude of a system state change.

DEFINITION 10: An END-STATE characterizes the desired measurements for all system attributes, or state variables.

EFFECTIVENESS gauges the magnitude of a system state change due to these influences. Thus, EFFECTIVENESS = $\Delta(\boldsymbol{x}_{A,\, t\, =\, 0}, \boldsymbol{x}_{A,\, t\, =\, T}) \approx \Delta(\boldsymbol{A}_{t\, =\, 0}, \boldsymbol{A}_{t\, =\, T})$ gauges the system impact from controllable and uncontrollable inputs between time $t = 0$ and time $t = T$. Finally, a key point is while an effect occurs on the empirical system, effectiveness is measured on the formal system. The precise definition of these concepts, along with their formalized language in terms of mathematical notation, is a cornerstone required for formal theory building (Wacker, 2004: 632). Figure 8 summarizes these concepts pictorially. In addition, Table 4 highlights definitions of key concepts to be extended in what follows.

Although Figure 8 addresses the concept of effectiveness in a generic sense, Figure 8 does not imply a general notion of effectiveness. This is a problem in the current literature which typically addresses effectiveness in generalities. For example, in response to an action, the question, "How effective was it?," has no meaning. In fact, it can be shown analytically a general-purpose, real-valued effectiveness measure, with the minimum assumption of an ordinal scale, does not exist.

THEOREM 1: A general notion of effectiveness does not exist.

PROOF: Let $\boldsymbol{S}$ be the set of all possible system states and $S_{i,\, t=T} \in \boldsymbol{S}$ be the system state at time $t = T$ resulting from input i. Additionally, let $S_{t=0}$ be the starting system state and S_{e} be the desired end-state. For independent system inputs, x and y at $t = 0$, system effectiveness is characterized by an empirical relation system which includes the relation $<_{E,\, S_{\mathrm{e}}}$, where $<_{E,\, S_{\mathrm{e}}}$ can be interpreted as "is less effective than, with respect to S_{e}" and E is the measure of effectiveness of the input with respect to S_{e} at $t = T$. However, for such a formalism to exist requires E: $\boldsymbol{S} \times S_{\mathrm{e}} \rightarrow \mathbb{R} \ni <_{E,\, S_{\mathrm{e}}}$ holds $\forall S \in \boldsymbol{S}$. This suggests $S_{x,\, t=T} <_{E,\, S_{\mathrm{e}}} S_{y,\, t=T} \Rightarrow E(x) < E(y)$. While $<_{E,\, S_{\mathrm{e}}}$ may clearly hold for some states, others states at time $t = T$ resulting from inputs x and y will not be comparable due to imprecision in the meaning of 'effectiveness'. This suggests $<_{E,\, S_{\mathrm{e}}}$ is not a total order on $\boldsymbol{S} \times S_{\mathrm{e}}$ while $<$ *is* a total order on $\mathbb{R}$. This violates Cantor's Theorem (Fenton, 1994:201);

specifically, the negative transitivity aspect of the strict weak order property: $\forall S_{x,y,z,\ t=T} \in \boldsymbol{S}$, $(S_{x,\ t=T} <_{E,\ S_e} S_{y,\ t=T} \Rightarrow (S_{x,\ t=T} <_{E,\ S_e} S_{z,\ t=T} \vee S_{z,\ t=T} <_{E,\ S_e} S_{y,\ t=T}))$. □

To illustrate the consequences of this theorem, using Figure 8 as a reference, let there be two actions, *y* and *z*. At $t = T$, *y* results in $\boldsymbol{x_A} = \langle -.2, .7, \ldots, .3, .5 \rangle$ while *z* results in $\boldsymbol{x_A} = \langle .8, .5, \ldots, -.4, .4 \rangle$. Which action, *y* or *z*, was more effective? **THEOREM 1** asserts this question cannot be answered. Thus, effectiveness measurements must *always* be with respect to specific system attributes from which it follows, E: $\boldsymbol{S} \times S_e$ must be mathematically complete (i.e. $\forall S_i, S_j \in \boldsymbol{S}, ((S_i \leq_{E,\ Se} S_j) \vee (S_j \leq_{E,\ Se} S_i))$).

Clearly, however, developing a universal set of system attribute effectiveness measures is futile. But, an axiomatic framework *can* provide a sound foundation and guidance for developing *all* specific system effectiveness measures. Thus, although there is no general notion of effectiveness, for specific effectiveness measures, there is a need to define effectiveness measurement concepts and define precisely the mathematical properties that characterize these concepts, regardless of the specific system attributes to which these concepts are applied.

In Measurement Theory, the empirical understanding of a system attribute is formalized through definition of an empirical relational system. A measure is valid if it is a homomorphism from the empirical relational system into a formal relational system, or in other words, if the measure maps system attributes into values such that all empirical relations among the attributes are preserved as formal relations among the measurement values (Poels, 2000:35). Clearly, the crux of the problem in effectiveness measurement is most aspects of the empirical relational system, such as links and mechanisms, are ill-defined or unknown. However, the empirical aspects of a system that are known can be formalized as a set of desirable properties for the system measures. Thus, instead of explicitly defining the formal relational system, an axiomatic approach defines properties for the formal system based on properties of the empirical relational system (Poels, 2000:35).

The entire field of mathematics is axiomatic-based where concepts are defined using necessary and sufficient sets of rules. One such concept, from Measure Theory, is called a metric. As noted earlier, in measurement practice, a metric generally represents a system of measurement composed of the system attributes, the units of measurement, and unit reference standards (Geisler, 2000:75). In mathematics, however, a metric has a precise definition which is developed in this section. First, however, to define a metric, or a 'measure of distance', a measurable space needs to be defined.

An algebra, on a set $\boldsymbol{S}$, is a collection, $\mathcal{A}$, of subsets of $\boldsymbol{S}$ where $\boldsymbol{S}$, $\emptyset \subset \mathcal{A}$, $A \in \mathcal{A} \Rightarrow \sim A \in \mathcal{A}$, where $\sim A$ is the complement of A, and $A_1, A_2, \ldots, A_n \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{A}$. In other words, an algebra is a collection of subsets of $\boldsymbol{S}$, which contains $\boldsymbol{S}$ and is closed under the complement and finite union. In this context, $\mathcal{A}$ is a measurable set. Further, $\mathcal{A}$ is a σ-algebra when $\forall i$, $i \in \mathbb{Z}^+$, $A_i \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$. Additionally, a measure, μ, is a non-negative set function on the σ-algebra, $\mathcal{A}$, where $\mu(\emptyset) = 0$, $\forall A, B \in \mathcal{A}$, $(A \cap B) = \emptyset$, $\mu(A \cup B) = \mu(A) + \mu(B)$, and $\mathcal{A} = \bigcup_{i=1}^{\infty} A_i \Rightarrow \mu(\mathcal{A}) = \sum_{i=1}^{\infty} \mu(A_i)$ (countably additive). It follows, $(\boldsymbol{S}, \mathcal{A}, \mu)$ is a measure space and $(\boldsymbol{S}, \mathcal{A})$ is a measurable space (Ruckle, 1991:80-81). A familiar example of these spaces is Cartesian space. With these fundamental constructs established, a metric can now be defined.

A metric, δ, is a type of measure that gauges distances between entities. Specifically, a metric on a set $\boldsymbol{S}$ is a function $\delta: \boldsymbol{S} \times \boldsymbol{S} \rightarrow \mathbb{R}^+ \ni \forall S_i, S_j, S_k \in \boldsymbol{S}$, $\delta(S_i, S_j) \geq 0$ (non-negativity), $\delta(S_i, S_j) = 0 \Leftrightarrow S_i = S_j$ (identity), $\delta(S_i, S_j) = \delta(S_j, S_i)$ (symmetry), and $\delta(S_i, S_j) \leq \delta(S_i, S_k) + \delta(S_k, S_j)$ (triangle inequality), where $\times$ denotes the Cartesian product, or all ordered pairs of vectors in $\boldsymbol{S}$ (Marlow, 1978:2). Additionally, if the second condition, $\delta(S_i, S_j) = 0 \Leftrightarrow S_i = S_j$, is replaced with $\delta(S_i, S_j) = 0 \Rightarrow S_i = S_j$, then δ is a semimetric or psudeo-metric (Cohn, 1980:8). It follows, $(\boldsymbol{S}, \delta)$ is a metric space. The above demonstrates how mathematics, and specifically Measure Theory, defines a measure via rules or axioms. Through the use of an axiomatic approach, measures can be 'validated', where sufficiency is guaranteed by proving invariance with respect to the rule set.

Measure Theory only addresses formal systems. Measurement Theory, on the other hand, is focused on mapping empirical systems to these formal structures. In other words, the formal representations are numerical structures used to represent the empirical systems. Dimensional metric models, which are numerical representation of qualitative structures with coordinate-vector representations, using primitives such as points and comparative distances, are often used as the formal structures. Dimensional metric models are based on two general concepts: 1) the representation of objects as points in a coordinate space and, 2) the use of metric distance to represent proximity between the points (Suppes, 1989:207). The most basic Dimensional metric model is a Geometrical model (spatial), which depicts objects as points in a space such that the proximity ordering of the objects is represented by the ordering of the metric distances among the respective points (Suppes, 1989:159). A familiar example of such a representation is points in n-dimensional Euclidean space, which is a particular type of metric space. The Measure Theory axioms required for a metric to be a measure on a formal structure were identified earlier. It will be shown in what follows, a metric is also a measure according to Measurement Theory when axioms defining a proximity structure are satisfied.

A proximity structure represents empirical relations, but is also a metric space in which any two points are joined by a straight line segment, along which distance is additive, yielding an ordering among the entities (Suppes, 1989:7). To further elaborate on proximity structures, let $\leq_{\boldsymbol{S}}$, $<_{\boldsymbol{S}}$, $=_{\boldsymbol{S}}$ be quaternary relations on $\boldsymbol{S}$ where $\forall S_i, S_j, S_k, S_l \in \boldsymbol{S}$, $(S_i, S_j) \leq_{\boldsymbol{S}} (S_k, S_l)$ means the difference, or conceptual distance, between S_i and S_j is at most as great as the distance between S_k and S_l, $(S_i, S_j) <_{\boldsymbol{S}} (S_k, S_l)$ implies the distance between S_i and S_j is not as great as the distance between S_k and S_l, and $(S_i, S_j) =_{\boldsymbol{S}} (S_k, S_l)$ suggests the distance between S_i and S_j is the same as the distance between S_k and S_l. It follows, $(\boldsymbol{S}, \leq_{\boldsymbol{S}})$ is a proximity structure if and only if $\forall S_i, S_j \in \boldsymbol{S}$, $((S_i \leq_{\boldsymbol{S}} S_j) \vee (S_j \leq_{\boldsymbol{S}} S_i))$ (strongly complete), $\forall S_i, S_j, S_k \in \boldsymbol{S}$, $(((S_i \leq_{\boldsymbol{S}} S_j) \wedge (S_j \leq_{\boldsymbol{S}} S_k)) \Rightarrow (S_i \leq_{\boldsymbol{S}} S_k))$ (transitive or consistent), $\forall S_i, S_j \in \boldsymbol{S}$, $((S_i \neq S_j) \Rightarrow ((S_i, S_i) <_{\boldsymbol{S}} (S_i, S_j)))$ (positivity), $\forall S_i, S_j \in \boldsymbol{S}$, $((S_i, S_i) =_{\boldsymbol{S}} (S_j, S_j))$ (minimality), and $\forall S_i, S_j \in \boldsymbol{S}$, $((S_i, S_j) =_{\boldsymbol{S}} (S_j, S_i))$ (symmetry). Thus, δ is both a formal and empirical metric, or measure of distance, if and only if, $\forall S_i, S_j, S_k, S_l \in \boldsymbol{S}$, $(S_i, S_j) \leq_{\boldsymbol{S}} (S_k, S_l) \Leftrightarrow \delta(S_i, S_j) \leq \delta(S_k, S_l)$ (Suppes, 1989:160).

This suggests every function satisfying the metric axioms is by definition a valid measure of distance when the system is a proximity structure. In a similar manner, weak ordering on a metric space gives rise to a proximity structure (Suppes, 1989:162). This implies effectiveness measures can be defined to measure the differences, or conceptual distances, between system states. Thus, what follows is a framework for system effectiveness measurement where measures, $\acute{\alpha}_j$, for empirical system attributes, α_j, are defined to hold for the properties of a metric giving rise to system state-spaces satisfying the properties of a proximity structure. System effectiveness measurement then, is the difference, or conceptual distance, from a given system state to some reference system state (e.g. end-state). By defining system attribute measures such that they yield system state-spaces characterized as proximity structures, differences in system states relative to a reference state over time can be gauged, resulting in an axiomatic definition of effectiveness measurement.

The proximity structure is not the only way to formally represent a system. There are numerous other types of structures. These include Grassmann structures (Krantz, 1971:229) and difference structures (Zuse, 1998:250) to name a few. These structures are essentially axiomatic system models. No particular structure is more correct than another. Choice of a structure, or formal model, depends on empirical system assumptions, empirical system hypotheses, and the measurement context. For example, to prove the properties of an extensive structure (Krantz, 1971:72) requires various combination rules such as concatenation (i.e. addition) hold. However, this implies the elements in the measure space represented by the extensive structure have meaning if combined. This may be true for many empirical systems, but for the effectiveness measurement framework presented here, there is no empirical meaning behind arbitrary combinations of systems states (i.e. points). That being said, the proximity structure *is* well-suited to providing insight on system states relative to a reference system state (e.g. end-state).

It should be noted, 'valid' as it is used here, implies theoretical validity suggesting the measure, $\acute{\alpha}_j$, satisfies all of the axioms established to define the formal system, or model. Although definition of system attributes as distances, during the ONA, should reflect empirical understanding of the system attributes, theoretical validity does not imply empirical validity. To define an empirically valid measure, however, requires certainty about the underlying structure of the empirical system to include attributes, links, and mechanisms. Clearly, for real-world systems, especially for those as complex as in the military realm, this information will be less than certain. Despite this uncertainty, to develop a framework to make quantitative statements about a qualitative, or empirical, system requires a specification, or product structure, for the system; in other words, a robust process for developing the system model, $\boldsymbol{x_A}$. Such processes can be found in Decision Theory, and specifically Value Focused Thinking (Keeney, 1992), where structured processes are used to reduce an abstract objective of a complex decision problem into values indicating why the problem is important and further, into quantifiable attributes that can be used to rank order alternatives to achieve the objective. One such process, modified to serve as a generic system state specification, or product structure, for the purpose of measuring effectiveness, is described in the steps below.

1. *System Identification.* This first step is crucial since it determines the system boundary. An empirical system, $\boldsymbol{A}$, and its formal representation, $\boldsymbol{x_A}$, should encompass all pertinent aspects of a desired end-state.

2. *Sub-system Identification.* For the identified empirical system, $\boldsymbol{A}$, and its formal representation, $\boldsymbol{x_A}$, identify empirical sub-systems,

$a_i \in \boldsymbol{a}$, and their formal representations, $x_i \in \boldsymbol{x}$, where $\boldsymbol{A} = \langle a_1,\ldots, a_n \rangle \approx \boldsymbol{x_A} = \langle x_1,\ldots, x_n \rangle$. An empirical system, $\boldsymbol{A}$, will likely have many possibilities for decomposition into smaller sub-systems, $a_i \in \boldsymbol{a}$. Choice of sub-systems should be limited to those that support the measurement context. Additionally, like the parent system, each sub-system will have its own boundary within the parent system. Ideally, the sub-systems should be defined in such a way that sub-systems are disjoint, or mutually exclusive, from other sub-systems. It should be noted, however, empirical systems of interest are often highly interconnected and mutually exclusivity may not be achievable (i.e. $\boldsymbol{A} = \langle a_1 \rangle$). Further, subject matter expert, mental models may have to be used when there is little understanding about system interconnectivity.

3. *Define Sub-system Relative Importance.* All identified sub-systems should be relevant to the measurement context; however, they may not all have the same level of relevancy. For all sub-systems, the relative importance among sub-systems must be defined. This amounts to weighting each of the sub-systems with respect to the other sub-systems. This can be done by developing a number (Keeney, 1992:148), w_{x_i}, for each sub-system, $x_i \approx a_i$, where $0 \leq w_{x_i} \leq 1$ and $\sum_{i=1}^{n} w_{x_i} = 1$.

4. *Attribute* (*Node*) *Identification.* Each sub-system, $x_i \approx a_i$, can be characterized by certain salient features, or attributes, α_j. Like the sub-systems, there will likely be a number of attributes from which to choose.

However, only attributes relevant to the measurement context should be used. Thus, for each sub-system, x_i, with attributes α_j, $x_i = \langle \alpha_1, \ldots, \alpha_m \rangle$ for $j = 1$ to m, where m is the number of relevant sub-system attributes.

5. *Define Attribute* (*Node*) *Relative Importance.* Like the sub-systems, all identified attributes, α_j, should be relevant to the measurement context but, they may not all have the same level of relevancy. For all attributes within a sub-system, the relative importance among the attributes must be defined. Again, this amounts to weighting each of the attributes with respect to the other attributes within a sub-system. This can be done by developing a number (Keeney, 1992:148), w_{α_j}, for each attribute defining a sub-system where $0 \leq w_{\alpha_j} \leq 1$ and $\sum_{i=1}^{n} w_{\alpha_{ji}} = 1$.

6. *Measure Development.* Each attribute, α_j, needs to be quantified. Attributes may need to be further reduced for quantification purposes. The basic measure development approach is to iteratively decompose the attribute into more basic attributes until they are so narrowly defined, a measure for attribute α_j, $\acute{\alpha}_j$, suggests itself (Sink, 1985:86). These will typically be in terms of counts (e.g. number of sightings in a day).

If the above reductionist approach does not yield atomic attributes with natural measures, constructed measures have to be used (Keeney, 1992:103). The first step in building a constructed measure for an attribute is to characterize the desired end-state, as well as the starting state, in terms of the attribute. Then, define possible intermediate states between the starting state and end-state, or in-other-words, construct a model of the distance between states of $\boldsymbol{x_A} \approx \boldsymbol{A}$ with respect to $\acute{\alpha}_j \approx \alpha_j$. Additional system states in the neighborhood of the starting state should also be defined to encompass possible negative consequences, or deliberate system inputs that lead away from the desired end-state. Definition of the intermediate states, essentially defines the units for the constructed measure $\acute{\alpha}_j$. Regardless of type of measure, however, natural or constructed, $\acute{\alpha}_j$ needs to hold for the properties of a metric. That is, each $\acute{\alpha}_j$ must hold for non-negativity, identity, symmetry, and the triangle inequality properties. A measure, $\acute{\alpha}_j$, meeting these properties will be identified by δ_{α_j} to signify it is both a measure of α_j and a metric. Thus, $\delta_{\alpha_j} \approx \alpha_j$.

Using this procedure as a system state specification, the following framework proposes $\forall\, S_k, S_e \in \boldsymbol{S}$, $\delta(S_k, S_e)$: $\boldsymbol{S} \times S_e \to \mathbb{R}^+$, or in other words, the proximity ordering by metric distance, be used as a measure of the difference, or conceptual distance, between state S_k, and the desired end-state, S_e, where $\boldsymbol{S}$ is the set of all possible system states. By defining attributes, $\delta_{\alpha_j} \approx \alpha_j$, based on the system state specification, $\delta(S_k, S_e)$ is a valid metric from both a Measure Theory and Measurement Theory perspective. However, as will be shown, assuming system state S_k characterizes the empirical system through two or more attributes, $\delta(S_k, S_e)$ is actually a semimetric, or pseudo-metric, since any two states in $\boldsymbol{S}$ can be different across attributes, but have the same conceptual distance to the reference state. Thus, only atomic measures, δ_{α_j}, are pure metrics in the mathematical sense.

Clearly, under this framework, a system state (i.e. point) by itself has no measurement. The concept of a difference, or conceptual distance, requires two states (i.e. the system state of interest, S_k, and a reference system state, such as the end-state, S_e). Thus, for a set of possible system states, $\boldsymbol{S}$, the empirical relation system consists of a set of entities and their relations. Comparison of all pairwise combinations of system states is denoted by $\boldsymbol{S} \times \boldsymbol{S}$. However, if the reference system state is S_e, this reduces to $\boldsymbol{S} \times S_e$, where each atomic attribute addresses a unique aspect (i.e. dimension) of system state difference. Thus, the focus here is on the differences, or conceptual distances, between system states, and more importantly for effectiveness measurement, a relation expressing a total order on $\boldsymbol{S} \times S_e$. The empirical ordering relation for system states can be expressed as $\leq_{\boldsymbol{S}}$, where $\leq_{\boldsymbol{S}}$ is a ternary relation mapping to the positive real numbers, $\mathbb{R}^+$. Thus, $\delta_{\boldsymbol{S}}:(\boldsymbol{S} \times S_e, \leq_{\boldsymbol{S}}) \rightarrow (\mathbb{R}^+, \leq)$ is a homomorphic mapping suggesting $\forall S_i, S_j, S_e \in \boldsymbol{S} \ni (S_i, S_e) \leq_{\boldsymbol{S}} (S_j, S_e) \Leftrightarrow \delta(S_i, S_e) \leq \delta(S_j, S_e)$, from which it follows $\delta_{\boldsymbol{S}}:(\boldsymbol{S} \times S_e, \leq_{\boldsymbol{S}})$ is of at least ordinal scale type. This useful result suggests,

> **THEOREM 2:** Effectiveness measures require at least an ordinal scale type.
>
> PROOF: For an ordinal scale effectiveness measure, $\delta_{\boldsymbol{S}}:(\boldsymbol{S} \times S_e, \leq_{\boldsymbol{S}}) \rightarrow (\mathbb{R}^+, \leq)$ implying $\delta_{\boldsymbol{S}}$ has both equivalence and rank order meaning on $\boldsymbol{S} \times S_e$. However, a nominal scale measure, $\mu_{\boldsymbol{S}}:(\boldsymbol{S} \times S_e, =_{\boldsymbol{S}}) \rightarrow (\mathbb{R}^+, =)$, only has equivalence meaning over $\boldsymbol{S} \times S_e$. Thus, $\forall S_i, S_j \in \boldsymbol{S}$, $((S_i \leq_{\boldsymbol{S}} S_j) \vee (S_j \leq_{\boldsymbol{S}} S_i)$ (strongly complete)) and $\forall S_i, S_j, S_k \in \boldsymbol{S}$, $(((S_i \leq_{\boldsymbol{S}} S_j) \wedge (S_j \leq_{\boldsymbol{S}} S_k)) \Rightarrow (S_i \leq_{\boldsymbol{S}} S_k))$ (transitive or consistent)) cannot be discerned with $\mu_{\boldsymbol{S}}$, the nominal scale type measure. □

To further illustrate, let A be an empirical system with one element a, where the element has one attribute, α, measured by ά. Thus, the model of $A = x_A = \langle \alpha \rangle \approx \langle \mu_\alpha \rangle$ and $\boldsymbol{S}$ is the space of all possible assignments to μ_α. Further, let $S_{t=0}$, the starting state, be $x_A = \langle \alpha \rangle \approx \langle \emptyset \rangle$ and S_e, the desired end-state, be $x_A = \langle \alpha \rangle \approx \langle \psi \rangle$. Additionally, let there be two actions, y and z. At $t = T$, y results in $x_A = \langle \chi \rangle$ while z results in $x_A = \langle \varphi \rangle$. Which action, y or z, was more effective in terms of α? **THEOREM 2** asserts this question can not be answered for the nominal system state measure $\mu_{\boldsymbol{S}}$. A key result following from **THEOREM 2**, in combination with the mathematical completeness implication of **THEOREM 1**, is $(\boldsymbol{S} \times S_e, \leq_{\boldsymbol{S}})$ is of weak order. It can further be shown however, $\delta_{\boldsymbol{S}}$ not only has ordinal meaning, but has meaning on the ratio scale as well.

> **THEOREM 3:** The effectiveness measure $\delta_{\boldsymbol{S}}:(\boldsymbol{S} \times S_e, \leq_{\boldsymbol{S}}) \rightarrow (\mathbb{R}^+, \leq)$ is of ratio scale type.
>
> PROOF: The admissible transformation for a ratio scale type measure is $x \rightarrow rx$, $r \in \mathbb{R}^+$. Because S_e is used as the second parameter in each pair for the ternary relation (i.e. $(S_i, S_e) \leq_{\boldsymbol{S}} (S_j, S_e)$), S_e acts as an absolute zero for $\delta_{\boldsymbol{S}}$. Thus, $\forall r \in \mathbb{R}^+$, for the relation $\forall S_i, S_j, S_e \in \boldsymbol{S} \ni (S_i, S_e) \leq_{\boldsymbol{S}} (S_j, S_e) \rightarrow r\delta_{\boldsymbol{S}}(S_i, S_e) \leq r\delta_{\boldsymbol{S}}(S_j, S_e) \Rightarrow \delta_{\boldsymbol{S}}(S_i, S_e) \leq \delta_{\boldsymbol{S}}(S_j, S_e)$. □

To further illustrate, let A be an empirical system with one element a, where the element has one attribute α, measured by ά. Thus, the model of $A = x_A = \langle \alpha \rangle \approx \langle \delta_\alpha \rangle$ and $\boldsymbol{S}$ is the space of all possible assignments to δ_α. Further, let $S_{t=0}$, the starting state, be $A = x_A = \langle \alpha \rangle \approx \langle 9 \rangle$ and S_e, the desired end-state be $A = x_A = \langle \alpha \rangle \approx \langle 2 \rangle$. Let there be two actions, y and z. At $t = T$, y results in $x_A = \langle 6 \rangle$ while z results in $x_A = \langle 3 \rangle$. In this example, **THEOREM 3** can be used to assert y is 50% less effective than z at α in achieving S_e.

An obvious question is, why not use the starting state $S_{t=0}$, as the reference state versus S_e? The starting state does not represent an absolute zero for δ_S thus, coming into conflict with the non-negativity property of a metric (i.e. $\delta_S \geq 0$). To illustrate, using the above example with starting state $x_A = \langle \alpha \rangle \approx \langle 9 \rangle$ and desired end-state $x_A = \langle \alpha \rangle \approx \langle 2 \rangle$, at $t = T$, let y result in $x_A = \langle 6 \rangle$, which is clearly an improvement from $S_{t=0}$ since it is closer to S_e. However, suppose z results in $x_A = \langle 15 \rangle$. If the state change is measured from $S_{t=0}$, resulting in 9 - 15 = -6, the non-negativity property is violated. Additionally, in an attempt to get around the non-negativity constraint, if the measure is referenced from $t = T$ resulting in 15 - 9 = 6, while non-negative, it is now not comparable to the result from y. The logic for using the end-state as a reference point is similar to that used in goal programming where outcomes are measured with respect to the desired goal (Deckro, 1988:152).

THEOREM 1 asserts a general notion of effectiveness does not exist. Clearly however, system attribute measures, δ_{α_j} need to be mathematically combined to derive a single, scalar system effectiveness measure, δ_S. Although a single scalar facilitates comparison of system states, whether this mathematical combination has empirical significance under the Representation Theorem in Measurement Theory is questionable. For example, suppose a set of boxes is of interest. The specific attributes of interest are length, width, and height. Instead of representing each box as a vector of length, width, and height, which would be a unique representation for each box, the product of the three is used. The problem is the derived measure does not provide an isomorphic mapping from the empirical world to the formal structure (e.g. a box 40cm wide, 30cm long, and 10cm high is the same as a box 20cm wide, 60cm long, and 10cm high).

To further illustrate, let $\delta_{\boldsymbol{S}}(S_i, S_e)$ measure the conceptual distance from system state S_i, to the desired end-state, S_e. Further, suppose $\forall S_i, S_e \in \boldsymbol{S}$ are characterized by $\langle \alpha_1, \ldots, \alpha_m \rangle \approx \langle \delta_{\alpha_1}, \ldots, \delta_{\alpha_m} \rangle$. Thus, the system effectiveness measure, or derived measure, could be represented as a combination of the individual system attributes measures as follows $\delta_{\boldsymbol{S}}(S_i, S_e) = f(a_1\delta_{\alpha_1}(S_{i_{\alpha_1}}, S_{e_{\alpha_1}}), a_2\delta_{\alpha_2}(S_{i_{\alpha_2}}, S_{e_{\alpha_2}}), \cdots, a_m\delta_{\alpha_m}(S_{i_{\alpha_m}}, S_{e_{\alpha_m}}))$, where $\forall i$, 1 to m, $\delta_{\alpha_i}(S_{i_{\alpha_i}}, S_{e_{\alpha_i}})$ is the difference, or conceptual distance, between system state S_i and the desired end-state, S_e, for a specific system attribute α_i and $\forall i$, 1 to m, $a_i \in \mathbb{R}^+$ are constants associated with $\delta_{\alpha_i}(S_{i_{\alpha_i}}, S_{e_{\alpha_i}})$ indicating relevancy of the attribute. It follows,

> **THEOREM 4:** A derived effectiveness measure, $\delta_{\boldsymbol{S}}(S_i, S_e)$, from a combination of individual effectiveness measures, $\delta_{\alpha_i}(S_{i_{\alpha_i}}, S_{e_{\alpha_i}})$, is a semimetric, or pseudo metric.
>
> PROOF: $\forall S_i, S_j, S_l, S_e \in \boldsymbol{S}$ and $\forall \delta_{\alpha_k} \in \delta_{\boldsymbol{S}}$, $\delta_{\alpha_k}(S_{i_{\alpha_k}}, S_{e_{\alpha_k}}) \geq 0 \Rightarrow \delta_{\boldsymbol{S}}(S_i, S_e) \geq 0$. Additionally, $\delta_{\alpha_k}(S_{i_{\alpha_k}}, S_{j_{\alpha_k}}) = 0 \Leftrightarrow S_{i_{\alpha_k}} = S_{j_{\alpha_k}}$. However, for a derived effectiveness measure, $\delta_{\boldsymbol{S}}(S_i, S_e)$, the following, $\forall S_i, S_j \in \boldsymbol{S}$, $\delta_{\boldsymbol{S}}(S_i, S_j) = 0 \Leftrightarrow S_i = S_j$, is not a true statement since $\forall S_i, S_j \in \boldsymbol{S}$, $\exists S_i, S_j \ni \delta_{\boldsymbol{S}}(S_i, S_e) = \delta_{\boldsymbol{S}}(S_j, S_e)$ where $S_i \neq S_j$. Continuing, $\delta_{\alpha_k}(S_{i_{\alpha_k}}, S_{j_{\alpha_k}}) = \delta_{\alpha_k}(S_{j_{\alpha_k}}, S_{i_{\alpha_k}}) \Rightarrow \delta_{\boldsymbol{S}}(S_i, S_j) = \delta_{\boldsymbol{S}}(S_j, S_i)$. Finally, $\delta_{\alpha_k}(S_{i_{\alpha_k}}, S_{j_{\alpha_k}}) \leq \delta_{\alpha_k}(S_{i_{\alpha_k}}, S_{l_{\alpha_k}}) + \delta_{\alpha_k}(S_{l_{\alpha_k}}, S_{j_{\alpha_k}}) \Rightarrow \delta_{\boldsymbol{S}}(S_i, S_j) \leq \delta_{\boldsymbol{S}}(S_i, S_l) + \delta_{\boldsymbol{S}}(S_l, S_j)$. It follows, $(\boldsymbol{S}, \delta_{\boldsymbol{S}})$ is a metric space. □

Within the formal system, it has been shown, the derived measure, $\delta_{\boldsymbol{S}}(S_i, S_e)$, is a pseudo metric (**THEOREM 4**) that can be measured on a ratio scale (**THEOREM 3**). Although this makes $\delta_{\boldsymbol{S}}(S_i, S_e)$ theoretically valid, there is no evidence to show it is empirically valid, or that it holds for the Representation Theorem in Measurement Theory (Poels, 1996:11). The limiting factor is the measurement context or defining exactly what is to be learned from the act of measurement. For example, continuing with the illustration using the boxes, if the ultimate aim was to compare the volume of the boxes, the scalar representation does have empirical significance. Thus, empirical validity of the scalar system representation comes via definition of the product structure, which is inline with the result of **THEOREM 1**. Further, the derived effectiveness measure, $\delta_{\boldsymbol{S}}(S_i, S_e)$, provides a basis as an overall system effectiveness measure.

Previously, a metric, δ, was defined as a measure of distance that holds for the non-negativity, identity, symmetry, and triangle inequality properties. Clearly, numerous measures of distance can be devised to hold for these properties. For example, for a non-empty set $\boldsymbol{S}$, $\forall x, y \in \boldsymbol{S}$, $\delta(x, y) = 0$: if $x = y$, and $\delta(x, y) = 1$: if $x \neq y$ is called the discrete metric (Apostol, 1974:61). The most common metrics are derivative of the power, or Minkowski metric, which is $\delta = (\sum_{i=1}^{n} | x_i - y_i |^r)^{1/r}$, where δ is a measure of distance between entities $\boldsymbol{x}$ and $\boldsymbol{y}$ each having n attributes and $r \in \mathbb{R}^+$ is an arbitrarily chosen value (Dillon, 1984:124). To illustrate, with $r = 1$, $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $\delta(\boldsymbol{x}, \boldsymbol{y}) = | x_1 - y_1 | + \cdots + | x_n - y_n |$ is the rectilinear distance, often called the 'city-block' distance (Love, 1988:5). However, in a mathematical sense, when discussing metric spaces, one typically is addressing Euclidean space, $\mathbb{R}^n$, and the commonly used metric for $\mathbb{R}^n$ is the Euclidean metric (Suppes, 1989:32).

To further elaborate, an ordered set of $n > 0$ real numbers, $(x_1, x_2, \ldots, x_n)$, is called an n-dimensional point. The number x_k is called the k^{th} coordinate of point $\boldsymbol{x}$. The set of all n-dimensional points is called n-dimensional Euclidean space, or n-space, and is denoted by $\mathbb{R}^n$ (Apostol, 1974:47). Algebraic operations on n-dimensional points include

a) equality: $\mathbf{x} = \mathbf{y} \Leftrightarrow x_1 = y_1, \ldots, x_n = y_n$
b) sum: $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n)$
c) multiplication by real numbers (scalars): $a\mathbf{x} = (ax_1, \ldots, ax_n)$
d) difference: $\mathbf{x} - \mathbf{y} = \mathbf{x} + (-1)\mathbf{y}$

e) origin or reference vector: $\mathbf{0} = (0, \ldots, 0)$

f) inner product: $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i$.

A final operation on n-dimensional points is called length, or norm. Although there are numerous types of norms (Nash, 1996:618), the Euclidean norm, denoted by $\| \mathbf{x} - \mathbf{y} \|$ and calculated as $(\sum_{i=1}^{n} (x_i - y_i)^2)^{1/2}$, is the most common and is interpreted as the Euclidean distance between **x** and **y** (Apostol, 1974:48). Clearly, the Euclidean norm is just the power metric with $r = 2$. The Euclidean norm, as well as all power metrics, are based on four fundamental assumptions: 1) Decomposability – The distance between points, driven by system inputs, is a function of the componentwise contributions of those inputs, 2) Intradimensional Subtractivity – Each component contribution is the absolute value of an appropriate scale difference, 3) Interdimensional Additivity – The distance is a function of the sum of componentwise input contributions, and 4) Homogeneity – Affine (straight) lines are additive segments (Suppes, 1989:175). Further, the Euclidean norm has the following additional properties (Ruckle, 1991:48):

a) $\forall \boldsymbol{x} \in \boldsymbol{S}$, $\| \boldsymbol{x} \| = 0 \Leftrightarrow \boldsymbol{x} = 0$ (identity)
b) $\forall \boldsymbol{x} \in \boldsymbol{S}$ and $\forall a \in \boldsymbol{F}$, the field of scalars, $\|ax\| = | a | \| x \|$ (scalar homogeneity)
c) $\forall \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{S}$, $\| \boldsymbol{x} + \boldsymbol{y} \| \leq \| \boldsymbol{x} \| + \| \boldsymbol{y} \|$ (triangle inequality)

For completeness, $\mathbb{R}^n$, as described earlier, is a linear space. A linear space, or vector space, over a field of $\mathbb{R}^+$, $\boldsymbol{F}$, is a set $\boldsymbol{S}$ and two functions; one from $\boldsymbol{S} \times \boldsymbol{S} \rightarrow \boldsymbol{S}$, denoted by +, and one from $\boldsymbol{F} \times \boldsymbol{S} \rightarrow \boldsymbol{S}$, denoted by $\cdot$, which can be characterized by the following properties (Ruckle, 1991:31):

a) $\forall x, y, z \in \boldsymbol{S}$, $x + (y + z) = (x + y) + z$ (associative for addition)
b) $\forall x, y \in \boldsymbol{S}$, $x + y = y + x$ (commutative for addition)
c) $\exists\, 0 \in \boldsymbol{S} \ni x + 0 = x \in \boldsymbol{S}$ (unique identity)
d) $\forall x \in \boldsymbol{S}$, $\exists\, \text{-}x \in \boldsymbol{S} \ni x + (\text{-}x) = 0$ (unique inverse)

e) $\forall a, b \in \boldsymbol{F}$ and $\forall x \in \boldsymbol{S}$, $a(bx) = (ab)x$ (associative for multiplication)
f) $\forall a, b \in \boldsymbol{F}$ and $\forall x \in \boldsymbol{S}$, $(a + b)\, x = ax + bx$ (right distributive)
g) $\forall a \in \boldsymbol{F}$ and $\forall x, y \in \boldsymbol{S}$, $a(x + y) = ax + ay$ (left distributive)
h) $\forall x \in \boldsymbol{S}$, $1x = x$ (multiplicative identity)

Finally, the norm has the following properties on a vector space, or more precisely on a normed vector space, $\mathbb{R}^n$ (Apostol, 1974:48):

a) $\| \boldsymbol{x} \| \geq 0$ (non-negativity) and $\| \boldsymbol{x} \| = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}$ (identity)
b) $\| a\boldsymbol{x} \| = | a | \, \| \boldsymbol{x} \| \; \forall a \in \mathbb{R}$ (scalar homogeneity)
c) $\| \boldsymbol{x} - \boldsymbol{y} \| = \| \boldsymbol{y} - \boldsymbol{x} \|$ (symmetry)
d) $| \boldsymbol{x} \cdot \boldsymbol{y} | \leq \| \boldsymbol{x} \| \, \| \boldsymbol{y} \|$ (triangle inequality for dot product)
e) $\| \boldsymbol{x} + \boldsymbol{y} \| \leq \| \boldsymbol{x} \| + \| \boldsymbol{y} \|$ (triangle inequality for addition)

Although the Euclidean norm serves as a robust and convenient way to aggregate measures, the units of the attributes will not likely be mathematically commensurate, or of the same magnitude in their initial form and thus, will require a transformation in order to be aggregated. Comparison of system states relative to an end-state implies individual system attributes are aggregated to make an overall statement about the system. Aggregation presents a special problem for the proposed effectiveness measurement framework, since the measures *will* likely be in different units and have differing magnitudes. Not addressing this issue of non-commensurate measures will result in a systemic error since combination of dissimilar measurements results in certain system attributes having a higher proportional weighting relative to other system attributes.

This aggregation problem is not unique to the proposed framework and is usually handled via a normalization transformation before aggregation of the measurements. In general, normalization is a mathematical transformation that maps from one scale to another, yielding a common scale. Numerous normalization techniques exist that will make dissimilar measures commensurate for purposes of aggregation. Some of these methods include percentage normalization and summation normalization (Tamiz, 1998:572). The most common techniques, however, attempt to scale each attribute to a common scale of zero to one and go by names such as 'zero-one' (Tamiz, 1998:573) or 'Bowles' (Zuse, 1998:232) normalization. For example, the normalization technique most commonly used in the literature is $\delta' = \frac{\delta - \delta_{\min}}{\delta_{\max} - \delta_{\min}}$, where δ is the value to be scaled and $\delta_{\min}$ and $\delta_{\max}$ are respectively the minimum and maximum values δ can be assigned where $\delta_{\max} - \delta_{\min} \neq 0$ (Kirkwood, 1997:58). Another technique often used when $\delta_{\min}$ and $\delta_{\max}$ are not known, but also produces a result from zero to one, can be calculated as $\delta' = \frac{\delta}{\sqrt{\delta^2 + a}}$, where $a \in \mathbb{R}^+$ is chosen arbitrarily large relative to δ.

These normalization techniques are useful in making dissimilar scales commensurate for purposes of aggregation. However, simply applying the transformation does not address all the issues. One issue concerns the meaning (i.e. scale type) associated with the numbers before and after the normalization transformation. Most normalization techniques result in reduced meaning after the transformation. Specifically, ratio meaning is usually lost (Kirkwood, 1997:241; Zuse, 1998:232). For example, let $\delta_1 = 17$ and $\delta_2 = 13$ be observations of metrics and thus of ratio scale type. The empirical relationship between these observations with ratio level meaning is $\frac{\delta_1}{\delta_2} = \frac{17}{13} = 1.30$. Assume $\delta_{\min} = 10$ and $\delta_{\max} = 20$. Thus, $\delta_1' = .7$ and $\delta_2' = .3$. Examining the empirical relationship after the normalization, $\frac{\delta'_1}{\delta'_2} = \frac{.7}{.3} = 2.33 \neq 1.30$, shows the ratio scale meaning was lost in the transformation.

Another issue involves the meaning of the measurements within a specific context. For example, a decision maker needs to evaluate projects in a portfolio for possible termination. Two projects are found that have exceeded their budgets by $1,000,000. For the cost attribute, each program is a distance of $1,000,000 from their respective desired end-states. However, let one of the programs have an original budget of $1,000,000 and the other have original budget of $200,000,000. Even though in general, the two projects are an equal distance from their end-states, from the decision maker's perspective, the interval distance from $1 million to $2 million likely has a different meaning from the interval distance from $200 million to $201 million. Further, simply looking at the distance as a percentage of the end-state may not yield equivalent distances (Keeney, 1992:115). It follows, the meaning (i.e. scale type) of numbers is context dependent (Kirkwood, 1997:241).

THEOREM 3 asserted effectiveness measures, as defined within the proposed framework, have ratio level meaning. A fundamental property of a ratio measure, building upon the properties of interval measures, is interval distances are equal (Stevens, 1946:679). From an applied standpoint, **THEOREM 3**, suggesting measures have ratio scale type in general, and the earlier statement about numerical meaning being context dependent, seems to be in conflict. This suggests, to achieve ratio level meaning, models of some system attributes (i.e. measures) may require a scale transformation to convert empirical observations such that they yield scales with equal intervals. For example, assume the following scenario:

> A specific system attribute is being monitored. A measure for the attribute has been developed with a lower bound of zero. Additionally, the desired end-state for the attribute has been defined as 10. Further, for the specific context, it is known the following relationship exists: a system attribute

value above 10 is twice as desirable as a value below 10. This implies an observed unit interval below 10 is equal to two observed unit intervals above 10.

For this scenario, two key issues have to be addressed before the problem of normalization for this measure can be solved. The first issue concerns the unequal unit intervals above and below the desired end-state (10). Since the relationships between the intervals are known, this problem can be handled with a scale transformation. For example, let $X_{OBSERVED} \in \mathbb{R} \geq 0$, be the observed system attribute measure (Figure 9). Further, let $X_{EQUAL} \in \mathbb{R} \geq 0$, be the equal interval transformation developed using the known relationship (Figure 10). $\mathbf{X_{EQUAL}}$ yields empirical observations with ratio level meaning. Clearly, the relationship presented in the scenario will not be known in general but will have to be discovered. This discovery process occurs by asking the decision maker, who will be making decisions based off the measurements, or subject matter experts on the system of interest, a series of lottery or certainty equivalent questions (Luce, 1957:21; Keeney, 1992:6) to identify indifference curves (Keeney, 1992:79; Clemen, 1996:540). This topic of achieving equal intervals is related to the concept of differentially value equivalence (Keeney, 1993:94) and is just an extension of the substitutability axiom of expected utility (Clemen, 1996:504).

The second issue concerns the non-monotinicity as a function of the observed values. Non-monotinicity suggests benefits or utility is not an increasing (or decreasing) function of observed values (i.e. if more is good, a lot more may not necessarily be better). The above scenario, even after adjusting for equal intervals, is non-monotinic since desirability of the system attribute is increasing from 0 to 10 and decreasing from 10 to ∞^+. However, the benefits or utility is not relative to the scale origin (0), but relative to the end-state (10), or more specifically the distance to the end-state. Measures of distance (i.e. metrics) are always monotonic (Apostol, 1974:60). This implies, all else being equal for the above scenario, a decision maker would be indifferent between system attribute values of 5 and 20, on the $X_{OBSERVED}$ scale, since they are the same equal interval distance from the desired end-state based on known empirical relationships.

These examples concerning budget overruns only looked at a single attribute (cost). However, when looking at multiple attributes simultaneously, or in other words, a derived effectiveness measure, **THEOREM 4** suggested the distance from a system state to the desired end-state was not unique to the state (i.e. a semi-metric). In light of **THEOREM 4** and the preceding discussion, it follows, strategically equivalent system states are equidistant from the desired end-state.

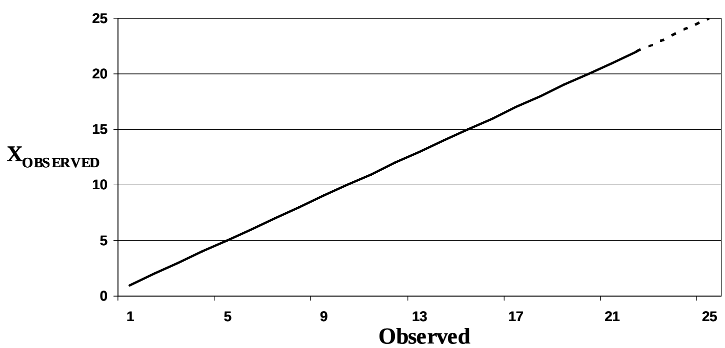


Figure 9. Observed System Attribute Assignments

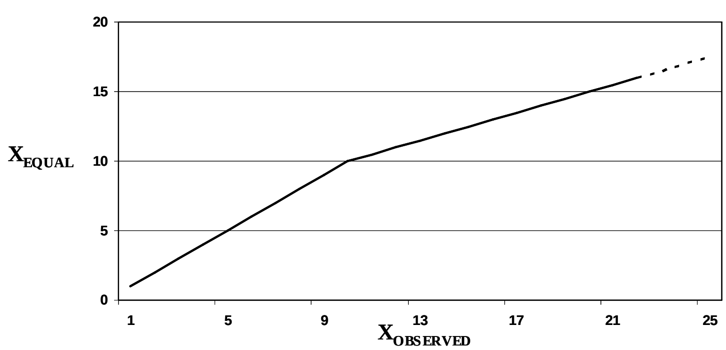


Figure 10. Observed After Equal Interval Transformation

A desirable characteristic of a normalization transformation is to preserve the scale meaning of the input values. Under the proposed effectiveness measurement framework, **THEOREM 3** asserted effectiveness measures have ratio scale meaning. The only allowable transformation that preserves ratio level information is multiplication by a scalar, $x \rightarrow rx,\ r \in \mathbb{R}^{+}$. A logical normalization approach, given the proposed framework, which adheres to this transformation, is to normalize the (equal interval) distance from the end-state with respect *to* the end-state ($X_{DESIRED}$). Following development of a system attribute measure, rules for such a normalization approach are outlined in Figure 11. A detailed example illustrating implementation of the technique on notional data is provided in Appendix B.

1.	Identify the desired end-state	$X_{DESIRED}$
2.	Establish certainty equivalent transformation	$X_{OBSERVED} \rightarrow X_{EQUAL}$
3.	Calculate distance from current system state to desired end-state	$X_{DISTANCE} = \lvert X_{EQUAL} - X_{DESIRED} \rvert$
4.	Normalize distance with respect to the end-state using a ratio preserving transformation	$X_{NORMALIZED} = k_j X_{DISTANCE}$
5.	Update previous observations using latest normalization constant (k_j)	$k_j = \dfrac{1}{\sum_{i=1}^{j} X_{DISTANCE}}$

Figure 11. Ratio Preserving Normalization

This proposed deterministic Theory of Effectiveness Measurement can be summarized as follows (Figure 12). Starting with the system state specification, or product structure, the system of interest is identified and, in particular, the system boundary is delineated. Continuing with the specification, the system model is developed to include all pertinent dimensions of the system. This is required, because as asserted in **THEOREM 1**, there is no general notion system effectiveness. Further, a key aspect required for the system model is for all developed measures to hold for the properties of a metric (i.e. non-negativity, identity, symmetry, and the triangle inequality). This step is not always straight forward, since meaning (i.e. scale type) is context dependent.

It follows, an instantiation of all system measures is an observation, or measurement, of the system yielding the system state. A crucial philosophical view, used in the theory presented here, is a change from one possible system state to another possible system state is an effect. By representing a system state as an *n*-dimensional vector, corresponding to each of the relevant system attributes (i.e. dimensions), the space of these points serves as the space of all possible system consequences. Clearly, each system dimension is a metric space via the definition of the product structure. However, as asserted in **THEOREM 4**, the space of all possible system states is also a metric space from which it follows, system states equidistance from the desired end-state are strategically equivalent. Further, a key result from **THEOREM 1** is that the space of all possible system states is strongly complete. Continuing, it follows from **THEOREM 2**, via the triangle inequality, the system state space is transitive. Combination of the transitivity and strongly complete properties yield another property, namely the weak order property. The significance of this derived property is that the state space being a metric space along with having weak ordering are sufficient conditions for the system state space to be a proximity structure (Suppes, 1989:162).

The state space being a proximity structure introduces the properties of positivity, minimality, and symmetry which are essentially reflections of the metric properties. These, along with weak ordering, allow for quaternary relations on the proximity structure (i.e. $(S_i, S_k) \leq_S (S_j, S_m)$). Under the proposed framework, however, the quaternary relations reduce to ternary relations since each side of the relation has a common parameter (i.e. the desired end-state, S_e, yielding $(S_i, S_e) \leq_{\boldsymbol{S}} (S_j, S_e)$). Finally, **THEOREM 3** suggests the state space is of ratio scale type allowing for meaningful comparison of inputs yielding system state changes (i.e. $\delta_{\boldsymbol{S}}:(\boldsymbol{S} \times S_e, \leq_{\boldsymbol{S}}) \rightarrow (\mathbb{R}^+, \leq)$).

The deterministic framework provides a set of necessary and sufficient conditions for conducting effectiveness measurement. However, by virtue of being 'deterministic', implied is the assumption of perfect information (i.e. what was seen is what actually happened). Clearly, application of measurement in any domain needs to address error and uncertainty, where uncertainty relates to the amount of knowledge available concerning a system attribute and error is the deviation of a system attribute measurement from the true, but unknown, value (Weise, 1992:1).

PROBABILISTIC FRAMEWORK

Error and uncertainty, as noted earlier, manifests itself in three forms to include observational, systemic, and random. With respect to the proposed effectiveness measurement framework, these forms, and their impact, are exemplified in Figure 13. In Figure 13, observational error is illustrated as germane system attributes not being identified. These missing attributes, in turn, do not appear in the system measurement model or the system vector representation. Additionally, random error is shown as an interval around a measured value in the system state estimate, $\boldsymbol{x}^*_{\boldsymbol{A}}$. Further, systemic error is portrayed as a shifted, or biased, interval around the observed value. Finally, Figure 13 displays the impact of these errors as an overall, unperceived error between the actual and observed system state. Thus, to address these errors and complement the deterministic framework, a probabilistic framework is needed for reasoning about these types of error and uncertainty while conducting effectiveness measurement.

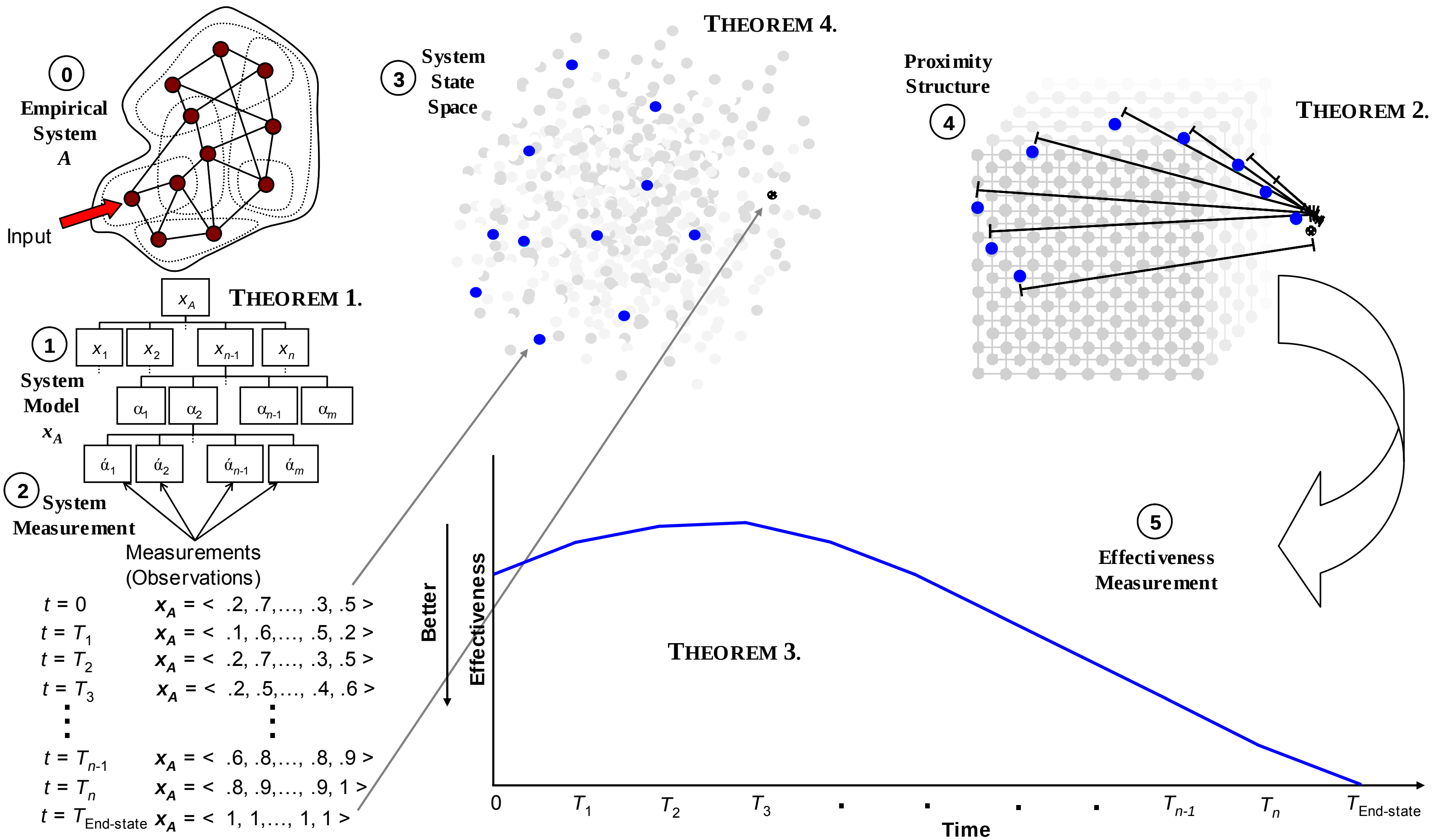


Figure 12. Framework Summary

Numerous probabilistic reasoning frameworks exist. These frameworks are essentially tools to address uncertainty and error in particular types of problems. Thus, what follows is a brief overview of some probabilistic reasoning techniques to identify a preferred approach to support the proposed deterministic effectiveness measurement framework.

In the application of measurement, it is unlikely everything about a domain of interest will be known. Because of this, it is not even possible to precisely quantify what is unknown. To get around this problem, Probability Theory can be used to generalize the unknown by assigning a degree of belief, or probability, to what is known (Weise, 1992:2). Additionally, domain knowledge consists of known truths about the domain of interest (Grassmann, 1996:60). It should be noted, however, 'degree of belief' is not the same as 'degree of truth' which is the realm of fuzzy logic (Russell, 2003:464). Assigning a degree of belief to a measurement implies an underlying distribution associated with all possible instantiations of the measure across the universe of discourse for the measure (Russell, 2003:469). The assignment can be based on different philosophical views concerning probability including empirical evidence (frequentist), proven theoretical assertions (objectivist), or a characterization without physical significance (subjectivist) (Russell, 2003:472). Regardless of how they are derived, these assignments form the basis for most probabilistic inference techniques.

One of these techniques is based on Dempster-Shafer Theory (DST). DST is closely aligned with the frequentist view in that, instead of computing the likelihood of an event based on theoretical hypotheses or expert opinion, DST derives probabilities

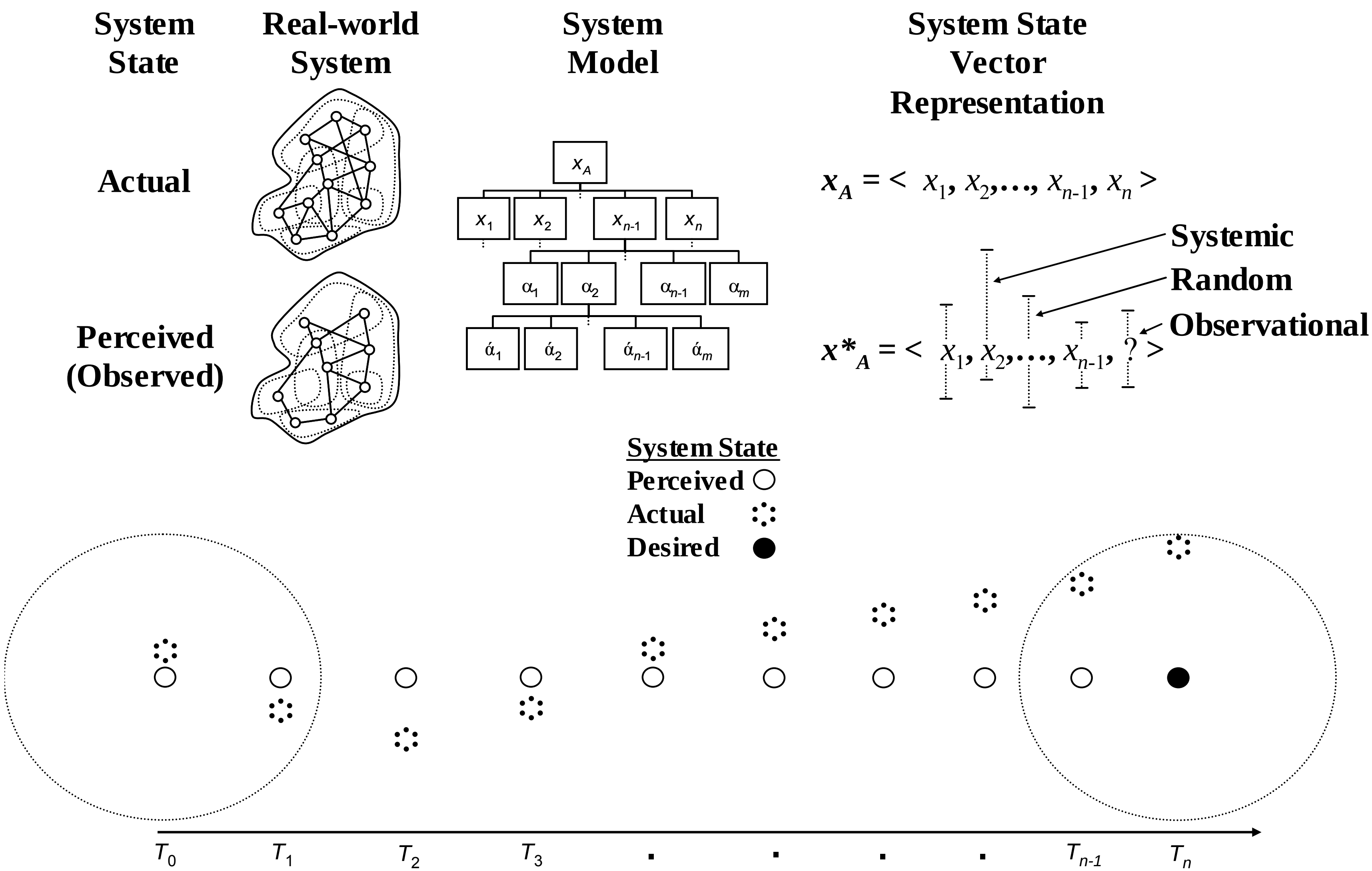


Figure 13. Error and Uncertainty in Effectiveness Measurement

based on measurements, or evidence, supporting a particular assertion about the domain of interest. This point suggests DST may be a good alternative for developing the probabilistic portion of an effectiveness measurement framework. However, many aspects of DST are not well understood and require further research (Russell, 2003:526).

Another probabilistic reasoning approach is based on Fuzzy Set Theory. Fuzzy Set Theory provides a means for specifying how well a system attribute meets the criteria of a given specification (Russell, 2003:526). A key feature of Fuzzy Set Theory lending itself for use in a measurement framework is its ability to handle qualitative, real-world observations without the need for precise system attribute quantification. However, this benefit is offset by representation problems of qualitative observations given subjective classification criteria (Russell, 2003:527). Additionally, as noted above, Fuzzy Set Theory is based on degrees of truth versus degrees of belief. From a real-world decision making point of view, this can lead to interpretation problems. For example, in Fuzzy Set Theory, the answer to the question, "Are we winning or losing?" is always, "both".

One inference framework closely aligned with the above deterministic effectiveness measurement framework, is known as filtering. Filtering uses system state representation in the form of vectors and is often used where the internal behavior of a system cannot be observed or is not known and must be inferred from the system's external behavior (Maybeck, 1979:4; Welch, 2004:1). In practice, filtering is the task of computing the current state of a system in the face of uncertainty as well as partial and noisy measurements (Zarchan, 2005:91). Mathematically, filtering is a recursive estimation technique and takes the form (Russell, 2003:541):

$$\mathbf{P}(\hat{\mathbf{x}}_k \mid \mathbf{x}^*) = f(\mathbf{x}^*_k, \mathbf{P}(\hat{\mathbf{x}}_{k-1} \mid \mathbf{x}^*)) \quad (8)$$

In other words, an estimate of the state of the system at the k^{th} measurement, $\hat{\mathbf{x}}_k$, given all measurements of the system, $\mathbf{x}^*$, is a function, f, of the latest measurement, $\mathbf{x}_k^*$, and the previous system state estimate, $\mathbf{P}(\hat{\mathbf{x}}_{k-1} \mid \mathbf{x}^*)$. Thus, even if the system of interest is not tangible, such as the collective will of a group of people, via filtering we could use existing measurements to estimate the current state of the system.

A popular filter for trying to estimate the state of a system based on uncertain and error prone measurements is known as the Kalman filter, first presented by Rudolf E. Kalman in 1960 (Kalman, 1960). The Kalman filter can take 'noisy' measurements and estimate the state of any system (Maybeck, 1979:4). A key assumption of the Kalman filter is the current system state estimate, $\hat{\mathbf{x}}_k$, is a linear function of the previous state estimate, $\hat{\mathbf{x}}_{k-1}$, plus some Gaussian noise (Maybeck, 1979:7). This is a reasonable assumption under the Central Limit Theorem. Specifically, as the number of measurements increases, the distribution tends to be Gaussian. Additionally, for the likely case of tracking numerous system attributes, the sum of independent random variables, regardless of individual density function, tends toward Gaussian as the number of random variables gets larger (Maybeck, 1979:8). In relation to mathematical techniques, the Kalman filter is essentially a Bayesian estimator that uses all available measurements, and their covariance, to arrive at a system state estimate (Maybeck, 1979:114).

Application of the Kalman filter assumes the system of interest can be described by a set of differential equations. Additionally, the equations must be in state-space notation. State-space notation implies any set of linear differential equations can be put into the form of the first-order matrix equation:

$$\hat{\mathbf{x}} = \boldsymbol{F}\boldsymbol{x} + \boldsymbol{G}\boldsymbol{u} + \boldsymbol{w} \tag{9}$$

where $\boldsymbol{x}$ is the system state vector, $\boldsymbol{F}$ is the system dynamics matrix, $\boldsymbol{u}$ is a deterministic input called a control vector, and $\boldsymbol{w}$ is a random forcing function, which is also known as process noise (Zarchan, 2005:33). It should be noted, $\boldsymbol{G}$ captures the relationships between the controls, $\boldsymbol{u}$, and the system states. However, since these relationships are commonly unknown, many Kalman filtering implementations set $\boldsymbol{G}$ to 0 (Zarchan, 2005:131).

In order for the above matrix differential equation to be used as a filter, it must be discretized, with measurements taken at a periodicity of T_s. This is achieved by deriving a fundamental matrix, $\boldsymbol{\Phi}$, via the system dynamics matrix, $\boldsymbol{F}$, using the following relationship $\boldsymbol{\Phi} = \mathcal{L}^{-1}\,[(s\boldsymbol{I} - \boldsymbol{F})^{-1}]$, where $\mathcal{L}^{-1}$ is the inverse Laplace transform and $\boldsymbol{I}$ is the identity matrix. In general, the solution to $\boldsymbol{\Phi}$ for a n-1 order filter is (Grewal, 1993:37):

$$\begin{bmatrix} 1 & T_s & \frac{1}{2}T_s^2 & \frac{1}{1\cdot 2\cdot 3}T_s^3 & \cdots & \frac{1}{(n-1)!}T_s^{n-1} \\ 0 & 1 & T_s & \frac{1}{2}T_s^2 & \cdots & \frac{1}{(n-2)!}T_s^{n-2} \\ 0 & 0 & 1 & T_s & \cdots & \frac{1}{(n-3)!}T_s^{n-3} \\ 0 & 0 & 0 & 1 & \cdots & \frac{1}{(n-4)!}T_s^{n-4} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{10}$$

Additionally, Kalman filtering assumes measurements are linearly related to the system states via a measurement matrix $\boldsymbol{H}$ along with an associated measurement noise $\boldsymbol{v}$. Finally, constants $\boldsymbol{K}$, also called Kalman gains, are needed to express the relationship between the new measurement and the current estimate. However, to calculate the gains, the errors in the state estimates before and after the most recent system measurement must be taken into account. This is accomplished using a covariance matrix, $\boldsymbol{M}_k$, representing the error before the measurement and a covariance matrix, $\boldsymbol{P}_k$, representing the error after the measurement in the following set of recursive matrix equations (Grewal, 1993:112):

$$\begin{aligned} \boldsymbol{M}_k &= \Phi_k \boldsymbol{P}_{k-1} \Phi_k^{\mathrm{T}} + \boldsymbol{Q}_k \\ \boldsymbol{K}_k &= \boldsymbol{M}_k \boldsymbol{H}^{\mathrm{T}} (\boldsymbol{H} \boldsymbol{M}_k \boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R}_k)^{-1} \\ \boldsymbol{P}_k &= (\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H}) \boldsymbol{M}_k \end{aligned} \tag{11}$$

It should be noted, $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$ relate to the process noise, $\boldsymbol{w}$, and the measurement noise, $\boldsymbol{v}$, respectively; specifically, through the following relationships:

$$\begin{aligned} \boldsymbol{Q} &= \mathbf{E}[\boldsymbol{w}\boldsymbol{w}^{\mathrm{T}}] \\ \boldsymbol{R} &= \mathbf{E}[\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}] \end{aligned} \tag{12}$$

Together, the above elements yield the Kalman filter equation (Zarchan, 2005:131):

$$\hat{\mathbf{x}}_k = \Phi_k \hat{\mathbf{x}}_{k-1} + \boldsymbol{K}_k (\mathbf{x}_k^* - \boldsymbol{H} \Phi_k \hat{\mathbf{x}}_{k-1}) \tag{13}$$

A key strength of the Kalman filter is there are no parameters requiring tuning for a particular problem. However, the order of the Kalman filter should fit the order of the real-world system, which is typically not known. The tradeoff is lower order filters are better at reducing measurement noise error in an estimate; however, lower order filters can also result in significant truncation error, a form of systemic error (Zarchan, 2005:127). For example, assume an unknown, system attribute behavior is actually a sine wave. A sine wave will be used since it is a familiar signal, but it also provides a challenging, nonlinear behavior to estimate with a Kalman filter. Additionally, for purposes of illustration, assume there is up to twenty-five percent 'noise' in the measurements, on which system estimates will be based.

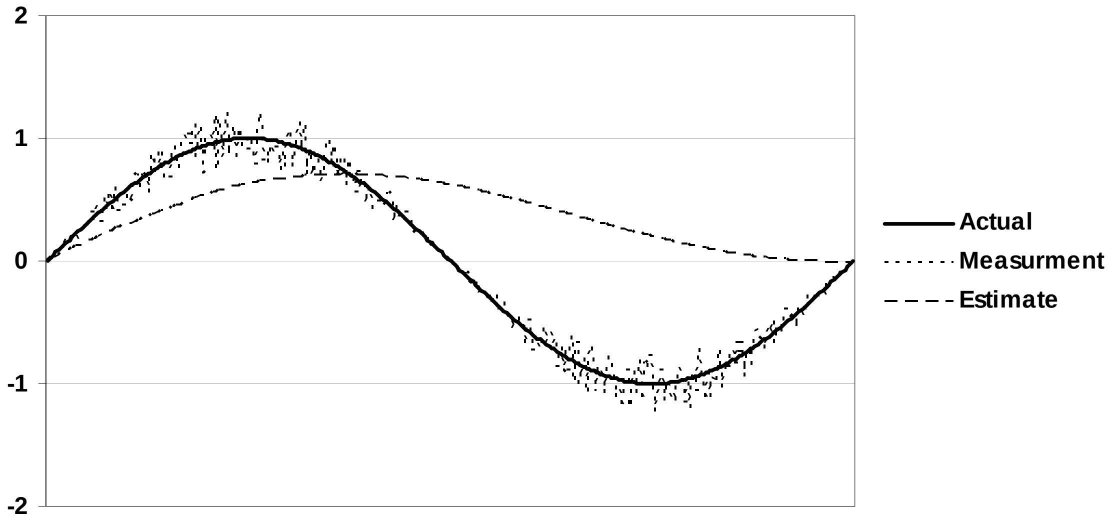


Figure 14. 0^{th} Order Kalman Filter Estimate of a Sine Wave

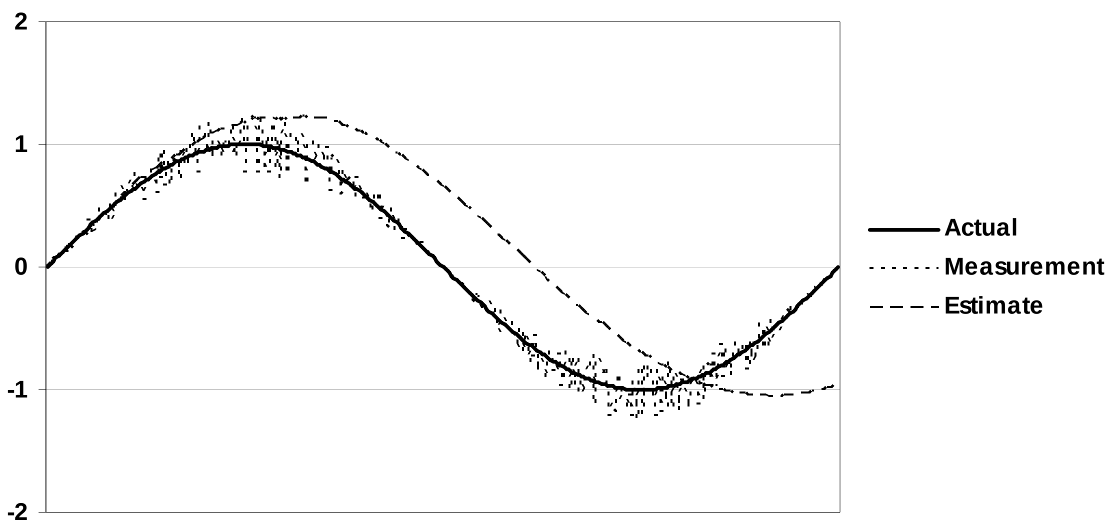


Figure 15. 1^{st} Order Kalman Filter Estimate of a Sine Wave

As a first attempt to estimate the unknown system attribute behavior, a 0^{th} order filter will be used. The results are shown in Figure 14. Clearly, the 0^{th} order filter would have yielded poor feedback on the behavior of the actual system. Next, a 1^{st} order filter is used. As can be seen in Figure 15, the 1st order filter does better at tracking the underlying system behavior, but significantly lags the actual behavior. As a final attempt to estimate the behavior, a 2nd order filter is used. The results are shown in Figure 16, in which the filter estimates the unknown behavior accurately through the first half off the sine wave. In the second half, however, the 2nd order filter tracks the underlying behavior trend, but with a significant divergence from the true values.

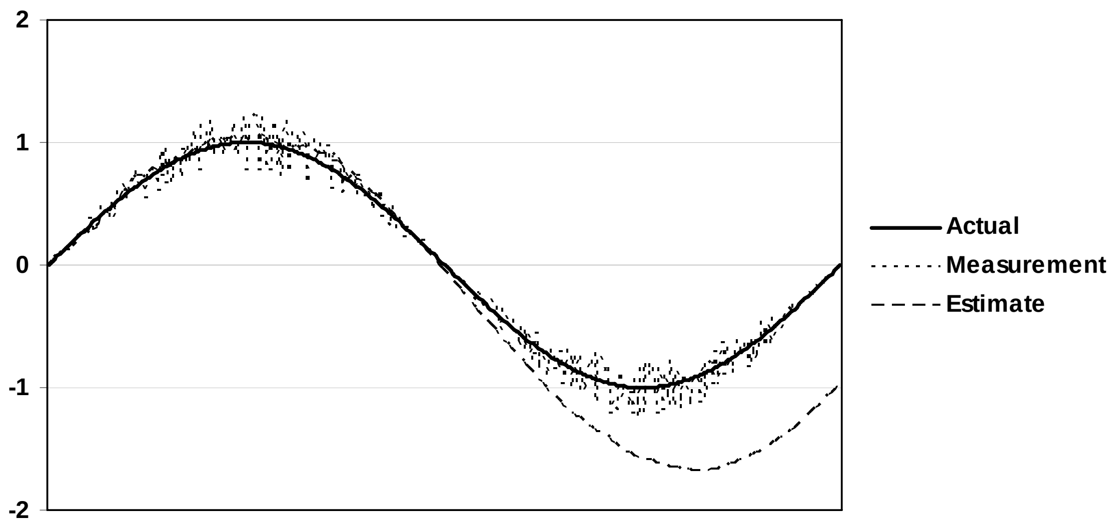


Figure 16. 2nd Order Kalman Filter Estimate of a Sine Wave

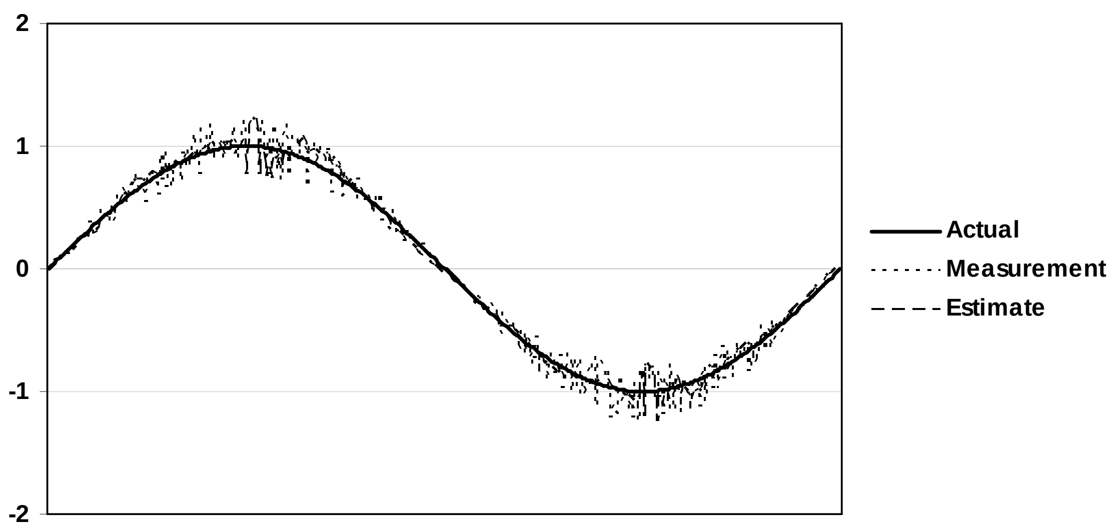


Figure 17. Fading Memory 2nd Order Kalman Filter Estimate of a Sine Wave

One approach to dealing with this divergence is called ‘fading memory’ (Maybeck, 1982:28). In the fading memory technique, only the most recent measurements are used to estimate the system state. The impact of this approach using an arbitrarily selected 90-period memory is shown in Figure 17. While Figure 17 shows a very close correlation between actual system behavior and the estimate, it should be noted, the need for the fading memory technique to address the estimate divergence from the actual system behavior, was driven because the actual system behavior was known to be a sine wave.

Since the true states of a system will not likely be available to validate a filter, a conservative approach is to use a second-order filter which is equivalent of keeping track of the position, velocity, and acceleration of system attribute movements. Thus, the matrix form of a second-order Kalman filter for a single system attribute, *x*, is:

$$\begin{bmatrix} \hat{x}_k \\ \hat{\dot{x}}_k \\ \hat{\ddot{x}}_k \end{bmatrix} = \begin{bmatrix} 1 & T_s & .5T_s^2 \\ 0 & 1 & T_s \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{x}_{k-1} \\ \hat{\dot{x}}_{k-1} \\ \hat{\ddot{x}}_{k-1} \end{bmatrix} + \begin{bmatrix} K_{1_k} \\ K_{2_k} \\ K_{3_k} \end{bmatrix} \left[x_k^* - \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & T_s & .5T_s^2 \\ 0 & 1 & T_s \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{x}_{k-1} \\ \hat{\dot{x}}_{k-1} \\ \hat{\ddot{x}}_{k-1} \end{bmatrix} \right] \quad (14)$$

An intuitive feature of the Kalman filter is errors in the estimates decrease as the number of measurements taken increases (Zarchan, 2005:148). It follows from this result, process noise can be assumed to be zero as more measurements are taken, which simplifies and allows for off-line calculation of the Kalman gains, ***K*** (Zarchan, 2005:156). A detailed example illustrating use of the second-order Kalman filter appears in Appendix C.

Finally, it should be emphasized, the above linear filter will be used to estimate the state of what is likely a non-linear, real world system. However, even if a model of the non-linear system was available, the above, basic linear filter is very robust and performs just as well as various non-linear Kalman filter variants, especially when the true underlying nature of the system is unknown (Zarchan, 2005:291, 329). Regardless, real-world non-linearity does reduce system state estimate accuracy as illustrated in Figure 16. This problem can best be addressed by increasing the periodicity, T_s, at which measurements are taken (Zarchan, 2005:291, 677). In other words, the more measurements one takes, the more accurate the estimate. For example, Figure 18 displays the result of estimating the sine wave based on the same measurements shown in Figure 16, but using $T_s = .\overline{3}$ versus $T_s = 1$.

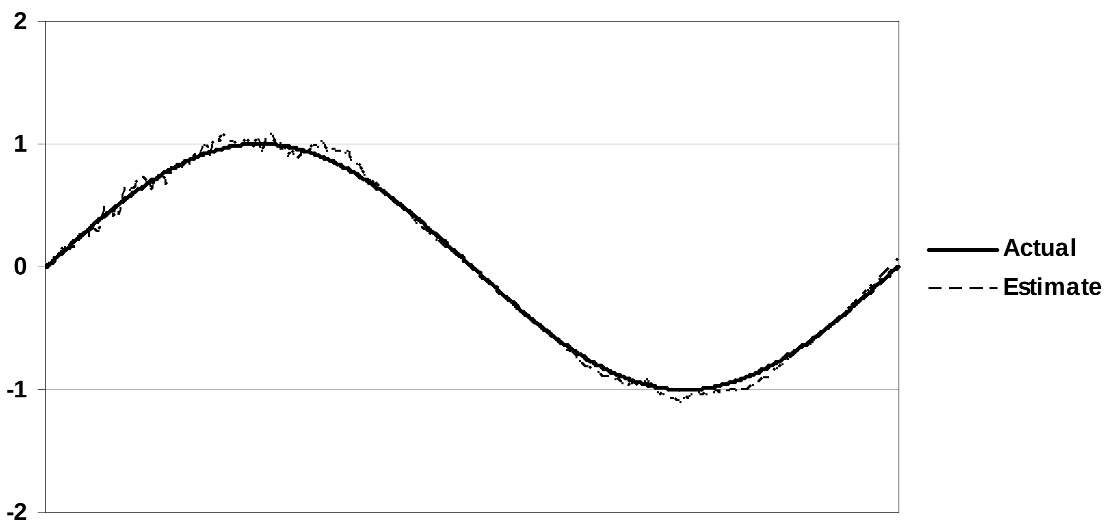


Figure 18. Impact of Increased Sample Rate on Sine Wave Estimate

IMPLEMENTATION OF FRAMEWORKS

The motivating driver for this research was to meet the needs of the practitioner, tasked with measuring progress towards abstract objectives, with the proper theory. While the above deterministic and probabilistic frameworks provide the proper theory, to be of use to the practitioner, the frameworks must also be pragmatic. The purpose of this section is to demonstrate the practical nature of the frameworks by implementing them on a simple, but realistic scenario.

A number of approaches are available to highlight the usefulness of the proposed theory. One of the alternatives explored included implementing the frameworks in a generic scenario using a simple, toy model developed using system dynamics or causal analysis techniques. While the simplicity would have provided transparency, the generic nature of the approach did not provide a level of realism likely required to which a practitioner could relate.

Another alternative examined was to use data collected on a historical battle. Numerous data sets exist such as those used in proving the theoretical assertions contained in The War Trap (Bueno de Mesquita, 1981). Unfortunately, most available datasets, as in The War Trap, only provide visibility on the starting state and the end-state, with no insight on events traversed between the two states. However, a few datasets do contain this level of fidelity. Highly detailed and comprehensive datasets on the World War II Battles of Kursk and Ardennes are available from the United States National Technical Information Service. While these datasets do provide time series data between the starting state and the end-state for individual units, the datasets are attrition oriented detailing only the unit location and strength level for each day of the respective battles. There is no insight into the cause for certain movements or declines in force strength.

A final alternative investigated, and ultimately used, for illustrating use of the frameworks, was to employ a high fidelity model depicting a realistic scenario. Numerous high fidelity models exist for the purposes of analysis and wargamming.

However, many of these models, such as the THUNDER campaign level warfare model and the Combat Forces Assessment Model (CFAM), are attrition oriented which is not in alignment with the tenets of EBO. One model though, specifically developed to support the concepts behind EBO is called Point of Attack 2 (POA2).

POA2 is a comprehensive and detailed, modern, tactical level, combat simulator that depicts engagements at the platoon and individual vehicle level, along with complete characterization of supporting artillery, air strikes, electronic warfare, engineering, chemical warfare, helicopter, naval, and psychological operations units (HPS, 2006). POA2 was designed to model the capabilities and effects of conventional weapons as well as developing technologies. POA2 was developed by HPS Simulations via funding from the Plasma Physics Program of the US Air Force Office of Scientific Research. The focus of the development effort was to create a state-of-the-art strategy wargame specifically designed to capture the effects of non-traditional weapons such as a high powered microwave (AFOSR, 2001).

While using an effects oriented model, such as POA2, was a necessary condition for demonstrating the theory, another critical factor was the scenario to be portrayed. Of key importance was finding a scenario that would highlight the strengths and limitations of the proposed theory. An additional characteristic was finding a scenario that would resonate with the practitioner. POA2 comes with several preprogrammed scenarios. One of these scenarios was selected and modified, as illustrated in what follows, for the purposes of demonstrating the frameworks. The scenario, involving a terrorist attack on a Continental United States Air Force Base, is highlighted here:

> Extremist attempt to breach a southwestern United States airfield with truck bombs and car bombs in an effort to destroy aircraft near a runway

as well as blowup a fuel depot. Additionally, using the truck/car bomb explosions as cover, extremist squads in off-road vehicles, try to infiltrate laboratories developing critical, near ready to be fielded technologies to support the Global War on Terror, in an effort to steal the technologies, destroy the laboratories supporting the technologies, and kill the people creating them. Base security forces, unaware of the impending attack, respond.

The base security forces (BLUE forces) are composed of the following objects, using the default characteristics and properties, as defined in the POA2 software:

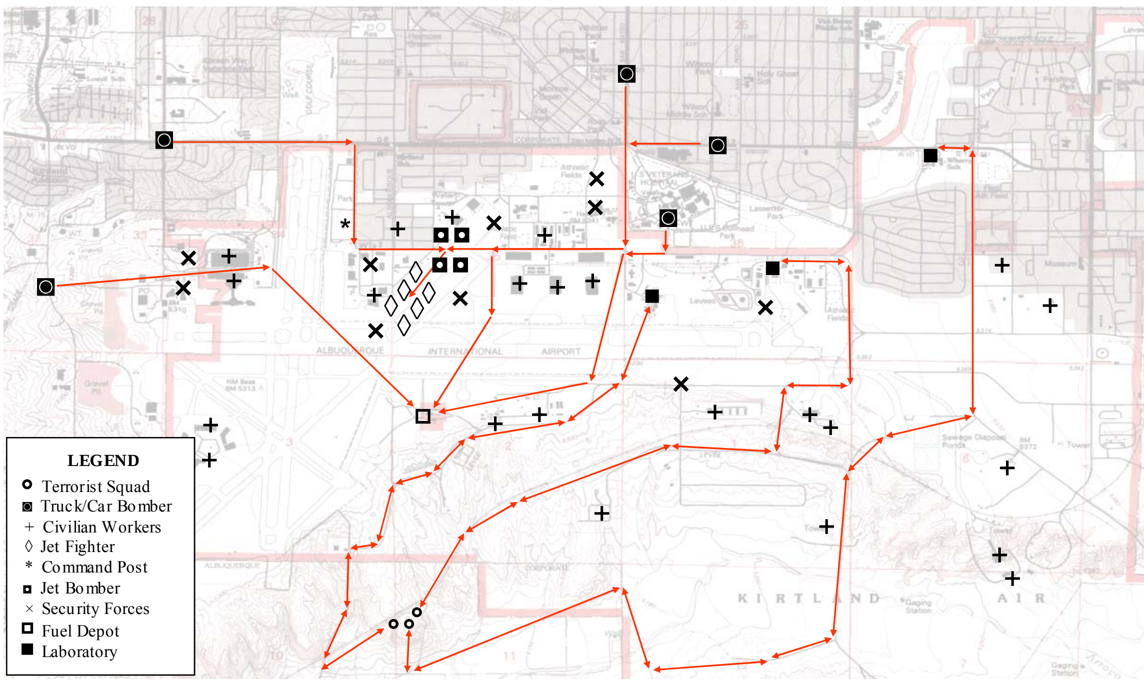


Figure 19. Scenario Details

1 x Command Post
10 x Armored Vehicle (HMMWV 988A2)
30 x Military Police (R)
4 x Parked Bomber (B-1)
6 x Parked Fighter (F-16)
1 x Fuel Depot (Large Masonry Building)
3 x Laboratory (Large Masonry Building)
195 x Civilian

Similarly, using existing objects as defined in the POA2 software, the extremist forces (RED forces) are:

2 x Truck Bomber
3 x Car Bomber
3 x Off-road Vehicle (CUCV 4 x 4)
9 x Terrorist (Infantry (R))

Details of the scenario are illustrated in Figure 19. Specifically, the fighters (◊) and the bombers (◘) are parked in the open, but are guarded by military police in armored vehicles (×). Additionally, military police in armored vehicles (×) are positioned near the major base entry gates. These gates are the entry points for the truck and car bombers (◙). While the fighters (◊) and the bombers (◘) are key targets for the truck and car bombers (◙), destruction of the fuel depot (□) is another target of the extremists. The paths to be traversed by the truck and car bombers (◙) to the fighters (◊), bombers (◘), and the fuel depot (□) are indicated by the arrows (→). The base security forces are controlled centrally through the command post (*), which can result in delays in the receipt and distribution of intelligence information as well as delays in the transmission of updated orders. While the extremists would consider completed attacks on the fighters (◊), bombers (◘), and the fuel depot (□) a victory, their true intent is to obtain advanced technologies being developed in laboratories (■) on the base. Terrorists in off-road vehicles (○) will traverse the paths indicated by the arrows (→) across the base to the laboratories (■) to obtain the technologies and if successful, use the same paths (→) to egress. The side of the base with the laboratories is also patrolled by military police in armored vehicles (×). Finally, scattered throughout the base are government civilian workers (+) that are potentially in harms way.

As noted, the base security forces are unaware of the impending attack. Despite being in the position of responding as events unfold, the base security forces can develop a desired end-state to focus actions. The end-state would be a function of what is valued as explained under the deterministic framework. For this scenario, the base security forces' notional desired end-state is illustrated in Figure 20. Specifically, the three primary objectives in the notional end-state are to secure the base, remain fully mission capable, and secure advanced technologies. Additionally, the 'secure the base' objective can be further broken down into the sub-objectives: all base sectors searched, all units/individuals identified, and all discovered terrorists captured/killed. In a similar manner, the 'remain fully mission capable' objective is composed of the sub-objectives: no personnel losses (both military and civilian), no equipment losses, and no infrastructure losses. Finally, the 'secure advanced technologies' objective, is made up of the sub-objectives: all base sectors searched for technologies and all stolen technologies recovered/destroyed. To complete the end-state characterization requires quantifying priorities among the objectives. This is done by assigning weights to the objectives. For this notional scenario, the weightings depicted in Figure 20 are used.

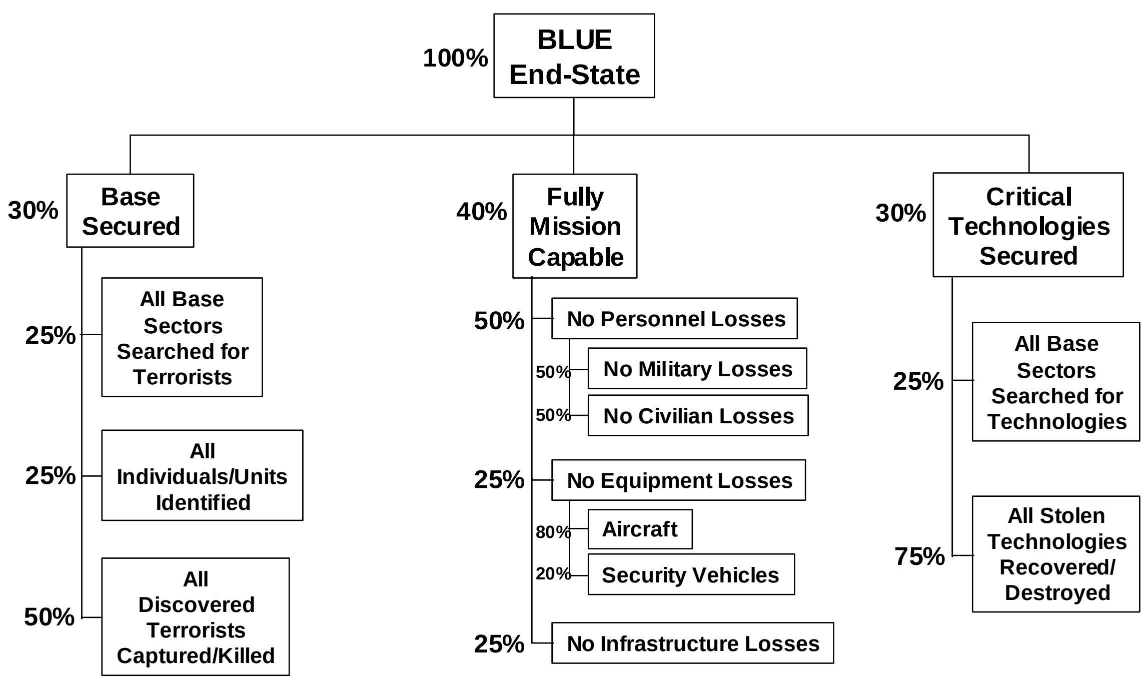


Figure 20. Base Security Forces End-State Characterization

As noted under the deterministic framework, a key aspect of the proposed methodology is being able to quantify abstract concepts. This is accomplished by identifying what is important and continuing to ask 'why it is important' until the concept cannot be further refined. This reductionist approach assisted in yielding the end-state characterization shown in Figure 20. However, in addition to breaking down abstract concepts, this methodology also simplifies the task of identifying attributes and their measures. Refinement of concepts to this fundamental level often yields natural and direct measures (Sink, 1985:86). This outcome can be seen in Table 5. Finally, on an Intel® Pentium 4® 3GHz based computer with 1GB of RAM, the scenario requires approximately 2½ hours to reach completion, which occurs when either RED escapes or is captured/killed. The significant scenario events occurring over the twenty-five minutes of simulated time are outlined in Table 6. The resulting twenty-five minutes generated the observations, at one minute intervals, shown in Table 7.

Table 5. Attributes and Measures Characterizing BLUE End-State

	Objective	Value	Attribute	Measure
BLUE End-State	Base Secured	All base sectors searched for terrorists	Sectors searched	Number of sectors searched out of 11 total.
		All individuals/units identified	Individuals/units identified	Number of individuals/units positively identified. The total number changes as the scenario progresses, but begins with the total military and civilian base population (30 + 195).
		All discovered terrorists captured/killed	Terrorists captured/killed	Number of positively identified terrorists captured/killed. The total number changes as the scenario progresses but begins at 0. Note: For this scenario, the number could be as high as 14 (9 terrorists + 3 car bombers + 2 truck bombers).
	Fully Mission Capable	No personnel losses	Military losses	Number of military losses out of a potential total of 30.
			Civilian losses	Number of civilian losses out of a potential total of 195.
		No equipment losses	Aircraft losses	Number of aircraft losses out of a potential total of 10 (4 bombers + 4 fighters).
			Security vehicle losses	Number of security vehicles losses out of a potential total of 10.
		No infrastructure losses	Infrastructure losses	Number of infrastructure losses out of a potential total of 5 (3 laboratories + 1 fuel depot + 1 command post).
	Critical Technologies Secured	All base sectors searched for technologies	Sectors searched	Number of sectors searched out of 11 total.
		All stolen technologies recovered/destroyed	Technologies recovered/destroyed	Number of technologies recovered/destroyed. The number changes as the scenario progresses but starts at 0 and can be as high as 3.

Table 7 represents the raw observations for each of the system attributes of interest. However, an inescapable feature of measurement is error and uncertainty

(Mitchell, 2003:301; Finkelstein, 2003:45). If the constraints of normality and linearity across errors are assumed, when combined with domain knowledge, a Kalman filter can be used to mitigate the impact of this error and uncertainty. The raw observations in Table 7 transformed through use of a 2nd order Kalman filter (T_s = 1), along with domain knowledge about the environment, are shown in Table 8.

One of the underlying themes of this research is that effectiveness is a relative concept. Thus, in order to measure effectiveness, a reference point is required. Under the framework being presented, the reference point is associated with the desired-end state. The reference point for each of the attributes of interest in this scenario is shown along with the filtered observations in Table 9.

Table 6. Scenario Significant Events

Time Period	Significant Events
1	- Terrorists commence with attack plan
4	- Base security forces encounter terrorists - Base security forces begin search for terrorists - Base security forces start friend/foe identification
8	- Base security forces kill a truck bomber - Base security forces kill a car bomber
9	- Base security forces kill a car bomber
10	- Base security forces kill a truck bomber
11	- Base security forces kill a car bomber
12	- Base security forces first encounter terrorist squads in off-road vehicles
13	- Base security forces kill 1 terrorist squad in an off-road vehicle
15	- Base security forces sustain first losses - Terrorists steal critical technology from a laboratory
16	- One base security vehicle destroyed by terrorists - Base security forces kill 1 terrorist squad in an off-road vehicle
17	- Base security forces complete search of all sectors for terrorists - All friendly forces accounted for - Terrorists destroy 1 laboratory - First civilian losses
19	- Base security forces complete search for critical technlogies
24	- Base security forces recover stolen critical technology
25	- All identified terrorists killed

Continuing, this research generically defined effectiveness as an attribute distance change relative to the desired end-state for the attribute. These distances are shown in

Table 10. While these distances allow for comparison across time periods for a given system attribute, more meaningful insight on progress towards the desired end-state is provided by comparing across system attributes. To obtain this type of insight requires normalizing the attribute observations. Although numerous normalization techniques exist, many do not preserve the scale meaning of the original observation (Kirkwood, 1997:241; Zuse, 1998:232). One technique that *does* preserve the scale meaning of the original observations is outlined in Figure 11. The algorithm in Figure 11 was used to transform the distances in Table 10 to the normalized distances shown in Table 11. Finally, the normalized distances can be combined to provide a single system effectiveness measure by multiplying the normalized attribute distances by the associated attribute weighting (Figure 20), which yields the results in Table 12.

These steps complete implementation of the frameworks developed in this research. However, display of the resulting information is also important. Although not a focus of this research effort, visualization of quantitative data is crucial in supporting decision-making based on effectiveness measures (Tufte, 1997:9). For the scenario results, three possible alternatives to visualize the data in Table 12 are shown in Figure 21, Figure 22, and the twenty-five figures of Appendix D.

These types of visualization techniques more clearly and readily communicate important system changes to the decision maker. For example, the bar charts in Appendix D provide time independent views of the system (scenario) at one minute intervals for the duration of the scenario. The charts not only highlight the significant events as delineated in Table 6, but more importantly portray the *effect* of those events relative to the desired end-state. Further, because of the mathematical concepts built into the proposed theory, the magnitude of the effect, or the *effectiveness*, can be assessed.

When the information contained in the bar charts is consolidated into a single view, additional insights can be gleaned as illustrated in Figure 21. The consolidation removes the time independence constraint and provides the decision maker a historical perspective on the effect and effectiveness of system (scenario) events over time. Further, Figure 21 not only indicates how individual system attributes are changing over time, but the last row in Figure 21 incorporates the attribute priorities, identified in Figure 20, to provide an overall system effectiveness assessment.

Another approach to viewing the information in Figure 21 is shown in Figure 22. The line chart view of Figure 22 also shows the overall system effectiveness assessment, but instead of going down to the attribute level, Figure 22 portrays the data only down to the primary objective level. Collectively, these three alternative views illustrate how the decision maker can control the granularity of the effectiveness measurements to best support decision making.

The overall effectiveness measurement process used for this notional scenario is illustrated in Figure 23. Comparing Figure 23 to Figure 1 highlights how the elements developed in this research build upon the established, basic measurement concepts. The overall process starts with the product structure, or measurement model, presented earlier. Embedded within the product structure process is the development of measures. If the developed measures hold for the metric properties, then the theoretical assertions

Table 7. Scenario Observations

Attributes	Time Periods																								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**
Sectors Searched for Terrorists	0	0	0	0	4	4	4	4	5	6	6	7	8	9	10	10	11	11	11	11	11	11	11	11	11
Individuals Identified (Red/Blue)	0 – 0	0 – 0	0 – 0	0 – 0	3 – 107	3 – 107	4 – 107	4 – 108	5 – 128	5 – 149	6 – 149	7 – 168	8 – 181	9 – 200	10 – 219	11 – 219	12 – 227	12 – 227	13 – 227	13 – 227	13 – 227	13 – 227	13 – 227	13 – 227	14 – 227
Terrorists Captured / Killed	0	0	0	0	0	0	0	2	3	4	5	5	8	8	8	11	11	11	11	11	11	11	11	13	14
Military Losses	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	3	3	3	3	3	3	3	3
Civilian Losses	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	20	20	20	20	20	20	20	20
Aircraft Losses	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Security Vehicle Losses	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
Infrastructure Losses	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
Sectors Searched for Technologies	0	0	0	0	0	0	0	0	0	0	0	4	5	6	7	8	9	10	11	11	11	11	11	11	11
Technologies Recovered / Destroyed	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1

Table 8. Kalman Filtered Observations

Attributes	Time Periods																								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**
Sectors Searched for Terrorists	0	0	0	0	4	5	5	5	5	6	6	7	8	9	10	10	11	11	11	11	11	11	11	11	11
Individuals Identified (Red/Blue)	0–0	0–0	0–0	0–0	3–95	3–122	4–127	5–126	5–134	5–149	6–155	7–166	8–178	9–193	10–210	11–220	12–225	13–225	13–225	14–225	14–225	14–225	14–225	14–225	14–225
Terrorists Captured / Killed	0	0	0	0	0	0	0	1	3	4	5	6	8	9	9	11	12	12	13	13	13	13	13	13	14
Military Losses	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	3	4	4	4	4	4	4
Civilian Losses	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	14	19	21	23	25	26	26	26
Aircraft Losses	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Security Vehicle Losses	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
Infrastructure Losses	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
Sectors Searched for Technologies	0	0	0	0	0	0	0	0	0	0	0	2	4	6	7	8	10	11	11	11	11	11	11	11	11
Technologies Recovered / Destroyed	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Table 9. Filtered Observations with Reference

Attributes	Time Periods																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Sectors Searched for Terrorists	0 – 0	0 – 0	0 – 0	0 – 0	4 – 11	5 – 11	5 – 11	5 – 11	5 – 11	6 – 11	6 – 11	7 – 11	8 – 11	9 – 11	10 – 11	10 – 11	11 – 11	11 – 11	11 – 11	11 – 11	11 – 11	11 – 11	11 – 11	11 – 11	11 – 11
Individuals / Units Identified	0 – 0	0 – 0	0 – 0	0 – 0	98 – 228	125 – 228	131 – 229	131 – 230	139 – 230	154 – 230	161 – 231	173 – 232	186 – 233	202 – 234	220 – 235	231 – 236	237 – 237	238 – 238	238 – 238	239 – 239	239 – 239	239 – 239	239 – 239	239 – 239	239 – 239
Terrorists Captured / Killed	0 – 0	0 – 0	0 – 0	0 – 0	0 – 3	0 – 3	0 – 4	1 – 5	3 – 5	4 – 5	5 – 6	6 – 7	8 – 8	9 – 9	9 – 10	11 – 11	12 – 12	12 – 13	13 – 13	13 – 14	13 – 14	13 – 14	13 – 14	14 – 14	14 – 14
Military Losses	0 – 0	0 – 0	0 – 0	0 – 0	0 – 30	0 – 30	0 – 30	0 – 30	0 – 30	0 – 30	0 – 30	0 – 30	0 – 30	0 – 30	0 – 30	1 – 30	2 – 30	3 – 30	3 – 30	4 – 30	4 – 30	4 – 30	4 – 30	4 – 30	4 – 30
Civilian Losses	0 – 0	0 – 0	0 – 0	0 – 0	0 – 195	0 – 195	0 – 195	0 – 195	0 – 195	0 – 195	0 – 195	0 – 195	0 – 195	0 – 195	0 – 195	0 – 195	8 – 195	14 – 195	19 – 195	21 – 195	23 – 195	25 – 195	26 – 195	26 – 195	26 – 195
Aircraft Losses	0 – 0	0 – 0	0 – 0	0 – 0	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10
Security Vehicle Losses	0 – 0	0 – 0	0 – 0	0 – 0	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	0 – 10	1 – 10	1 – 10	1 – 10	1 – 10	1 – 10	1 – 10	1 – 10	1 – 10	1 – 10
Infrastructure Losses	0 – 0	0 – 0	0 – 0	0 – 0	0 – 5	0 – 5	0 – 5	0 – 5	0 – 5	0 – 5	0 – 5	0 – 5	0 – 5	0 – 5	0 – 5	0 – 5	0 – 5	1 – 5	1 – 5	1 – 5	1 – 5	1 – 5	1 – 5	1 – 5	1 – 5
Sectors Searched for Technologies	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	2 – 11	4 – 11	6 – 11	7 – 11	8 – 11	10 – 11	11 – 11	11 – 11	11 – 11	11 – 11	11 – 11	11 – 11	11 – 11	11 – 11
Technologies Recovered / Destroyed	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 0	0 – 1	0 – 1	0 – 1	0 – 1	0 – 1	0 – 1	0 – 1	0 – 1	0 – 1	0 – 1	1 – 1

Table 10. Distance from End-State

Attributes	Time Periods																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Sectors Searched for Terrorists	0	0	0	0	7	6	6	6	6	5	5	4	3	2	1	1	0	0	0	0	0	0	0	0	0
Individuals / Units Identified	0	0	0	0	130	103	98	99	91	76	70	59	47	32	15	5	0	0	0	0	0	0	0	0	0
Terrorists Captured / Killed	0	0	0	0	3	3	4	4	2	1	1	1	0	0	1	0	0	1	0	1	1	1	1	1	0
Military Losses	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	3	3	3	3	3	3	3
Civilian Losses	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	14	19	21	23	25	26	26	26
Aircraft Losses	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Security Vehicle Losses	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
Infrastructure Losses	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
Sectors Searched for Technologies	0	0	0	0	0	0	0	0	0	0	0	9	7	5	4	3	1	0	0	0	0	0	0	0	0
Technologies Recovered / Destroyed	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0

Table 11. Normalized Distance from End-State

Attributes	Time Periods																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Sectors Searched for Terrorists	0.00	0.00	0.00	0.00	0.68	0.58	0.57	0.58	0.53	0.46	0.43	0.37	0.30	0.21	0.12	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Individuals / Units Identified	0.00	0.00	0.00	0.00	0.57	0.45	0.43	0.43	0.39	0.33	0.30	0.25	0.20	0.14	0.06	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Terrorists Captured / Killed	0.00	0.00	0.00	0.00	1.00	1.00	1.00	0.80	0.40	0.20	0.17	0.14	0.00	0.00	0.10	0.00	0.00	0.08	0.00	0.07	0.07	0.07	0.07	0.07	0.00
Military Losses	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.07	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
Civilian Losses	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.07	0.10	0.11	0.12	0.13	0.13	0.13	0.13
Aircraft Losses	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Security Vehicle Losses	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
Infrastructure Losses	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
Sectors Searched for Technologies	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.82	0.64	0.45	0.36	0.27	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Technologies Recovered / Destroyed	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00

Table 12. Weighted Normalized Distance from End-State

Weights					Attributes	Time Periods 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
100%	30%	25%		7.5%	Sectors Searched for Terrorists	0.000	0.000	0.000	0.000	0.051	0.044	0.043	0.043	0.040	0.035	0.032	0.028	0.022	0.016	0.009	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
		25%		7.5%	Individuals / Units Identified	0.000	0.000	0.000	0.000	0.043	0.034	0.032	0.032	0.030	0.025	0.023	0.019	0.015	0.010	0.005	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
		50%		15.0%	Terrorists Captured / Killed	0.000	0.000	0.000	0.000	0.150	0.150	0.150	0.120	0.060	0.030	0.025	0.021	0.000	0.000	0.015	0.000	0.000	0.012	0.000	0.011	0.011	0.011	0.011	0.011	0.000
	40%	50%	50%	10.0%	Military Losses	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.007	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010
			50%	10.0%	Civilian Losses	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.007	0.010	0.011	0.012	0.013	0.013	0.013	0.013
		25%	80%	8.0%	Aircraft Losses	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
			20%	2.0%	Security Vehicle Losses	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
		25%		10.0%	Infrastructure Losses	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.020	0.020	0.020	0.020	0.020	0.020	0.020	0.020
	30%	25%		7.5%	Sectors Searched for Technologies	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.061	0.048	0.034	0.027	0.020	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
		75%		22.5%	Technologies Recovered / Destroyed	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.225	0.225	0.225	0.225	0.225	0.225	0.225	0.225	0.225	0.225	0.000

Attributes	Time Periods 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Sectors Searched for Terrorists	100%	100%	100%	100%	32%	42%	43%	42%	47%	54%	57%	63%	70%	79%	88%	93%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Individuals / Units Identified	100%	100%	100%	100%	43%	55%	57%	57%	61%	67%	70%	75%	80%	86%	94%	98%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Terrorists Captured / Killed	100%	100%	100%	100%	0%	0%	0%	20%	60%	80%	83%	86%	100%	100%	90%	100%	100%	92%	100%	93%	93%	93%	93%	93%	100%
No Military Losses	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	97%	93%	90%	90%	90%	90%	90%	90%	90%	90%
No Civilian Losses	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	96%	93%	90%	89%	88%	87%	87%	87%	87%
No Aircraft Losses	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
No Security Vehicle Losses	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	90%	90%	90%	90%	90%	90%	90%	90%	90%
No Infrastructure Losses	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	80%	80%	80%	80%	80%	80%	80%	80%
Sectors Searched for Technologies	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	18%	36%	55%	64%	73%	91%	100%	100%	100%	100%	100%	100%	100%	100%
Technologies Recovered / Destroyed	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%
Overall	100%	100%	100%	100%	76%	77%	78%	80%	87%	91%	92%	87%	91%	94%	72%	74%	76%	72%	73%	72%	72%	72%	72%	72%	95%

0% - 25% | 25% - 60% | 61% - 94% | 95% - 100%

Figure 21. Table Visualization of Time Series Results

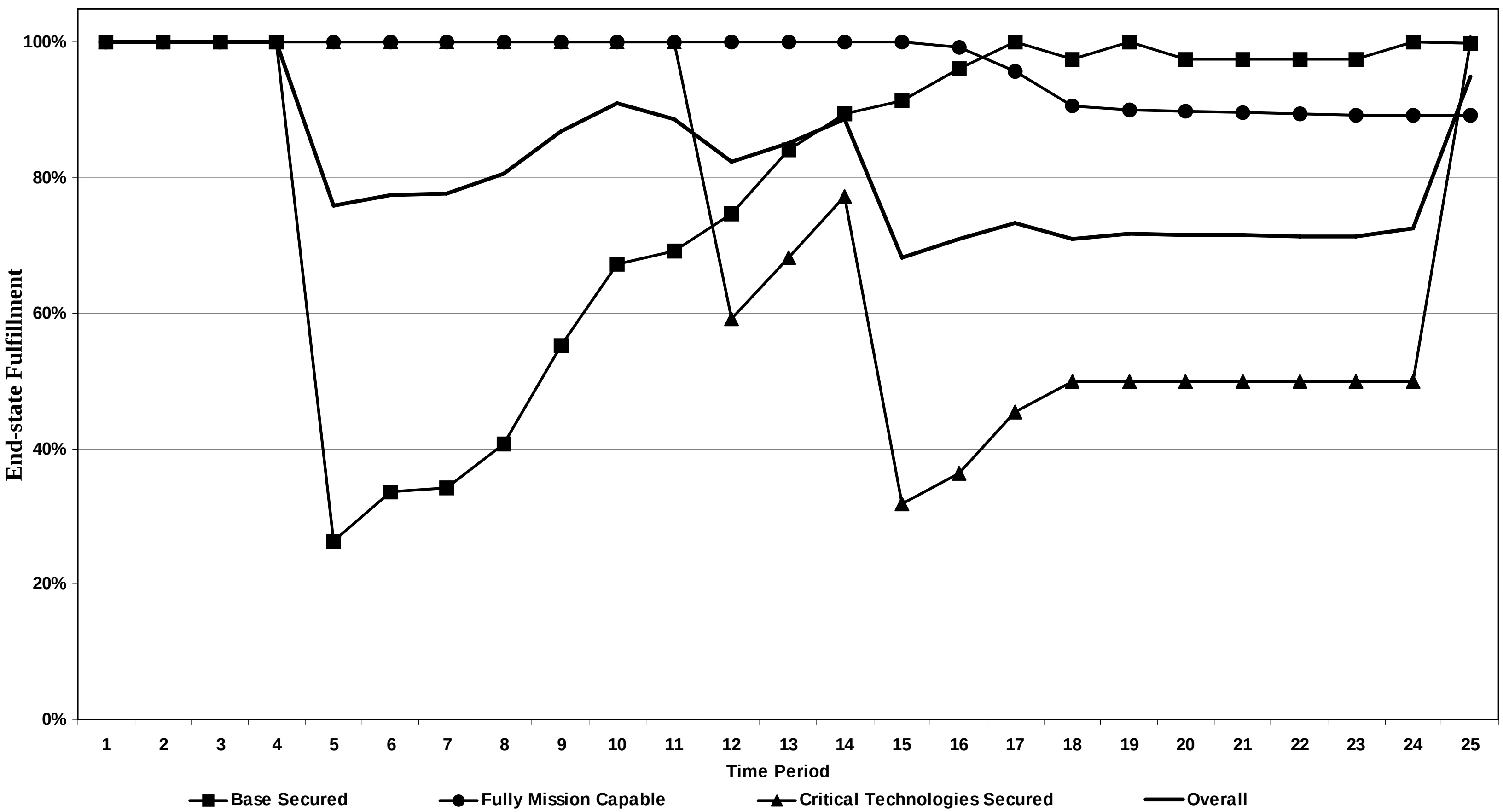


Figure 22. Line Chart Visualization of Results

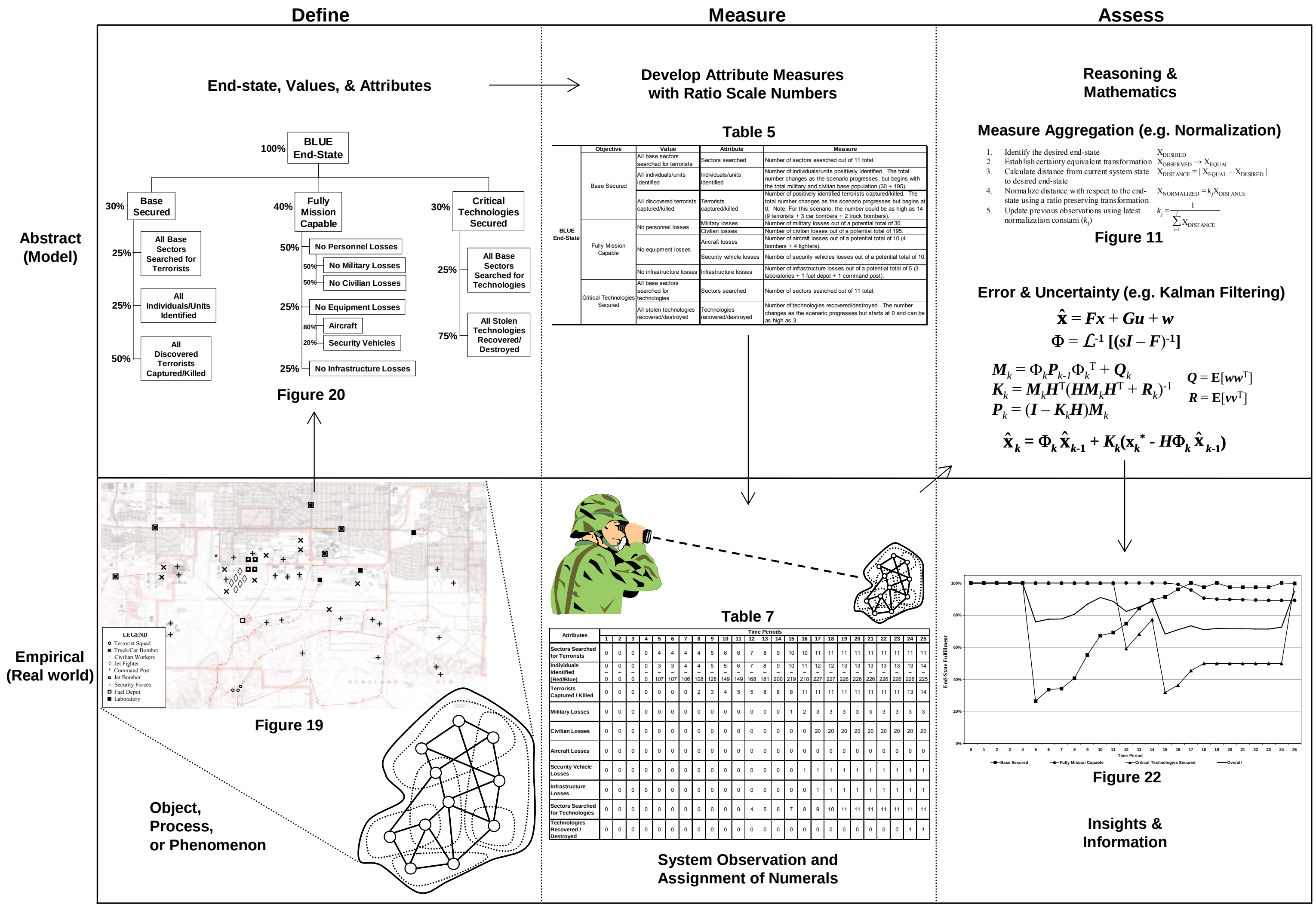


Figure 23. Effectiveness Measurement Process

presented in this research also hold. Additionally, after conducting system observations, mathematical techniques, such as that in Figure 11, can be used to assist in aggregation of low level data. Further, probabilistic reasoning techniques, like the Kalman Filter presented in this research, can be used to address the error and uncertainty associated with effectiveness measurement. As displayed in Figure 23, the collective measurements can then be used to provide insights about the system of interest and specifically, how the system is progressing towards the desired end-state.

To stress a final point, in the scenario used to demonstrate the frameworks proposed in this research, a single course-of-action was used. Specifically, pre-positioned base security forces patrolled pre-defined areas of responsibility until terrorists were encountered. After encountering terrorists, the base security forces engaged the terrorists and pursued them even if pursuit took the base security forces beyond their pre-defined area of responsibility. An additional element of the course-of-action was to maximize base security force coverage of the entire base (highlighted area in Figure 19). This resulted in some base security forces leaving their pre-defined area of responsibility to provide support even if terrorists were not encountered. This course-of-action used is notional and is clearly one of many that could have been used to respond to the terrorist attack.

The proposed frameworks as presented were intended for use in assessing the effectiveness of a single course-of-action that was being executed. A natural extension in the use of the frameworks, however, is to determine which course-of-action to use among a number of developed courses-of-action. The proposed frameworks provide a foundation for developing a common basis for comparison on not only course-of-action fulfillment of the desired end-state but on timeliness of fulfillment as well. To further assist in development of courses-of-action for achievement of a desired end-state, the proposed frameworks could be combined with various Operations Research techniques such as Response Surface Methodology, to identify common strengths and weaknesses among courses-of-action being evaluated, or Linear/Goal Programming, to optimize timing and sequencing of action within a selected course-of-action.

V. CONCLUSIONS

CONTRIBUTIONS & RECOMMENDATIONS

This dissertation has synthesized elements from the broad field of measurement, reviewed and identified limitations of various measurement approaches, and introduced a theoretical foundation, as well as a corresponding framework for effectiveness measurement. While the primary motivation for this research was measurement of military campaign advancement, effectiveness measurement is of broad interest and applicable to many fields of endeavor. The methods developed in this research address the need for a rigorous, mathematically grounded basis for monitoring progress towards abstract goals and objectives.

This research began by exploring fundamental issues related to measurement as well as foundational concepts established via Measurement Theory. These theoretical topics were then balanced by an examination of various views on the application of measurement. Next, attention focused on the key driver of this research, Effects-based Operations. Despite the literature in the field of Effects-based Operations being highly disjoint, commonalities were identified to establish a foundation for effects concepts. Finally, building upon the measurement, Measurement Theory, and Effects-based Operations ideas, Measure Theory concepts were introduced, which provided the mathematical means for real-world system modeling. Culmination of these concepts resulted in an axiomatic-based Theory of Effectiveness Measurement.

To meet the practical needs of measurement application, a probabilistic framework to address error and uncertainty associated with effectiveness measurement was also introduced. Numerous techniques exist for probabilistic reasoning. While each technique has advantages and disadvantages, all are suitable to handle the error and uncertainty encountered in effectiveness measurement. However, one established approach, Kalman Filtering, stood out as being best suited for mitigating these probabilistic problems, as well as being an excellent match for integration with the axiomatic-based Theory of Effectiveness Measurement developed in this research. As a final means of making this introduced mathematical construct pragmatic, mechanical details on the implementation of the Theory of Effectiveness Measurement were demonstrated using a notional scenario.

While this research introduced a new, comprehensive theory, there are a number of areas for further research. First, this research assumed a course-of-action had been developed and was being executed. The developed effectiveness measurement framework then provides feedback to determine the effectiveness of the course-of-action. A key step of an effects-based approach, however, is planning, or determining the best course-of-action from a number of developed courses-of-action (USJFC, 2006:viii). During planning, the developed effectiveness measurement framework could provide a common basis for comparison of candidate courses-of-action. In a similar vein, during the planning process, the developed effectiveness measurement framework could be combined with Operations Research techniques such as Linear and Goal Programming to optimize the sequencing and timing of a selected best course-of-action.

Another assumption of this research is that observations are based strictly on outward behavioral system attributes. This passive approach will always have more error and uncertainty associated with it since the time between the observed behavior and the action that produced the observed behavior will rarely be instantaneous. If the effectiveness measurement framework developed in this research could be linked with internal models of the system of interest, not only could the error and uncertainty be significantly mitigated, but system state changes from a given action could be forecasted.

The goal of this research was to provide a framework for effectiveness measurement from both a theoretical and practical view. An axiomatic-based measurement theory was presented and a generic measurement methodology explored. The most important contribution of this effort is a theory for effectiveness measurement; however, there are empirical benefits as well. The intent was to develop fundamental effectiveness measurement principles and to give theoretical, as well as practical guidelines for implementation of effectiveness measurement.

The theory provides a standardized framework for thinking about effects regardless of the domain. The framework includes precise definitions of the qualitative concepts within the frameworks, along with their corresponding quantitative notation. Additionally, there is a mechanism for interpreting numbers, criteria for selecting measures, conditions for comparing measures, theoretical foundations for validating measures, as well as approaches for handling uncertainty.

From an academic standpoint, the most significant contribution is the 'theory', however, from a practical standpoint, the most important contribution, is meeting the needs of the practicing analyst *with* the proper theory. In summary, a theoretically-based effectiveness measurement approach provides effects assessment practitioners a level of precision on par with the level of precision with which Effects-based Operations are conducted.

Theory Of Effectiveness Measurement

Appendix A: Game Theory

Game Theory addresses decision contexts where there are two or more decision makers with competing or conflicting objectives and the outcome for each decision maker depends on the choices made by the others (Winston, 1994:824). Additionally, each decision maker knows their outcome is influenced by the choices of other decision makers which in turn, influences their preferences. Essentially, Game Theory assists a decision maker in arriving at a better decision. In Game Theory, the decision depends on the choices available to the decision maker and the decision maker's preferences on the outcomes of each of those alternatives. Additionally, the decision maker's beliefs about what actions are available to each of the other decision makers, beliefs about how each of those decision makers rank the outcomes of their choices, and beliefs about every other decision maker's beliefs (Luce, 1957:5) also influence the decision at hand.

Game Theory provides a framework for thinking about strategic interaction and helps formulate an optimal strategy by forecasting the outcome of strategic situations (Dresher, 1961:1). Thus, Game Theory concerns games of strategy versus games of pure chance such as slot machines or non-interactive games like solitaire. The idea of a general theory of games was introduced by John von Neumann and Oskar Morgenstern in 1944, in their book, <u>Theory of Games and Economic Behavior</u>. They describe a game as a competitive situation among two or more decision makers, or groups with a common objective, conducted under a prescribed set of rules and known outcomes (von Neumann, 1944:49). The objective of Game Theory is to determine the best strategy for a given decision maker under the assumption the other decision makers are rational, or consistently make decisions in alignment with some well-defined objective, and will make intelligent countermoves, where intelligent implies all decision makers have the same information and are capable of inferring the same insights from that information (von Neumann, 1944:51).

A cornerstone concept of Game Theory is each decision maker will act to maximize their expected outcome. For example, possible outcomes are generally characterized by a numeric representation or on a utility scale. It is assumed, given a set of possible choices and associated outcomes, for any two alternatives, decision makers can discern preference or indifference among the alternatives, allowing them to rank the set of alternatives with respect to each other. A key result from von Neumann and Morgenstern is there exists a way of assigning utility numbers to the outcomes such that the decision maker would always choose the option that maximizes their expected utility (Luce, 1957:4). Thus, while Decision Theory assumes decision makers are self-interested and selfish, Game Theory extends this to assume everyone else is too.

Games can be characterized in a wide variety of ways. Some of the attributes that can be used to classify games include players, structure, outcome, interaction, timing, and information. Game Theory typically addresses contexts with *n*-decision makers, where *n* is two or more. However, some sources address 1-person games as games against 'nature', which is the realm of Decision Theory.

Games can be characterized by three basic structures: Simultaneous, Sequential, and Repeated. In Simultaneous games (also known as Static or Stage games), all decision makers reveal their decision to other decision makers simultaneously (Myerson, 1991:47). Thus, Simultaneous games amount to trying to forecast the decisions of other decision makers. Even in situations where decisions are not made simultaneously, any game where decisions are made without knowledge of other decision maker's choices are considered Simultaneous games.

Sequential games (also known as Dynamic or Multi-stage games) require a sequence of decisions for preferences over a set of alternatives where a different set of alternatives is presented at each decision point; usually dependent on decisions made in previous stages. Thus, ordering of decisions is important. Finally, Repeated games, like sequential games, require a sequence of decisions, but each decision point is similar to a simultaneous game where choices available and their outcomes may be dependent on previous choices. In contrast to a 'one-shot' Simultaneous game, in Repeated games, all past decisions for previous decision points are known to all other decision makers (Fudenberg, 1993:107).

Another attribute of games concerns outcome. A constant-sum game is where the sum of the outcomes for all decision makers is constant (Winston, 1994:827). A common instantiation is where the constant is zero. Thus, a zero-sum game is one where the decision makers' interests are in direct conflict and what one decision maker 'loses', another decision maker 'wins'. A constant-sum game is in contrast to a variable sum game, or general-sum game, where the sum of the outcomes for all decision makers is not constant (Owen, 1968:136).

Interaction is another way to categorize games. Interaction addresses the level of cooperation among decision makers. In a non-cooperative game, each decision maker pursues their own interests (Luce, 1957:89). In cooperative games, however, decision makers are free to form coalitions and make agreements, essentially combining their decision making problems (Myerson, 1991:244).

Table 13. Attributes of Games

Attribute	Game Type
Players	1-player n-players
Structure	Simultaneous (Static or One-stage) Sequential (Dynamic or Multi-stage) Repeated
Outcome	Constant-sum Variable-sum
Interaction	Cooperative Non-cooperative
Time	Duel Non-duel Differential
Information	Perfect Incomplete Imperfect

Yet another way to classify Game Theory problems is by time. In general, Game Theory does not put a time constraint on the decision maker to make a decision at a decision point. However, if the passage of time does impact the expected outcome, the game is referred to as a Duel (Dresher, 1961:128). Another type of game where time is a factor is a differential game. Differential games address multi-decision maker problems in dynamic situations where the position, or state, of the players develops continuously in time (Friedman, 1971:19).

A final way to characterize games is by information. A game of perfect information is one where each decision maker has the same information. This includes information on all previous decisions for sequential games (Shubik, 1982:232). If the game does not allow for perfect information, it is termed a game of imperfect information or a Bayesian game (Fudenberg, 1993:209). Finally, in a game of incomplete information, only some elements of information are unknown. Table 13 summarizes these various game attributes.

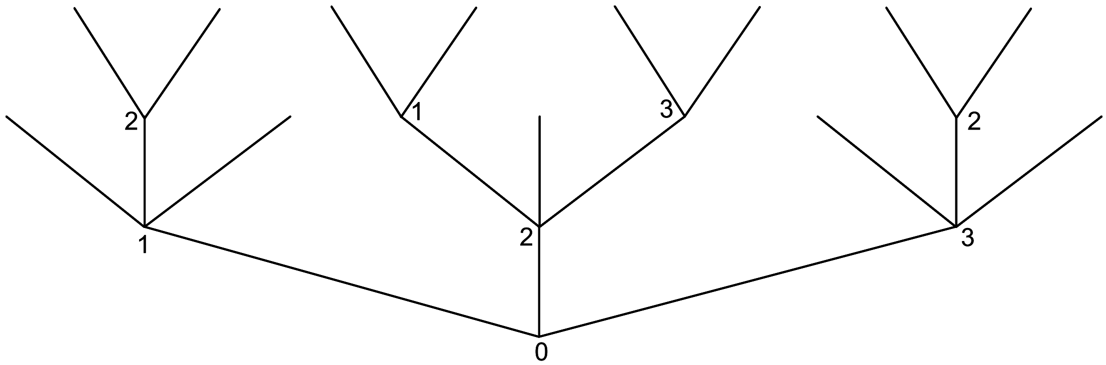


Figure 24. Extensive Form (Game Tree)

Regardless of game type, a conceptual model is needed in order to analyze a game. There are three primary models, or forms, that can be used: extensive form, coalitional form, and strategic form. Games in extensive form are usually depicted as a multi-player decision tree (Luce, 1957:40) as shown in Figure 24. The extensive form describes sequentially what each decision maker might do and the possible outcomes. For example, in Figure 24, the number at each node represents the player making the decision. By convention, node '0' is a chance node. In the literature, the lesser discussed of the three forms is the coalitional form, which is focused on examining the value of belonging to a coalition (Shubik, 1983:4).

Games in strategic form, or normal form, are the most common for examining games. In strategic form, in contrast to the extensive form, details of the game, such as position and move, are not shown. The key aspects available in the strategic form are the decision makers, their strategies, and the possible outcomes (Luce, 1957:53). All the forms generally assume the number of decision makers is finite. Additionally, all the forms usually assume the number of strategies available to each decision maker is also finite. However, extensions with the strategic form can accommodate a decision maker with infinite strategies (Fudenberg, 1993:5). Mathematically, the simplest way to describe a game is the strategic form. Thus, the strategic form will be the focus in the remainder of this review.

A specific strategic form game can be formalized by three elements: the set of decision makers or players $i \in \boldsymbol{I}$, which is assumed to be finite (i.e. $\boldsymbol{I} = \{ 1, 2,\ldots, n \}$); the set of alternatives available to each of the decision makers, or the pure strategy space $\boldsymbol{S}_i = \{ s_{i1},\ldots, s_{im} \}$ $m < \infty$; and the outcome functions $u_i : \boldsymbol{S}_i \rightarrow \mathbb{R}$ giving player i's von Neumann-Morgenstern utility $u_i(s)$ where $s = (s_{1w}, s_{2x},\ldots, s_{iy},\ldots, s_{nz})$ is the strategy profile, or specific instance of choices made by all decision makers (Fudenberg, 1993:4).

Many n-person decision contexts with competing decision makers have multiple, conflicting objectives. Typically, however, the objectives will have a common ordering among all the decision makers which has the effect of polarizing the decision context and essentially making it a 2-person game (Isaacs, 1965:306). Examples of this phenomenon are numerous. A historical example is the unlikely alliance of the USA, Soviet Union, and China during World War II. Although these 'players' had fundamental differences at the time, an important, shared objective brought them together, polarizing the situation. Additionally, simplifying an n-person ($n > 2$) game to a 2-peron game, allows a characterization of the strategic form to be displayed as 'game matrix' (Fudenberg, 1993:5).

The following Gridiron game (Table 14) is used to demonstrate additional properties of games. The Gridiron game is a non-cooperative, 2-person, simultaneous, zero-sum, non-duel game with perfect information. The two players are Offense and

Defense. Offense has four alternatives, or pure strategies, to choose from and Defense has three. A pure strategy is a predetermined sequence of moves and countermoves made during the game (Kaplan, 1982:105). In this game, Offense knows its pure strategies, the pure strategies of Defense, and the outcome when one pure strategy is played against another. Defense has the same information. When all players know the same fact, it is called mutual knowledge. Further, Offense knows Defense knows what it knows and Defense knows Offense knows what it knows. When all players know a fact and all know that all know it, it is called common knowledge (Fudenberg, 1993:541).

Table 14. Gridiron Game

		Defense		
		Run	**Pass**	**Blitz**
Offense	**Run**	-3	5	5
	Short Pass	3	0	3
	Medium Pass	7	0	6
	Long Pass	10	0	-10

In yards gained by Offense

If some pure strategy is strictly preferred over another strategy *s*, regardless of what other players do, *s* is a dominating strategy (Luce, 1957:79). For Offense, 'Medium Pass' is always preferred over 'Short Pass'. Thus, 'Short Pass' is dominated. If there were a strategy, *s*, that dominated all other strategies regardless of what other players were doing, *s* would be a dominant strategy (Owen, 1968:25). Neither Offense nor Defense has a dominant strategy. If both had dominant pure strategies, their intersection would be the classic saddle point (von Neumann, 1944:95).

If no players have a dominant solution, they must select the 'best' strategy based on what they know (and what they think all the other players know). However, if one player always chooses the same pure strategy or chooses pure strategies in a fixed order, opponents in time will recognize the pattern and exploit the information to defeat the player. Thus, when no dominant pure strategy exists, the most effective strategy is a mixed strategy (Owen, 1968:16). A mixed strategy is defined by a probability distribution over the set of pure strategies. Under a mixed strategy, each player will form a probabilistic assessment over what other players will do. Thus, when a player chooses one of their own strategies, they are choosing a lottery over other player's mixed strategy profiles. Further, a player can interpret other player's mixed strategies as expectations of how they are likely to play (Luce, 1957:74).

The notation presented thus far can be extended as follows: A mixed strategy σ_i is a probability distribution over the pure strategies and $\boldsymbol{\sigma}$ is the space of mixed strategy profiles. σ_i (s_i) is the probability that σ_i assigns to s_i where $\sum_i \sigma_i (s_i) = 1$ and $\sigma_i (s_i) \geq 0$. Additionally, $u_i(\sigma_i)$ = player i's outcome under the mixed strategy (Fudenberg, 1993:5).

Assumed in Game Theory is that players will select the strategy that maximizes their outcome given the other players' strategies. If every player is playing their best strategy, there is no incentive for any player to unilaterally change their strategy and thus, the players are at strategic equilibrium or Nash equilibrium (Fudenberg, 1993:11). An important result for mixed strategies is *every* finite game has at least one strategic equilibrium, which was proved by John Nash in his dissertation, Non-cooperative games (1950). This equilibrium, or optimum set of strategies, can be found using the Minimax Theorem. The key result of this theorem is when one player attempts to minimize their opponent's maximum outcome, while their opponent attempts the contrary; the result is the minimum of the maximum outcomes equals the maximum of the minimum outcomes (von Neumann, 1944:93).

A solution is a description of an outcome that may emerge from a game. The optimal solution guaranteed by the Minimax Theorem can be solved via linear programming for 2-player, constant-sum games in a program of the form shown in Figure 25, which yields the optimal strategy for the column player. The row player strategy can be found via the dual solution (Winston, 1994:840).

$$
\begin{array}{ll}
\text{maximize:} & z = -y_{n+1} \\
\text{subject to:} & u_{11}y_1 + u_{12}y_2 + \ldots + u_{1n}y_n - y_{n+1} \leq 0 \\
 & u_{21}y_1 + u_{22}y_2 + \ldots + u_{2n}y_n - y_{n+1} \leq 0 \\
 & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
 & u_{m1}y_1 + u_{m2}y_2 + \ldots + u_{mn}y_n - y_{n+1} \leq 0 \\
 & y_1 + y_2 + \ldots + y_n = 1 \\
 & y_1, y_2, \ldots, y_n > 0
\end{array}
$$

Figure 25. Program for Column Player's Strategy

The Gridiron game resulted in the following mixed strategies for Offense (0.50, 0, 0.31, 0.19) and for Defense (0.31, 0.63, 0.06). Thus, Game Theory is 'conditionally normative' and suggests how each side ought to play to achieve certain ends (Luce, 1957:63). Here, Offense should Run 50% of the time, never play the Short Pass, play Medium Pass 31% of the time, and go Long 19% of the time. The Defensive strategies can be interpreted in a similar manner. Additionally, the value of the game, *z*, is 2.5. By convention, the row player maximizes and the column player minimizes. Thus, the game value suggests Offense is expected to gain 2.5 yards, on average, on each play of Gridiron.

The Gridiron game demonstrated a zero-sum game solution. For non-constant sum games, solutions in pure strategies can be found via the algorithm in Figure 26. However, not all non-constant sum games have solutions in pure strategies. Although every game has at least one equilibrium point in mixed strategies, finding these points for non-constant sum games requires more complex solution techniques such as the reverse search algorithm for polyhedral vertex enumeration and convex hull problems (Avis, 2002:350).

For player I with outcome matrix **A** and player II with outcome matrix **B**:

1. In each column of **A**, underline the largest value in the column.
2. In each row of **B**, underline the largest value in each row.
3. Positions *ij* in **A** and **B** in which both a_{ij} and b_{ij} were underlined are equilibrium points in pure strategies.

Example:

$$\begin{array}{ccccc} & & x & y & z \\ \textbf{Player I} & i & 0 & 0 & \underline{8} \\ \textbf{Outcome} & j & 2 & 1 & 0 \\ \textbf{Matrix} & k & \underline{7} & \underline{2} & 0 \end{array} \qquad \begin{array}{ccccc} & & x & y & z \\ \textbf{Player II} & i & 0 & \underline{5} & 1 \\ \textbf{Outcome} & j & \underline{6} & 0 & 0 \\ \textbf{Matrix} & k & 0 & \underline{3} & 1 \end{array}$$

Figure 26. Non-constant Sum Pure Strategy Solution Algorithm

(Kaplan, 1982:154)

The aim of the above review was to illustrate the benefits and uses of Game Theory. Specifically, Game Theory is a mathematically robust approach to thinking strategically in conflict situations involving other decision-making entities with conflicting objectives.

APPENDIX B: RATIO NORMALIZATION

The purpose of this appendix is to illustrate a normalization technique which makes dissimilar measures commensurate for the purposes of aggregation, while preserving the original ratio level meaning of the individual measurements. The example concerns a notional system with five, context relevant attributes and their associated measures. In line with the algorithm in Figure 11, it is assumed a desired end-state has been specified and equal interval measurement scales have been developed for the five system attributes. System measurements are taken at five points in time. The desired system attribute assignments, as well as the equal interval transformations for each of the five system attributes are shown in Table 15. Further, Table 16 shows the distance from the desired end-state value for each attribute. Using the values in Table 16 as inputs to the equation in step 5 of Figure 11, the normalization constants (k_j) for each system attribute, at each time step (j) are shown in Table 17.

As seen in Table 17, the normalization 'constants' are changing at each time step with each new system observation. Because the constants are changing from one time step to another, normalized values for the current system state are not comparable to normalized values of previous system states based on different constants. In order to compare the current state to previous system states, all previous attribute values must be normalized using the calculated constants from the most recent system observation. For the five observations in this illustration, the result is shown in Table 18. While the values shown in Observation$_4$ of Table 18 are normalized within each attribute dimension, the attributes are not yet normalized with respect to each other. This is achieved by scaling the values in each attribute dimension relative to a common system observation. For this illustration, Observation_0 is used as the common reference and the results are shown in Table 19. Additionally, Table 19 shows the aggregated attribute observations (assuming equal weighting among the attributes) using the Power metric for r = 1 (rectilinear) and r = 2 (straight-line).

Table 15. Observed System Attribute Assignments

System Attribute:	A	B	C	D	E
$X_{DESIRED}$	10	2	100	15	0.5
$Time_0$	2	100	80	25	10
$Time_1$	4	75	85	20	5
$Time_2$	6	50	92	15	1
$Time_3$	8	10	96	10	0.7
$Time_4$	9	5	98	12	0.2

Table 16. Attribute Distance from Desired

System Attribute:	A	B	C	D	E
$X_{DESIRED}$	0	0	0	0	0
$Time_0$	8	98	20	10	9.5
$Time_1$	6	73	15	5	4.5
$Time_2$	4	48	8	0	0.5
$Time_3$	2	8	4	5	0.2
$Time_4$	1	3	2	3	0.3

Table 17. System Attribute Normalization Constants

System Attribute:	A	B	C	D	E
$X_{DESIRED}$	0.000	0.000	0.000	0.000	0.000
$Time_0$	0.125	0.010	0.050	0.100	0.105
$Time_1$	0.071	0.006	0.029	0.067	0.071
$Time_2$	0.056	0.005	0.023	0.067	0.069
$Time_3$	0.050	0.004	0.021	0.050	0.068
$Time_4$	0.048	0.004	0.020	0.043	0.067

Table 18. Normalized Values for each System Observation

$Observation_0$	A	B	C	D	E
$X_{DESIRED}$	0	0	0	0	0
$Time_0$	1	1	1	1	1
$Observation_1$	A	B	C	D	E
$X_{DESIRED}$	0	0	0	0	0
$Time_0$	0.57	0.57	0.57	0.67	0.68
$Time_1$	0.43	0.43	0.43	0.33	0.32
$Observation_2$	A	B	C	D	E
$X_{DESIRED}$	0	0	0	0	0
$Time_0$	0.44	0.45	0.47	0.67	0.66
$Time_1$	0.33	0.33	0.35	0.33	0.31
$Time_2$	0.22	0.22	0.19	0.00	0.03
$Observation_3$	A	B	C	D	E
$X_{DESIRED}$	0	0	0	0	0
$Time_0$	0.40	0.43	0.43	0.50	0.65
$Time_1$	0.30	0.32	0.32	0.25	0.31
$Time_2$	0.20	0.21	0.17	0.00	0.03
$Time_3$	0.10	0.04	0.09	0.25	0.01
$Observation_4$	A	B	C	D	E
$X_{DESIRED}$	0	0	0	0	0
$Time_0$	0.38	0.43	0.41	0.43	0.63
$Time_1$	0.29	0.32	0.31	0.22	0.30
$Time_2$	0.19	0.21	0.16	0.00	0.03
$Time_3$	0.10	0.03	0.08	0.22	0.01
$Time_4$	0.05	0.01	0.04	0.13	0.02

Table 19. Normalized State Values

System Attribute:	A	B	C	D	E	r = 1	r = 2
$Time_0$	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$Time_1$	0.75	0.74	0.75	0.50	0.47	0.64	0.66
$Time_2$	0.50	0.49	0.40	0.00	0.05	0.29	0.36
$Time_3$	0.25	0.08	0.20	0.50	0.02	0.21	0.27
$Time_4$	0.13	0.03	0.10	0.30	0.03	0.12	0.15

Theory Of Effectiveness Measurement

Appendix C: Kalman Filtering

The following example is designed to illustrate the detailed mechanics involved in implementing a Kalman filter. The notional system of interest, ***X***, will have two attributes, *A* and *B*. No model is available for ***X*** and the underlying behavior and relationships between *A* and *B* are unknown. A 2nd order Kalman filter will be used to estimate the state of the system based on observations, or measurements, of the system's two attributes.

It is assumed numerous measurements will be taken, thus, process noise can be set to zero. Additionally, it is assumed there is no a priori information on how to initialize the filter. These two assumptions greatly simplify the filter state estimate calculations. The matrix form of a second-order Kalman filter for a single system attribute (14) in expanded form is:

$$\begin{aligned}
\hat{x}_k &= \hat{x}_{k-1} + T_s\hat{\dot{x}}_{k-1} + .5T_s^2\hat{\ddot{x}}_{k-1} + K_{1_k}(x_k^* - \hat{x}_{k-1} - T_s\hat{\dot{x}}_{k-1} - .5T_s^2\hat{\ddot{x}}_{k-1}) \\
\hat{\dot{x}}_k &= \hat{\dot{x}}_{k-1} + T_s\hat{\ddot{x}}_{k-1} + K_{2_k}(x_k^* - \hat{x}_{k-1} - T_s\hat{\dot{x}}_{k-1} - .5T_s^2\hat{\ddot{x}}_{k-1}) \\
\hat{\ddot{x}}_k &= \hat{\ddot{x}}_{k-1} + K_{3_k}(x_k^* - \hat{x}_{k-1} - T_s\hat{\dot{x}}_{k-1} - .5T_s^2\hat{\ddot{x}}_{k-1})
\end{aligned} \qquad (15)$$

The common term, $(x_k^* - \hat{x}_{k-1} - T_s\hat{\dot{x}}_{k-1} - .5T_s^2\hat{\ddot{x}}_{k-1})$, is the residual between the most current measurement and a projection of the preceding estimate to the current time (Zarchan, 2005:113). Thus, making the substitution $res_k = (x_k^* - \hat{x}_{k-1} - T_s\hat{\dot{x}}_{k-1} - .5T_s^2\hat{\ddot{x}}_{k-1})$ produces the following set of recursive equations:

$$
\begin{aligned}
\hat{x}_k &= \hat{x}_{k-1} + T_s \hat{\dot{x}}_{k-1} + .5T_s^2 \hat{\ddot{x}}_{k-1} + K_{1_k} res_k \\
\hat{\dot{x}}_k &= \hat{\dot{x}}_{k-1} + T_s \hat{\ddot{x}}_{k-1} + K_{2_k} res_k \\
\hat{\ddot{x}}_k &= \hat{\ddot{x}}_{k-1} + K_{3_k} res_k
\end{aligned}
\tag{16}
$$

Additionally, the above two assumptions allow the gains, ***K***, to be calculated via the following equations for $k = 1, 2, \ldots, n$ (Zarchan, 2005:145):

$$
\begin{aligned}
K_{1_k} &= \frac{3(3k^2 - 3k + 2)}{k(k+1)(k+2)} \\
K_{2_k} &= \frac{18(2k-1)}{k(k+1)(k+2)T_s} \\
K_{3_k} &= \frac{60}{k(k+1)(k+2)T_s^2}
\end{aligned}
\tag{17}
$$

The above equations, (16) and (17), were implemented in a Microsoft® EXCEL® spreadsheet and used on the notional system attribute measurements with $T_s = 1$ as shown in Table 20.

Table 20. Notional Data and Results

Time	Measurement #	Kalman Gains			Observation x_k*		res_k		Estimates of *A*			Estimates of *B*		
Period	*k*	K_{1k}	K_{2k}	K_{3k}	*A*	*B*	*A*	*B*	x_k	x_dot_k	x_wdot_k	x_k	x_dot_k	x_wdot_k
0	1	1.0000	3.0000	10.0000	10.0000	1000.0000	10.0000	1000.0000	10.0000	30.0000	100.0000	1000.0000	3000.0000	10000.0000
1	2	1.0000	2.2500	2.5000	10.0972	927.7364	-79.9028	-8072.2636	10.0972	-49.7813	-99.7570	927.7364	-5162.5930	-10180.6589
2	3	1.0000	1.5000	1.0000	10.9337	907.6948	100.4963	10232.8808	10.9337	1.2061	0.7393	907.6948	6.0693	52.2219
3	4	0.9500	1.0500	0.5000	11.3544	833.2678	-1.1551	-106.6071	11.4122	0.7326	0.1618	838.5982	-53.6463	-1.0817
4	5	0.8857	0.7714	0.2857	11.7389	750.5178	-0.4867	-33.8932	11.7946	0.5189	0.0227	754.3913	-80.8742	-10.7654
5	6	0.8214	0.5893	0.1786	11.8059	733.1754	-0.5189	65.0409	11.8986	0.2358	-0.0700	721.5609	-53.3119	0.8490
6	7	0.7619	0.4643	0.1190	12.2847	690.6145	0.1853	21.9410	12.2406	0.2519	-0.0479	685.3905	-42.2760	3.4610
7	8	0.7083	0.3750	0.0833	12.4315	669.7430	-0.0370	24.8980	12.4423	0.1901	-0.0510	662.4811	-29.4782	5.5359
8	9	0.6606	0.3091	0.0606	13.6162	650.3141	1.0093	14.5433	13.2737	0.4511	0.0102	645.3782	-19.4471	6.4173
9	10	0.6182	0.2591	0.0455	14.1758	635.6134	0.4460	6.4738	14.0055	0.5768	0.0305	633.1416	-11.3526	6.7115
10	11	0.5804	0.2203	0.0350	14.7179	616.8418	0.1203	-8.3031	14.6674	0.6338	0.0347	620.3256	-6.4700	6.4212
11	12	0.5467	0.1896	0.0275	15.7338	575.7121	0.4153	-41.3541	15.5455	0.7472	0.0461	594.4578	-7.8879	5.2851
12	13	0.5165	0.1648	0.0220	16.4561	539.0188	0.1403	-50.1937	16.3882	0.8164	0.0492	563.2882	-10.8764	4.1820
13	14	0.4893	0.1446	0.0179	17.1267	502.1468	-0.1024	-52.3559	17.1790	0.8507	0.0473	528.8858	-14.2674	3.2470
14	15	0.4647	0.1279	0.0147	18.2221	485.6167	0.1687	-30.6252	18.1318	0.9196	0.0498	502.0102	-14.9386	2.7967
15	16	0.4424	0.1140	0.0123	18.8789	476.9400	-0.1974	-11.5299	18.9889	0.9469	0.0474	483.3691	-13.4560	2.6554
16	17	0.4221	0.1022	0.0103	19.5474	431.2822	-0.4121	-39.9587	19.7856	0.9522	0.0431	454.3749	-14.8831	2.2430
17	18	0.4035	0.0921	0.0088	21.3488	390.6959	0.5894	-49.9174	20.9972	1.0496	0.0483	420.4712	-17.2377	1.8051
18	19	0.3865	0.0835	0.0075	22.6429	376.6328	0.5719	-27.5032	22.2920	1.1456	0.0526	393.5070	-17.7280	1.5983
19	20	0.3708	0.0760	0.0065	23.2938	366.2288	-0.1701	-10.3494	23.4009	1.1853	0.0515	372.7409	-16.9159	1.5311
20	21	0.3563	0.0695	0.0056	24.5539	360.5393	-0.0580	3.9488	24.5913	1.2328	0.0512	357.9975	-15.1105	1.5534
21	22	0.3429	0.0637	0.0049	24.9328	327.2480	-0.9168	-16.4157	25.5353	1.2255	0.0466	338.0350	-14.6033	1.4723
22	23	0.3304	0.0587	0.0043	26.8109	304.6944	0.0268	-19.4734	26.7930	1.2738	0.0468	317.7331	-14.2740	1.3877
23	24	0.3188	0.0542	0.0038	29.1431	282.1684	1.0529	-21.9845	28.4258	1.3776	0.0508	297.1433	-14.0786	1.3031
24	25	0.3080	0.0503	0.0034	32.0189	258.3226	2.1900	-25.3936	30.5035	1.5385	0.0583	275.8941	-14.0517	1.2163
25	26	0.2979	0.0467	0.0031	33.8027	235.7770	1.7316	-26.6735	32.5870	1.6776	0.0636	254.5039	-14.0811	1.1349

26	27	0.2885	0.0435	0.0027	35.5963	220.3984	1.2999	-20.5918	34.6714	1.7978	0.0671	235.0505	-13.8423	1.0785
27	28	0.2796	0.0406	0.0025	37.2554	198.8763	0.7526	-22.8712	36.7131	1.8955	0.0690	215.3537	-13.6933	1.0222
28	29	0.2712	0.0380	0.0022	38.9933	188.5023	0.3501	-13.6692	38.7381	1.9778	0.0698	198.4646	-13.1911	0.9918
29	30	0.2633	0.0357	0.0020	41.1540	185.9133	0.4032	0.1439	40.8570	2.0620	0.0706	185.8073	-12.1942	0.9921
30	31	0.2559	0.0335	0.0018	39.1087	192.3844	-3.8455	18.2753	41.9703	2.0036	0.0635	178.7851	-10.5891	1.0256
31	32	0.2488	0.0316	0.0017	39.0853	192.5722	-4.9204	23.8635	42.7813	1.9117	0.0553	174.6467	-8.8099	1.0654
32	33	0.2422	0.0298	0.0015	38.7025	201.7955	-6.0182	35.4260	43.2633	1.7877	0.0461	174.9486	-6.6890	1.1196
33	34	0.2359	0.0282	0.0014	37.7542	209.9635	-7.3199	41.1440	43.3477	1.6278	0.0359	178.5234	-4.4111	1.1772
34	35	0.2299	0.0266	0.0013	37.6724	216.1331	-7.3210	41.4322	43.3106	1.4686	0.0264	184.2244	-2.1302	1.2305
35	36	0.2242	0.0252	0.0012	37.4115	225.0739	-7.3809	42.3644	43.1379	1.3087	0.0177	192.2058	0.1700	1.2807
36	37	0.2187	0.0240	0.0011	36.9120	231.7934	-7.5434	38.7772	42.8055	1.1456	0.0094	201.4981	2.3800	1.3232
37	38	0.2136	0.0228	0.0010	35.7966	234.3232	-8.1592	29.7836	42.2133	0.9693	0.0012	210.9003	4.3814	1.3533
38	39	0.2086	0.0217	0.0009	34.4374	245.3459	-8.7458	29.3876	41.3585	0.7809	-0.0070	222.0894	6.3715	1.3809
39	40	0.2039	0.0206	0.0009	33.8877	257.3322	-8.2482	28.1808	40.4540	0.6036	-0.0142	234.8980	8.3342	1.4054
40	41	0.1994	0.0197	0.0008	33.4984	260.3816	-7.5521	16.4467	39.5445	0.4407	-0.0203	247.2147	10.0634	1.4187
41	42	0.1951	0.0188	0.0008	31.9890	270.8083	-7.9860	12.8208	38.4169	0.2702	-0.0264	260.4889	11.7232	1.4284
42	43	0.1910	0.0180	0.0007	31.1188	278.9536	-7.5551	6.0273	37.2311	0.1081	-0.0317	274.0774	13.2600	1.4327
43	44	0.1870	0.0172	0.0007	30.4596	280.6821	-6.8637	-7.3717	36.0397	-0.0416	-0.0362	286.6751	14.5659	1.4278
44	45	0.1832	0.0165	0.0006	30.4336	284.4283	-5.5464	-17.5266	34.9637	-0.1691	-0.0396	298.7436	15.7051	1.4170
45	46	0.1796	0.0158	0.0006	30.0575	295.9551	-4.7173	-19.2022	33.9277	-0.2832	-0.0423	311.7089	16.8191	1.4059
46	47	0.1761	0.0151	0.0005	29.7373	300.7992	-3.8860	-28.4317	32.9391	-0.3844	-0.0445	324.2248	17.7944	1.3905
47	48	0.1727	0.0145	0.0005	28.3917	303.6619	-4.1408	-39.0526	31.8173	-0.4891	-0.0466	335.9699	18.6170	1.3706
48	49	0.1695	0.0140	0.0005	28.2124	311.5005	-3.0926	-43.7717	30.7809	-0.5788	-0.0481	347.8547	19.3759	1.3495
49	50	0.1663	0.0134	0.0005	27.2388	319.6022	-2.9392	-48.3032	29.6892	-0.6664	-0.0494	359.8709	20.0763	1.3277
50	51	0.1633	0.0129	0.0004	27.0393	332.4836	-1.9588	-48.1275	28.6781	-0.7411	-0.0502	372.7507	20.7815	1.3071
51	52	0.1604	0.0125	0.0004	26.5488	346.3668	-1.3631	-47.8190	27.6932	-0.8083	-0.0508	386.5148	21.4929	1.2879
52	53	0.1576	0.0120	0.0004	26.4572	362.8238	-0.4024	-45.8278	26.7961	-0.8639	-0.0509	401.4286	22.2305	1.2704
53	54	0.1549	0.0116	0.0004	26.3184	376.6544	0.4117	-47.6398	25.9705	-0.9101	-0.0508	416.9146	22.9492	1.2532
54	55	0.1523	0.0112	0.0003	25.4977	388.2686	0.4626	-52.2218	25.1055	-0.9557	-0.0506	432.5376	23.6188	1.2354
55	56	0.1498	0.0108	0.0003	24.9223	405.0131	0.7978	-51.7610	24.2439	-0.9977	-0.0504	449.0224	24.2956	1.2186
56	57	0.1473	0.0104	0.0003	24.2614	405.5084	1.0404	-68.4189	23.3743	-1.0372	-0.0500	463.8482	24.8007	1.1975
57	58	0.1449	0.0101	0.0003	23.3611	413.2443	1.0490	-76.0034	22.4641	-1.0767	-0.0497	478.2314	25.2320	1.1753
58	59	0.1427	0.0098	0.0003	23.0711	420.2371	1.7086	-83.8139	21.6063	-1.1098	-0.0493	492.0949	25.5899	1.1520
59	60	0.1404	0.0094	0.0003	22.183	430.4876	1.7106	-87.7733	20.7121	-1.1429	-0.0488	505.9350	25.9134	1.1288
60	61	0.1383	0.0091	0.0003	21.3	436.7222	1.7553	-95.6906	19.7876	-1.1756	-0.0484	519.1813	26.1675	1.1047
61	62	0.1362	0.0089	0.0002	21.013	456.7506	2.4250	-89.1505	18.9180	-1.2025	-0.0478	533.7602	26.4827	1.0833
62	63	0.1342	0.0086	0.0002	20.912	460.7689	3.2199	-100.0156	18.1236	-1.2227	-0.0470	547.3667	26.7073	1.0604
63	64	0.1322	0.0083	0.0002	20.036	465.9368	3.1585	-108.6674	17.2949	-1.2434	-0.0464	560.2395	26.8630	1.0367
64	65	0.1303	0.0081	0.0002	19.091	471.0184	3.0622	-116.6025	16.4273	-1.2650	-0.0457	572.4300	26.9577	1.0123
65	66	0.1284	0.0078	0.0002	18.874	480.7311	3.7347	-119.1628	15.6190	-1.2814	-0.0450	584.5908	27.0356	0.9886
66	67	0.1266	0.0076	0.0002	18.201	493.2815	3.8860	-118.8392	14.8071	-1.2968	-0.0442	597.0735	27.1192	0.9659
67	68	0.1249	0.0074	0.0002	18.111	496.1271	4.6226	-128.5485	14.0654	-1.3069	-0.0434	608.6247	27.1340	0.9424
68	69	0.1232	0.0072	0.0002	17.98	504.2782	5.2435	-131.9517	13.3826	-1.3125	-0.0425	619.9792	27.1275	0.9193
69	70	0.1215	0.0070	0.0002	17.818	511.3157	5.7687	-136.2507	12.7497	-1.3147	-0.0415	631.0125	27.0942	0.8965
70	71	0.1199	0.0068	0.0002	16.965	520.7042	5.5508	-137.8507	12.0797	-1.3184	-0.0406	642.0295	27.0531	0.8743
71	72	0.1183	0.0066	0.0002	16.959	540.7521	6.2184	-128.7677	11.4767	-1.3179	-0.0396	654.2859	27.0752	0.8544
72	73	0.1168	0.0064	0.0001	16.247	558.7942	6.1079	-122.9942	10.8522	-1.3182	-0.0387	667.4262	27.1374	0.8362
73	74	0.1153	0.0063	0.0001	15.491	578.9129	5.9761	-116.0688	10.2036	-1.3194	-0.0379	681.6015	27.2455	0.8197
74	75	0.1138	0.0061	0.0001	15.474	581.0453	6.6085	-128.2116	9.6174	-1.3169	-0.0370	694.6637	27.2817	0.8022
75	76	0.1124	0.0060	0.0001	15.146	601.7007	6.8640	-120.6458	9.0535	-1.3131	-0.0361	708.7858	27.3655	0.7863
76	77	0.1110	0.0058	0.0001	15.008	622.7958	7.2860	-113.7487	8.5312	-1.3069	-0.0352	723.9167	27.4916	0.7719
77	78	0.1097	0.0057	0.0001	14.757	639.5241	7.5497	-112.2702	8.0347	-1.2993	-0.0342	739.4822	27.6282	0.7583
78	79	0.1083	0.0055	0.0001	14.154	640.7331	7.4359	-126.7564	7.5240	-1.2925	-0.0334	753.7561	27.6867	0.7434
79	80	0.1071	0.0054	0.0001	13.712	657.9354	7.4976	-123.8791	7.0175	-1.2855	-0.0325	768.5523	27.7629	0.7294
80	81	0.1058	0.0053	0.0001	13.637	679.4352	7.9208	-117.2447	6.5537	-1.2764	-0.0317	784.2754	27.8760	0.7167
81	82	0.1046	0.0051	0.0001	13.348	706.1581	8.0868	-106.3516	6.1072	-1.2665	-0.0308	801.3884	28.0469	0.7055
82	83	0.1034	0.0050	0.0001	13.256	727.8767	8.4305	-101.9114	5.6967	-1.2551	-0.0300	819.2533	28.2416	0.6952
83	84	0.1022	0.0049	0.0001	12.818	738.1386	8.3914	-109.7039	5.2842	-1.2440	-0.0291	836.6310	28.3998	0.6845
84	85	0.1011	0.0048	0.0001	12.497	754.4575	8.4716	-110.9155	4.8817	-1.2326	-0.0283	854.1648	28.5537	0.6740
85	86	0.0999	0.0047	0.0001	12.042	780.5586	8.4075	-102.4969	4.4751	-1.2217	-0.0276	872.8129	28.7486	0.6647
86	87	0.0988	0.0046	0.0001	11.794	800.8892	8.5546	-101.0046	4.0851	-1.2101	-0.0268	891.9111	28.9517	0.6558
87	88	0.0978	0.0045	0.0001	11.747	809.2624	8.8856	-111.9282	3.7303	-1.1973	-0.0261	910.2484	29.1072	0.6463
88	89	0.0967	0.0044	0.0001	12.842	737.2199	10.3222	-202.4589	3.5182	-1.1782	-0.0252	920.0987	28.8686	0.6296
89	90	0.0957	0.0043	0.0001	13.038	729.5463	10.7105	-219.7357	3.3522	-1.1576	-0.0244	928.2568	28.5585	0.6121
90	91	0.0947	0.0042	0.0001	13.347	695.5929	11.1646	-261.5285	3.2395	-1.1353	-0.0235	932.3604	28.0763	0.5919
91	92	0.0937	0.0041	0.0001	14.45	630.0393	12.3579	-330.6933	3.2503	-1.1082	-0.0226	929.7489	27.3138	0.5673
92	93	0.0927	0.0040	0.0001	15.203	588.4304	13.0725	-368.9160	3.3430	-1.0783	-0.0216	923.1374	26.4018	0.5406
93	94	0.0918	0.0039	0.0001	15.749	578.244	13.4952	-371.5655	3.4925	-1.0470	-0.0207	915.7060	25.4835	0.5146
94	95	0.0909	0.0038	0.0001	16.665	559.0496	14.2296	-382.3972	3.7280	-1.0130	-0.0197	906.7032	24.5276	0.4887
95	96	0.0899	0.0038	0.0001	17.865	546.6477	15.1602	-384.8274	4.0688	-0.9756	-0.0187	896.8600	23.5665	0.4634
96	97	0.0891	0.0037	0.0001	18.766	498.1982	15.6819	-422.4600	4.4805	-0.9364	-0.0177	883.0338	22.4703	0.4364
97	98	0.0882	0.0036	0.0001	20.284	468.3824	16.7486	-437.3399	5.0122	-0.8936	-0.0167	867.1542	21.3246	0.4094
98	99	0.0873	0.0035	0.0001	22.273	430.9043	18.1628	-457.7791	5.6965	-0.8459	-0.0156	848.7044	20.1105	0.3819
99	100	0.0865	0.0035	0.0001	24.11	405.8956	19.2677	-463.1102	6.5093	-0.7945	-0.0145	828.9496	18.8822	0.3549
100	101	0.0857	0.0034	0.0001	24.529	401.7415	18.8212	-446.2678	7.3200	-0.7448	-0.0134	809.7771	17.7155	0.3297

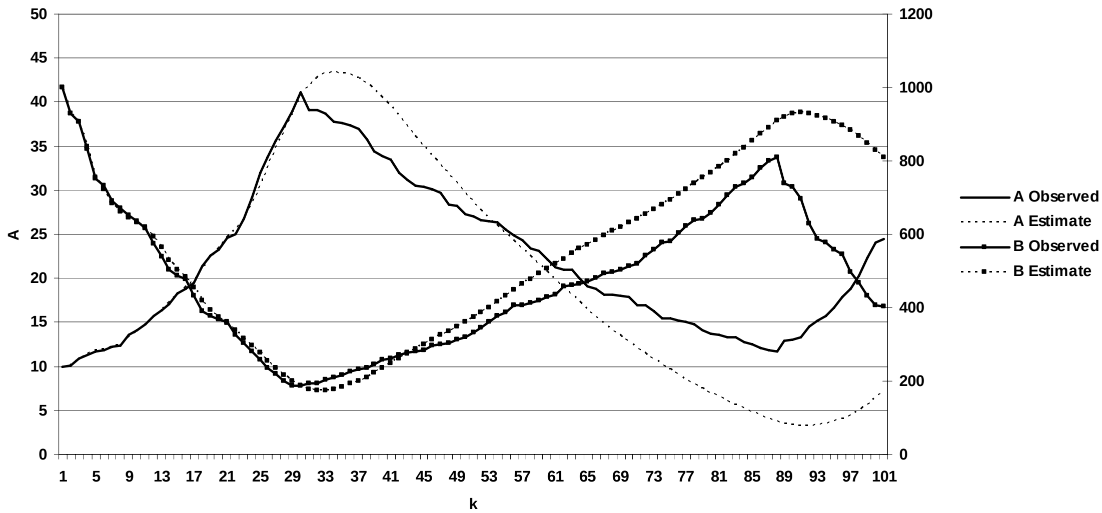


Figure 27. 2nd Order Kalman Filter Estimate of a Two Attribute Notional System

Figure 27 highlights the results of the 2nd order Kalman filter on the notional, two attribute system. Although the true system state is unknown, the filter diverges from the measurements. As previously noted, this occurs from use of a linear estimator on non-linear behavior. Also previously noted, this divergence can be addressed by increasing the sampling rate. Figure 28 shows the improved filter performance using a sampling rate of $T_s = .\overline{3}$ vs. $T_s = 1$.

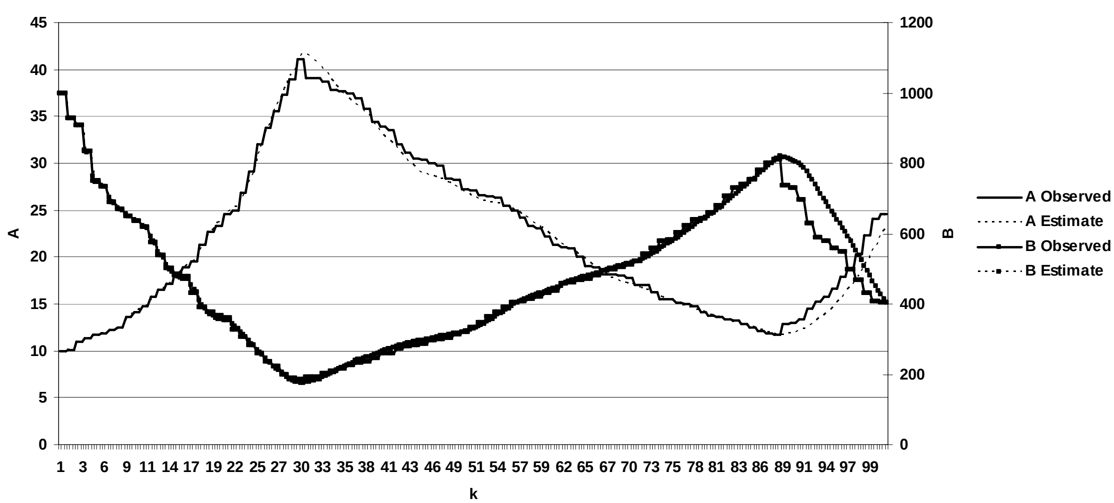
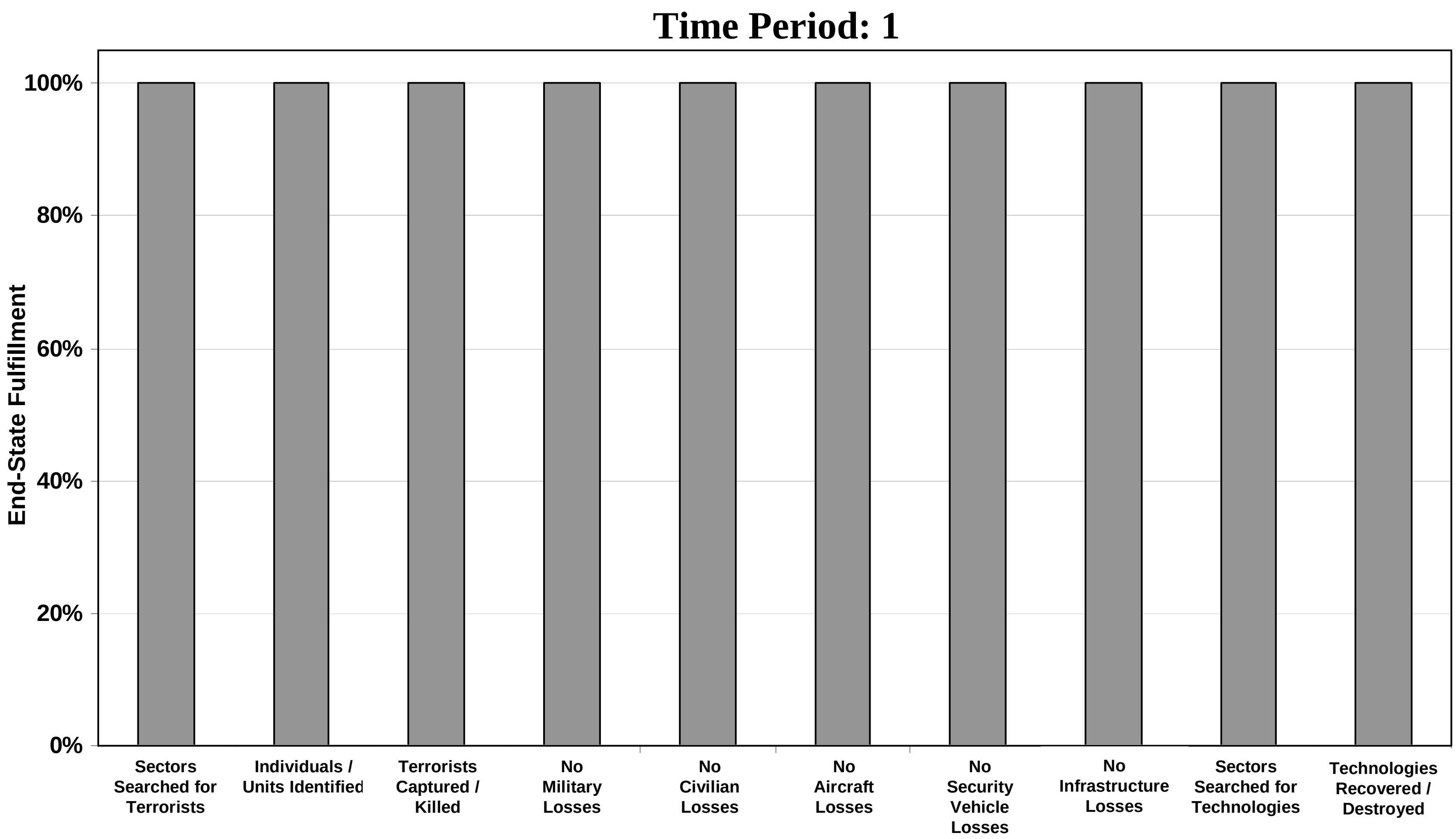


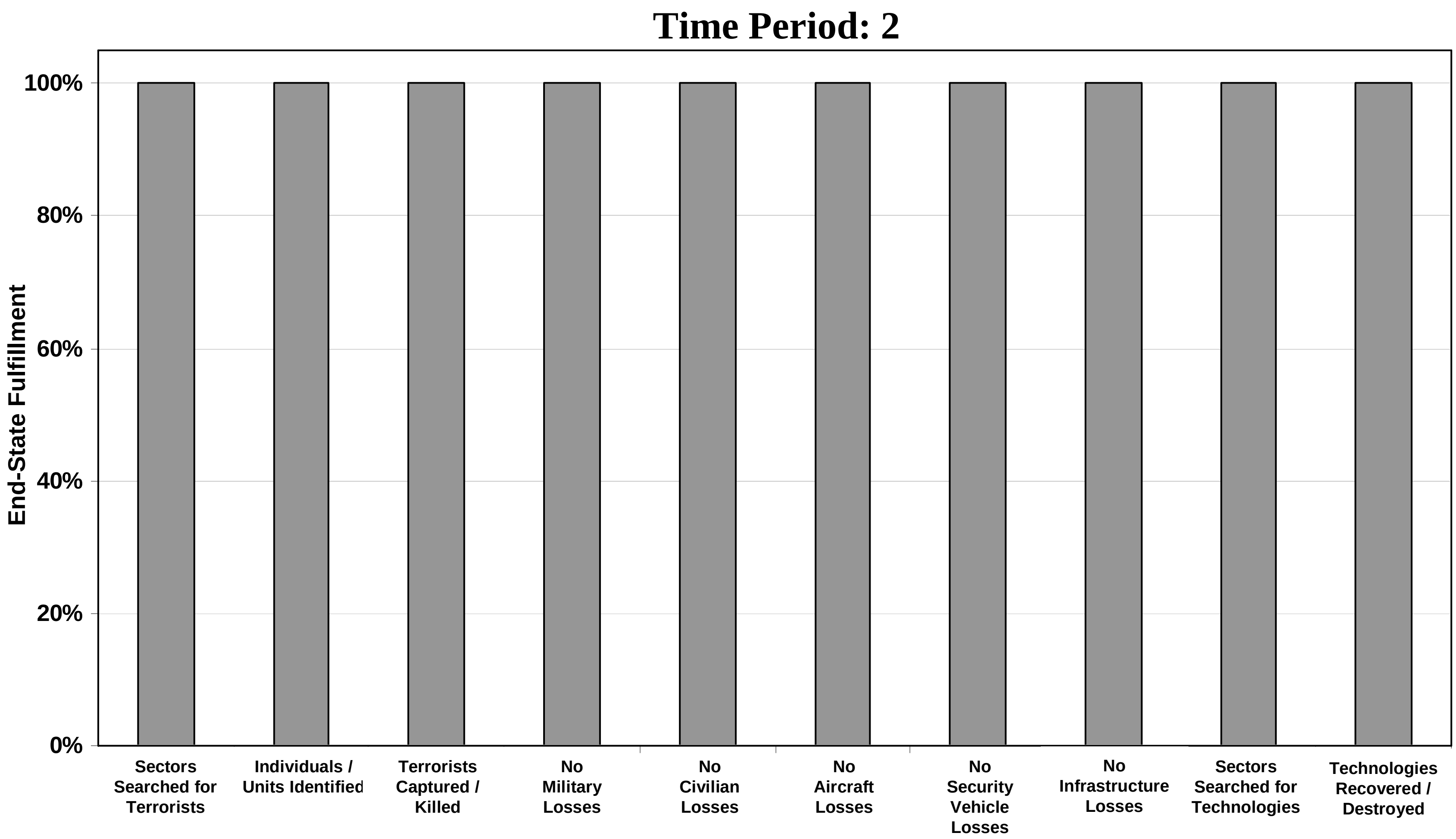
Figure 28. 2nd Order Kalman Filter Estimate with Increased Sampling

THEORY OF EFFECTIVENESS MEASUREMENT

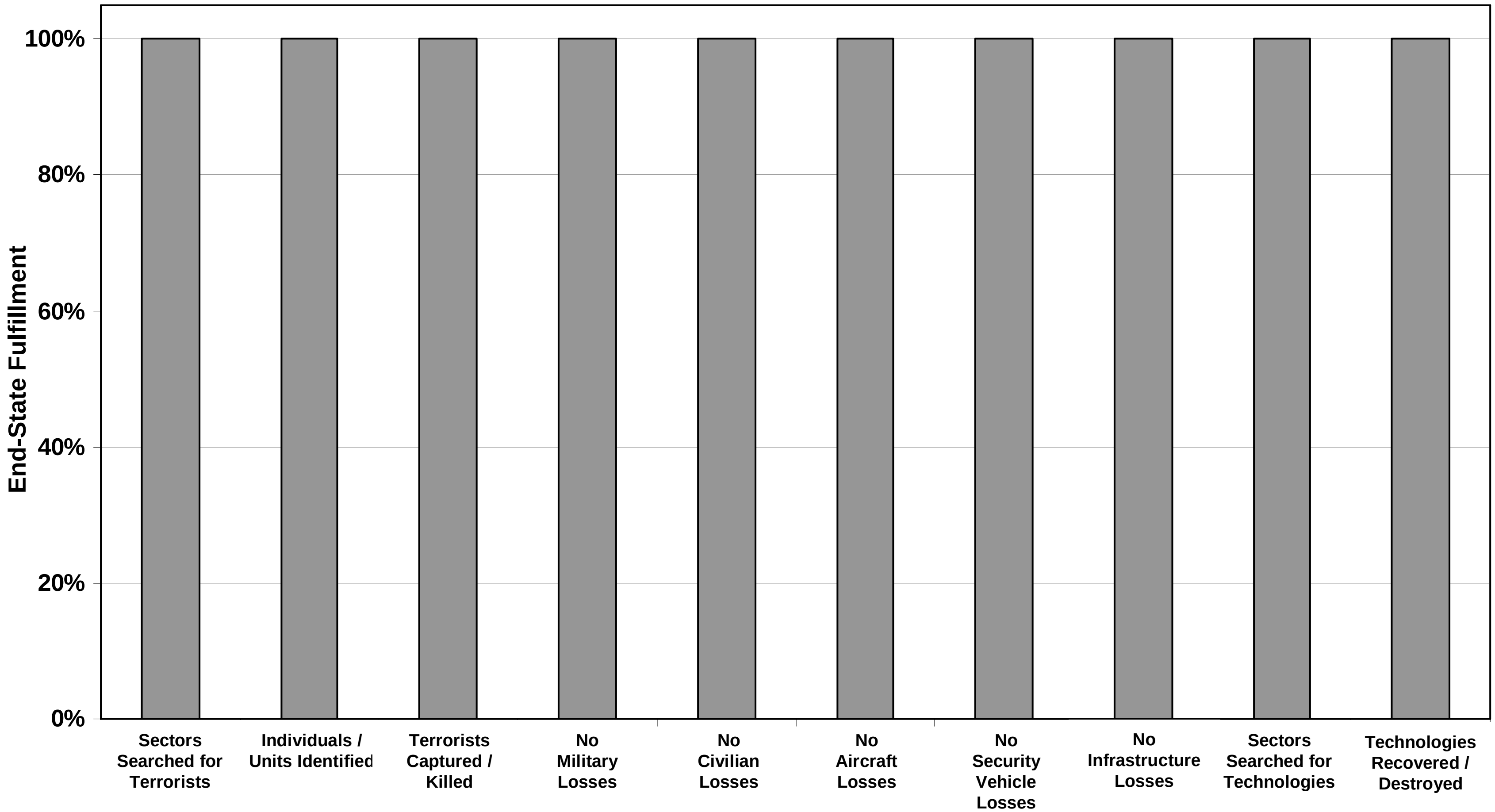
APPENDIX D: IMPLEMENTATION RESULTS (BAR CHARTS)

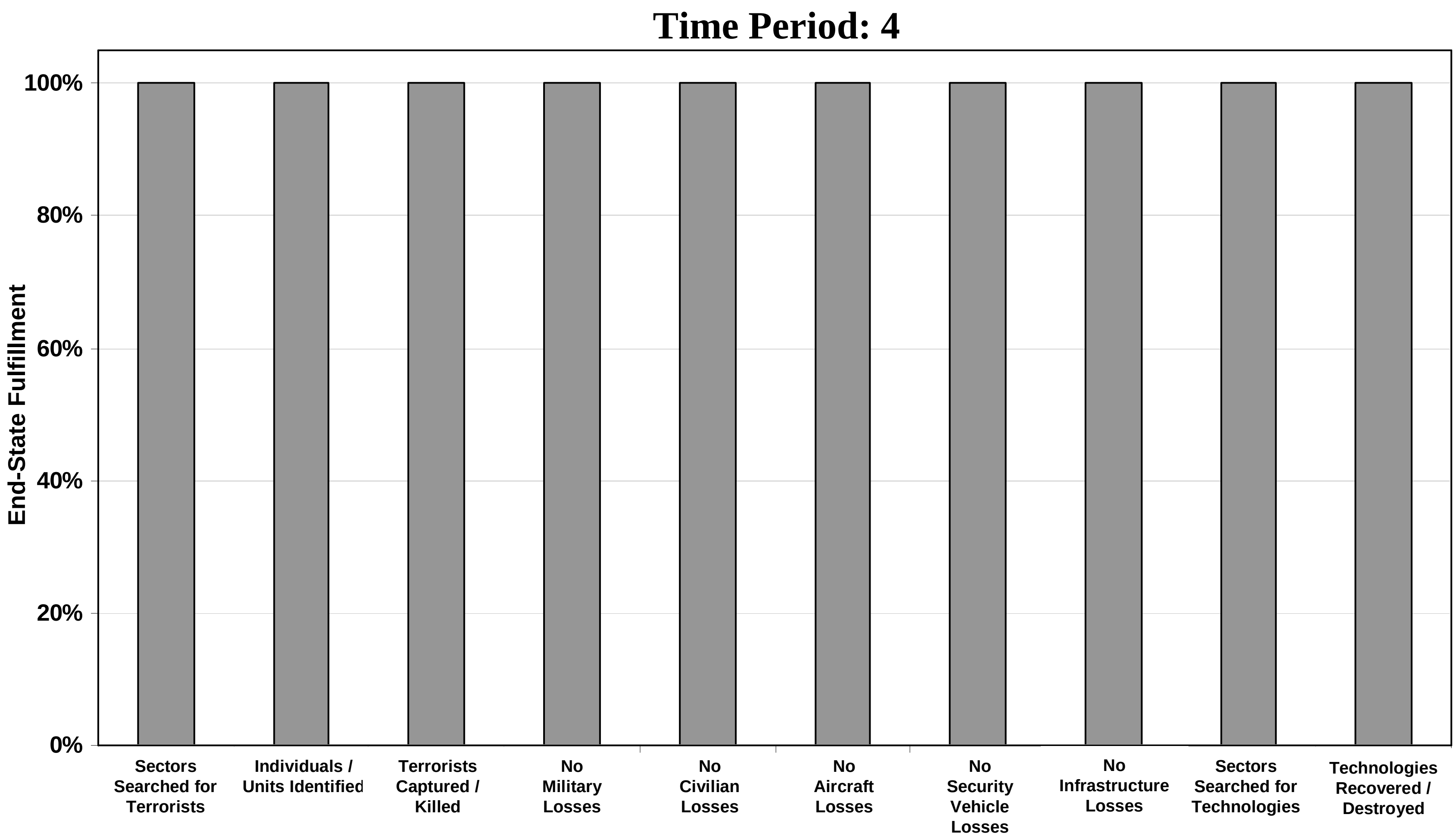
The following twenty-five figures highlight the results of the illustrative example presented in this research. The data in the charts was derived via the developed effectiveness measurement framework (Figure 23). The figures provide time independent views of end-state fulfillment, at one minute intervals, as the scenario unfolds.

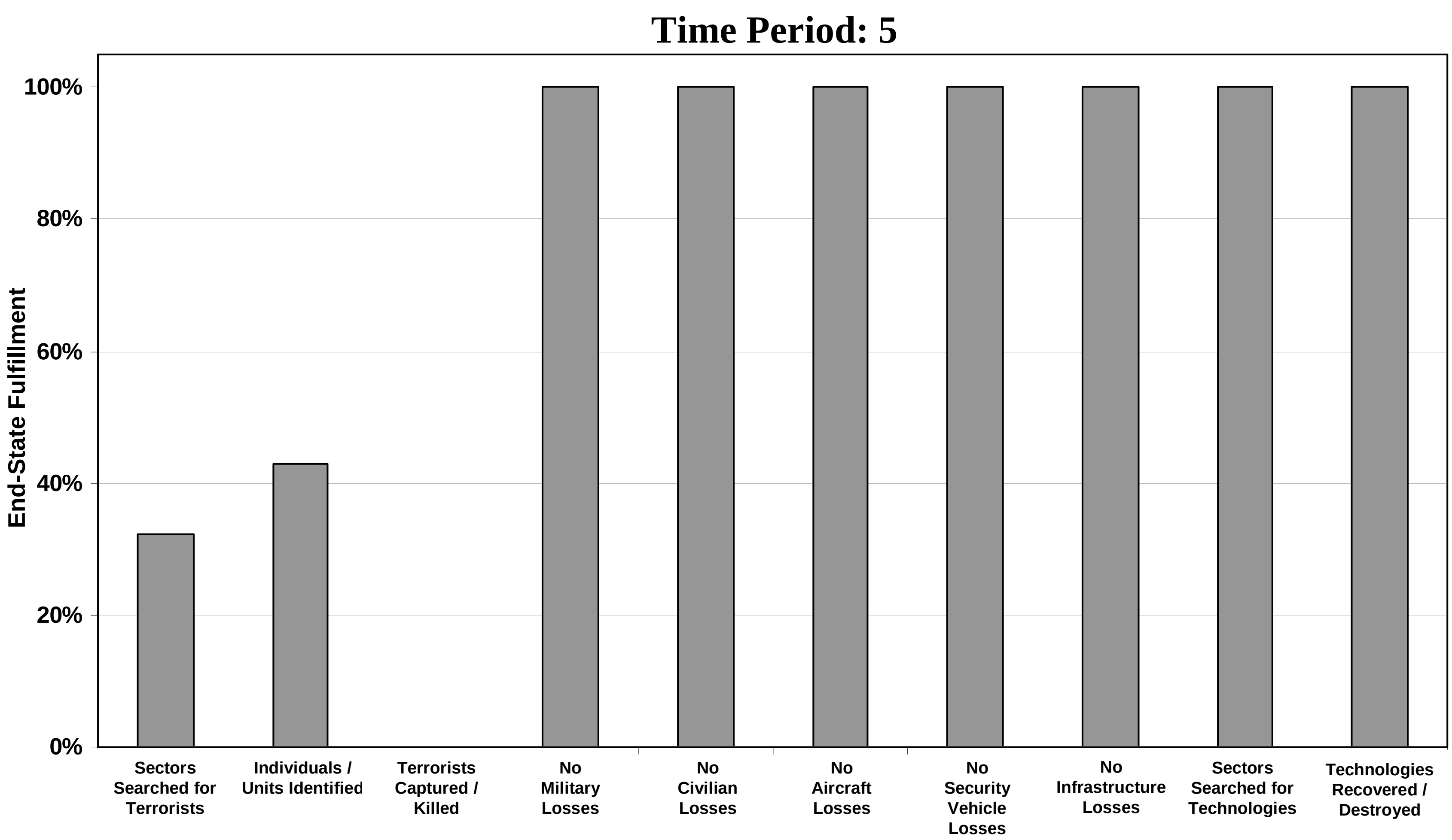

Time Period: 1
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed

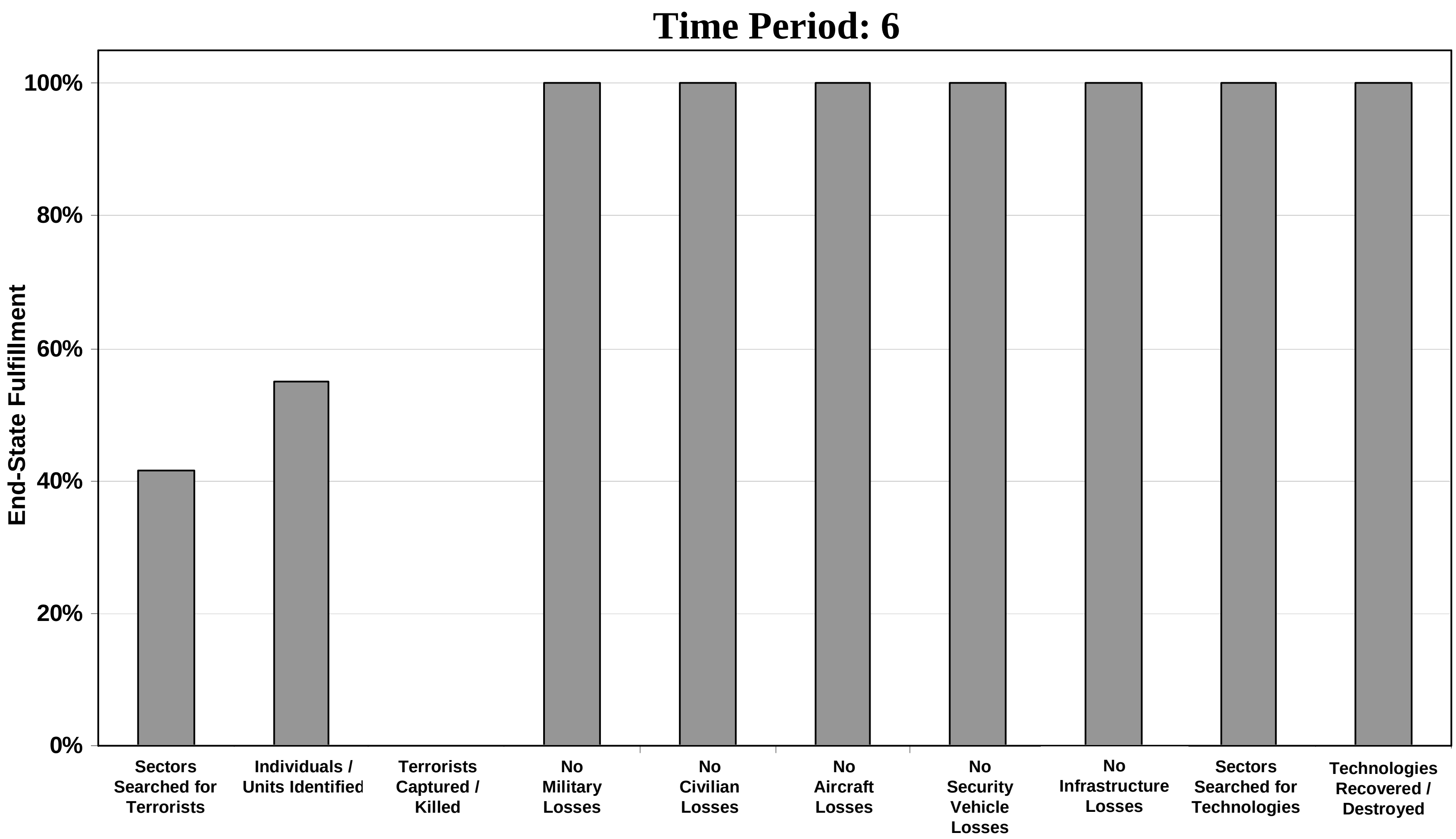

Time Period: 2
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed

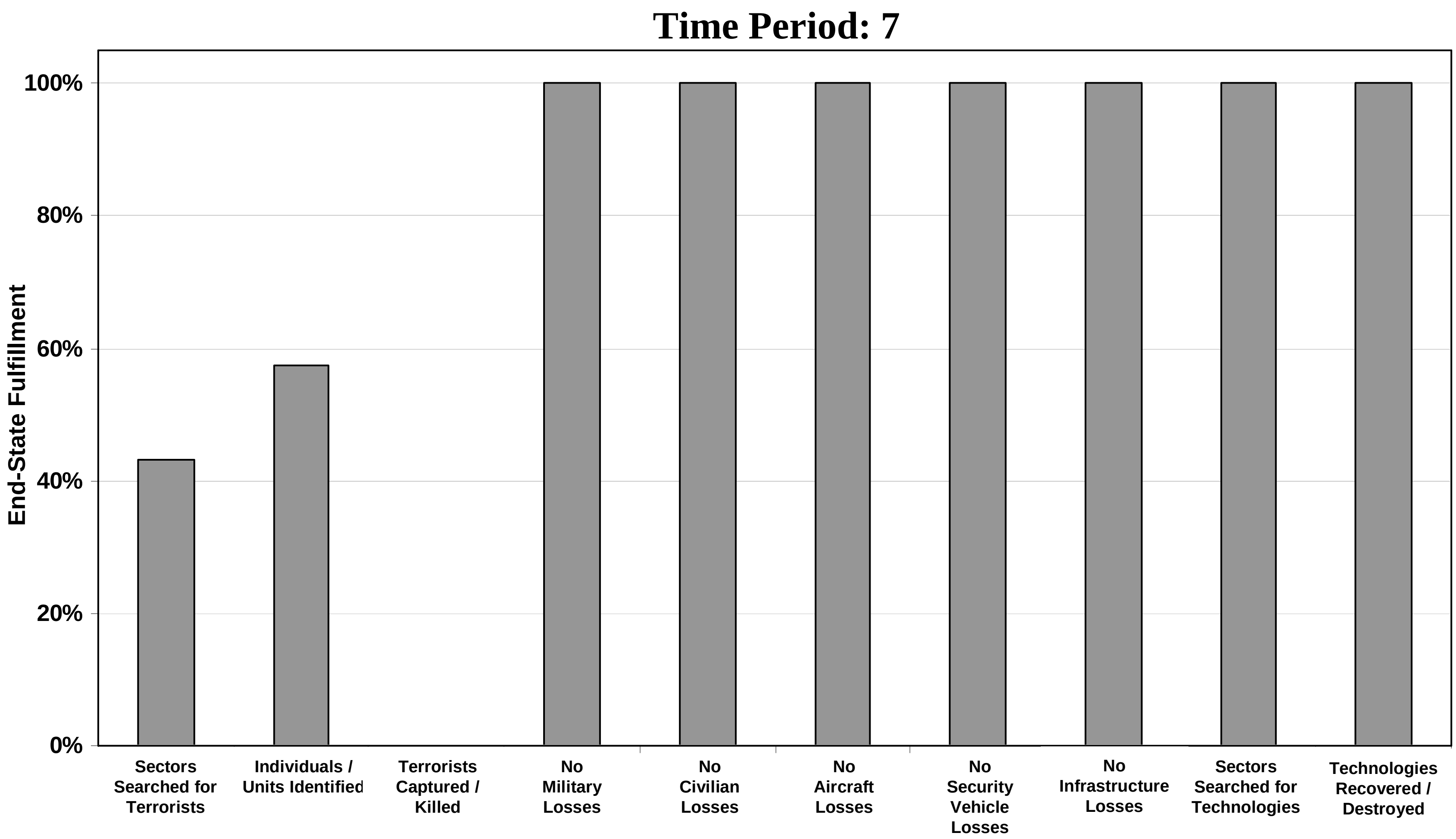
Time Period: 3

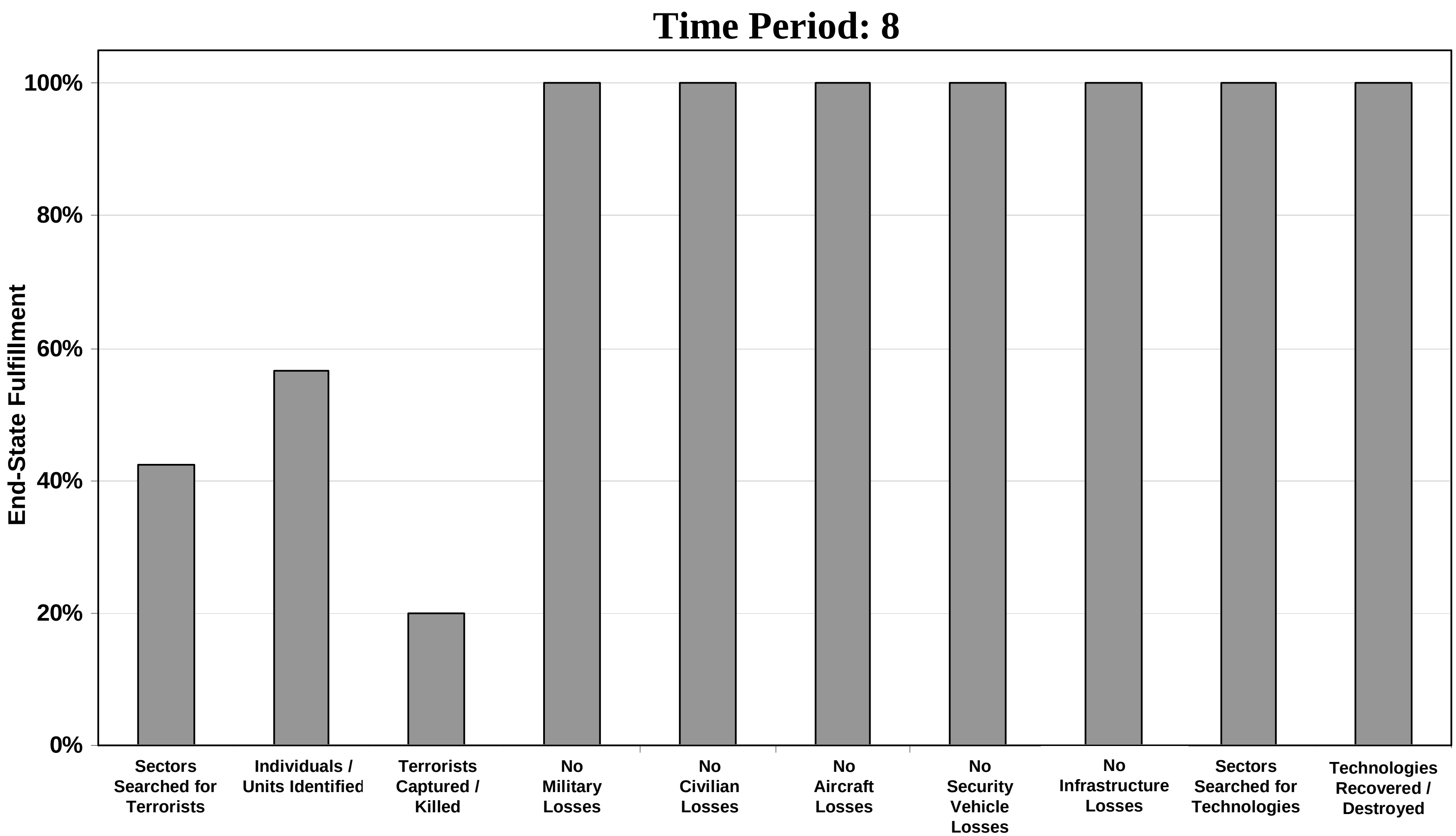



Time Period: 4
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed

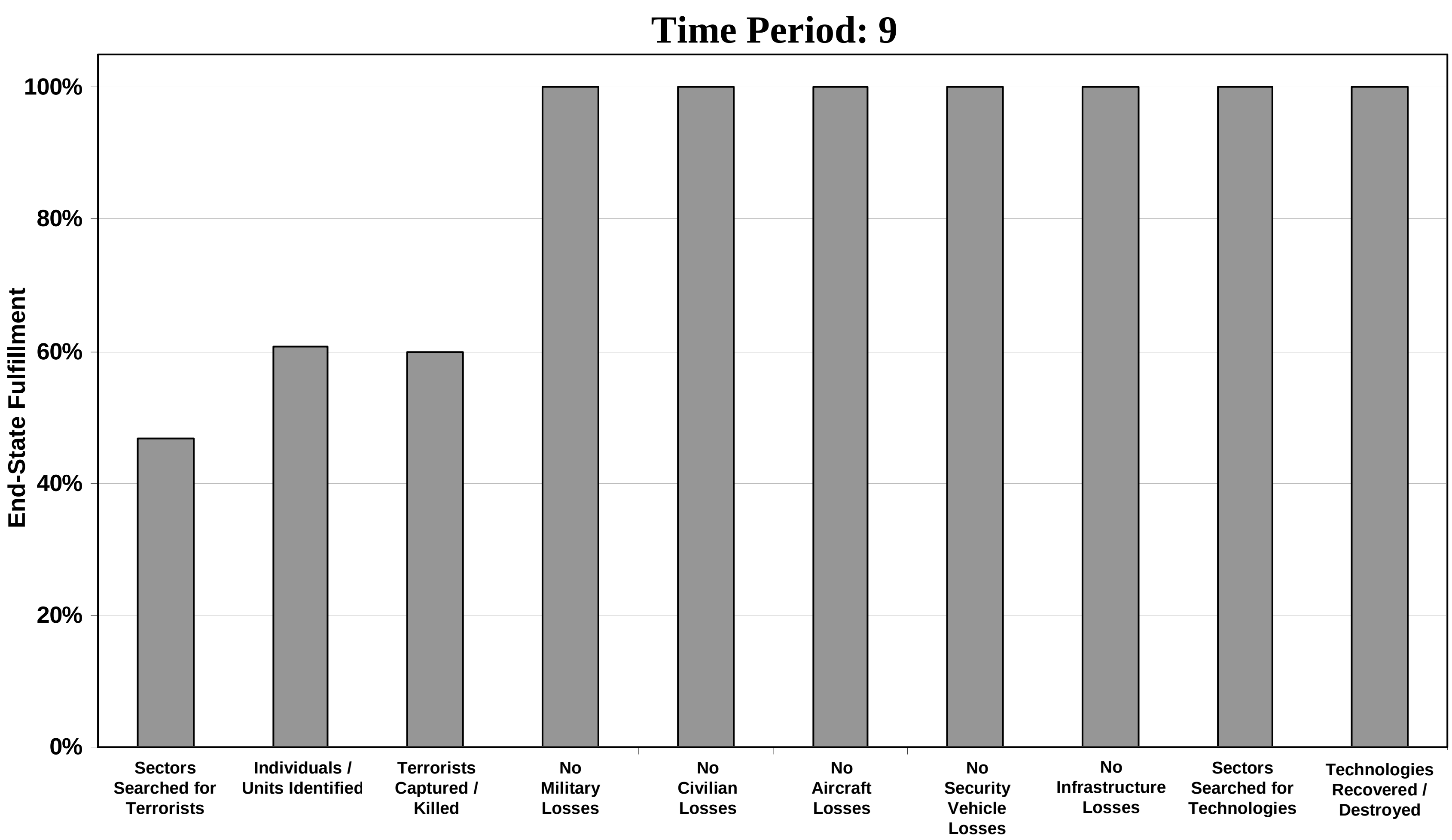

Time Period: 5
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed

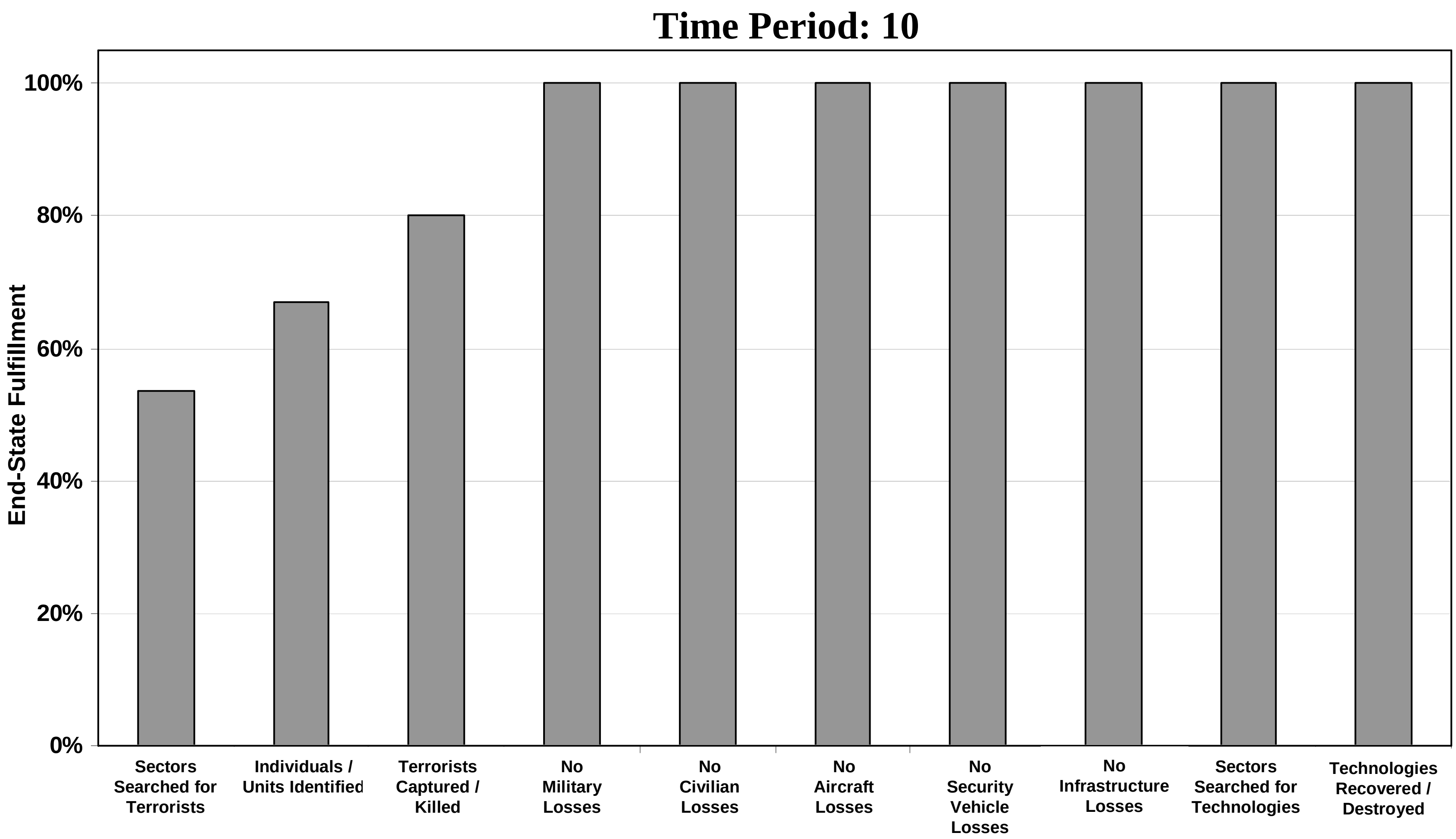

Time Period: 6
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed

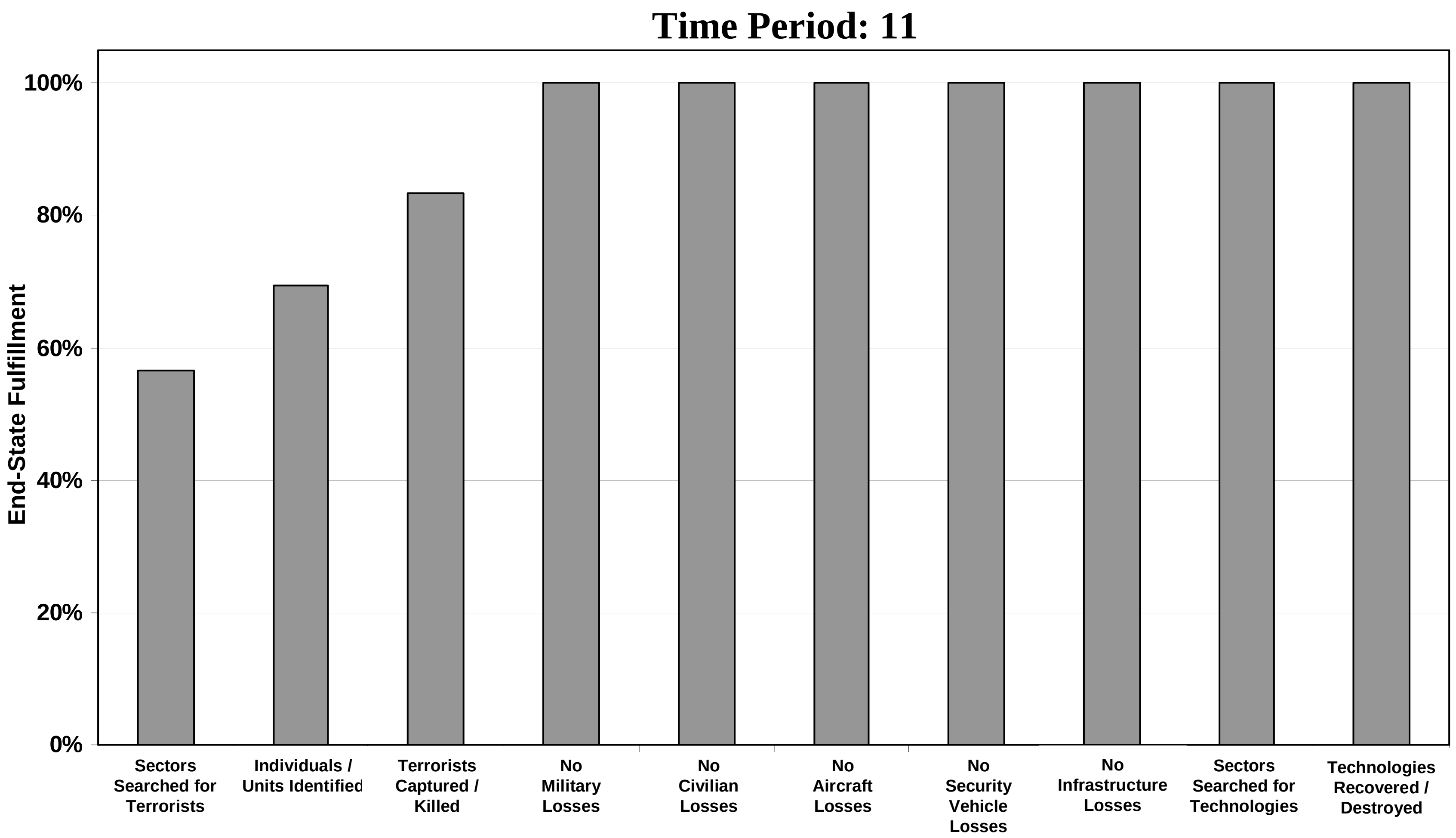

Time Period: 7
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed

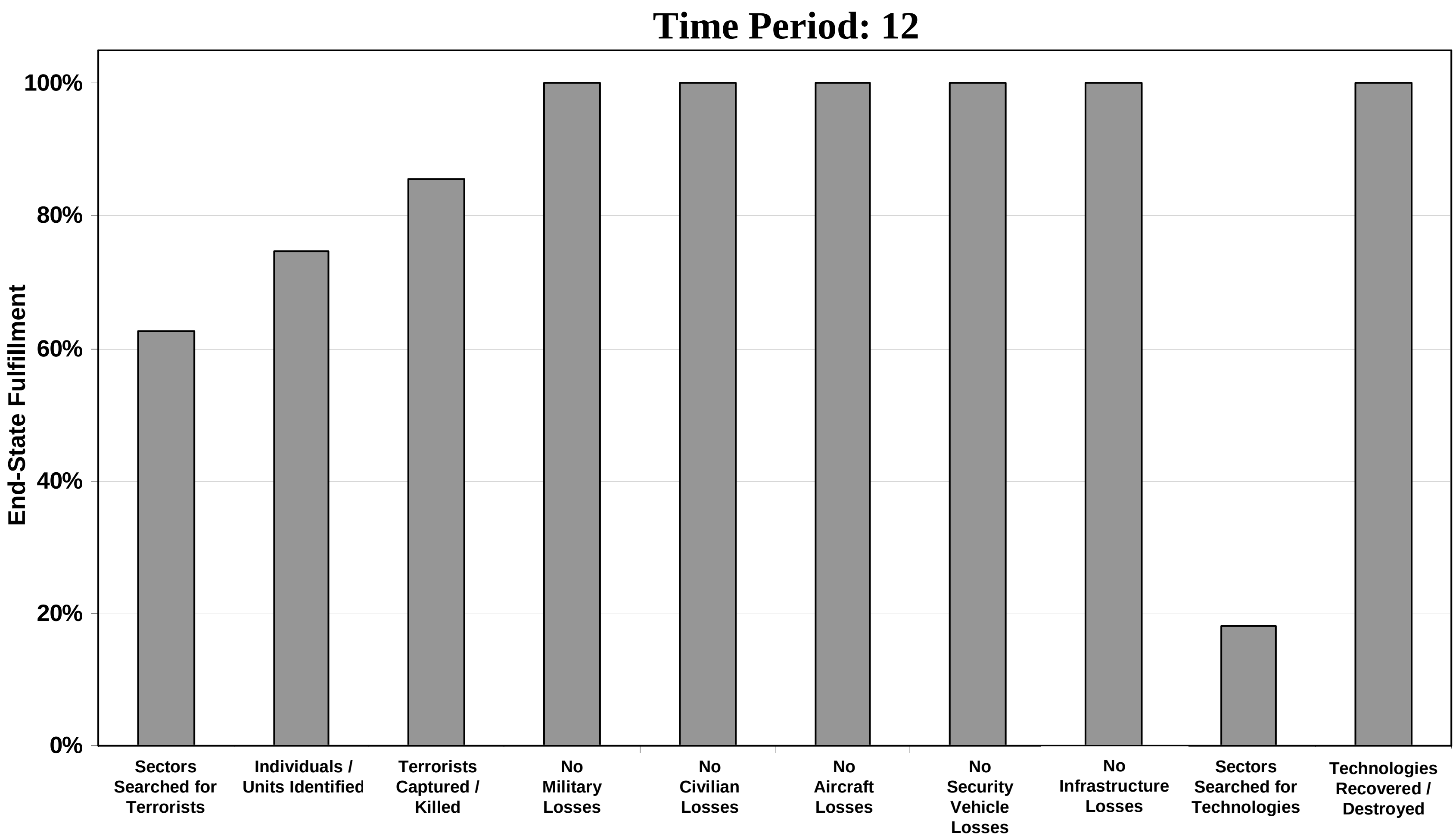


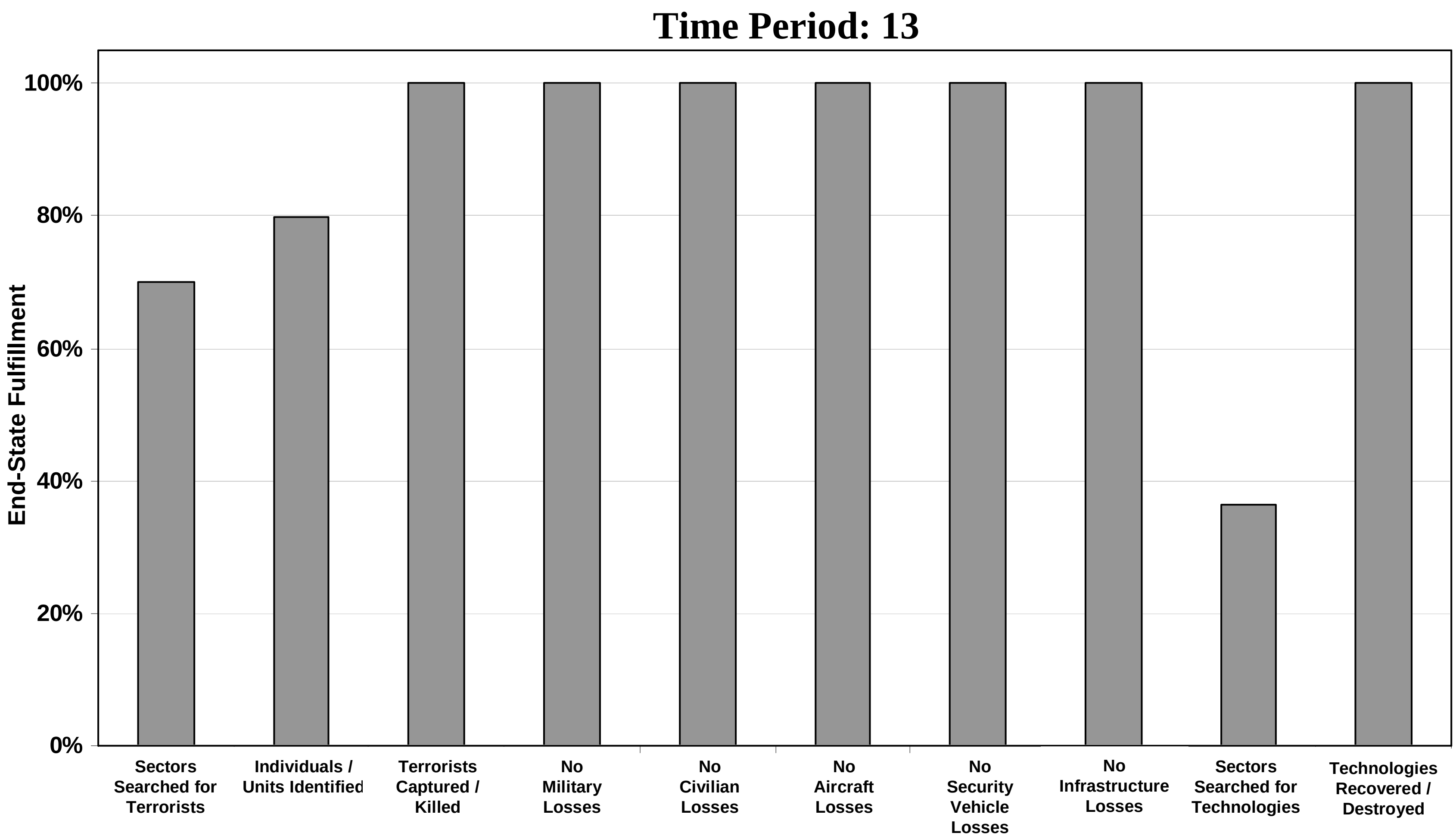
Time Period: 8
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed

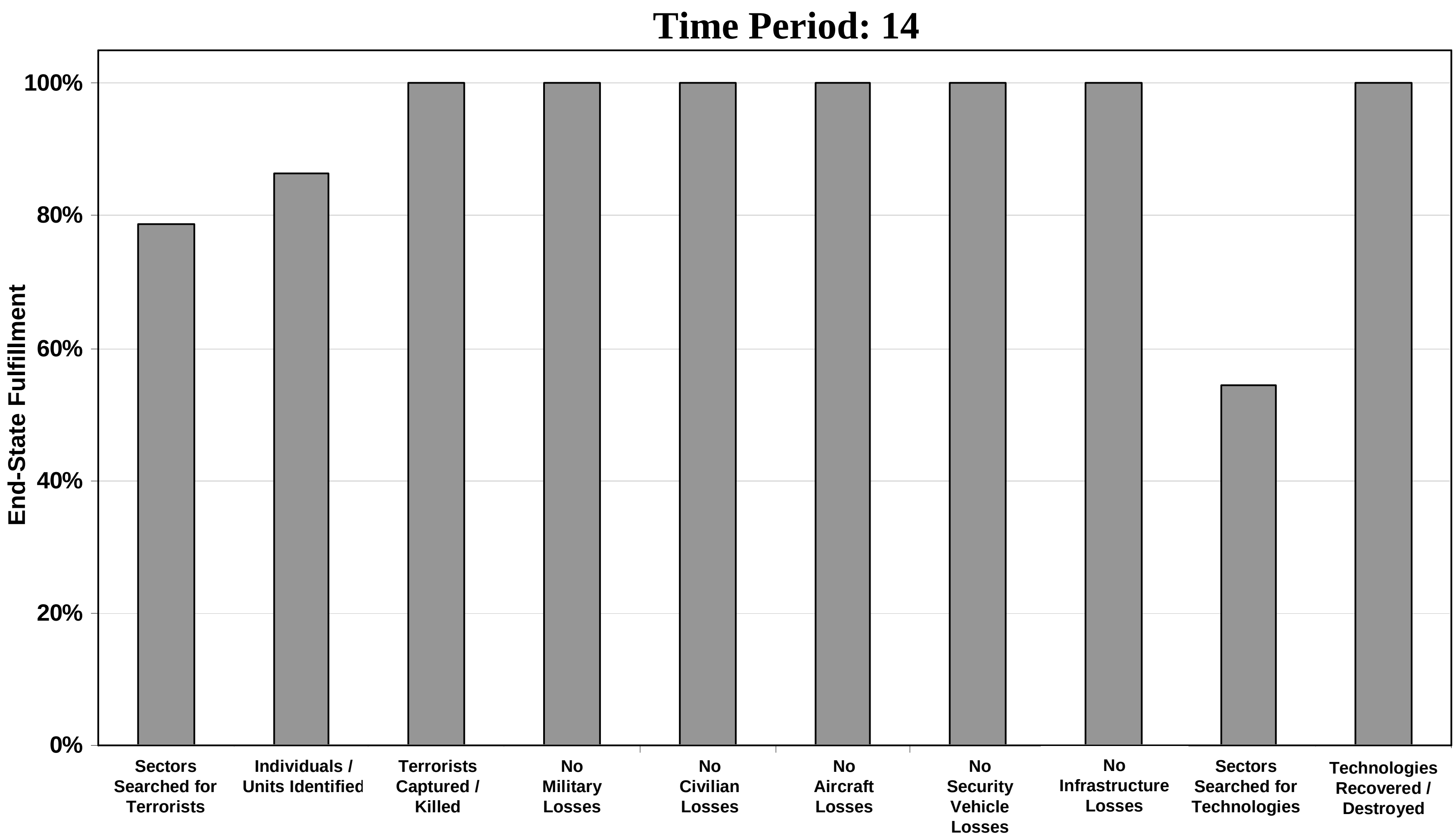

Time Period: 9
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed

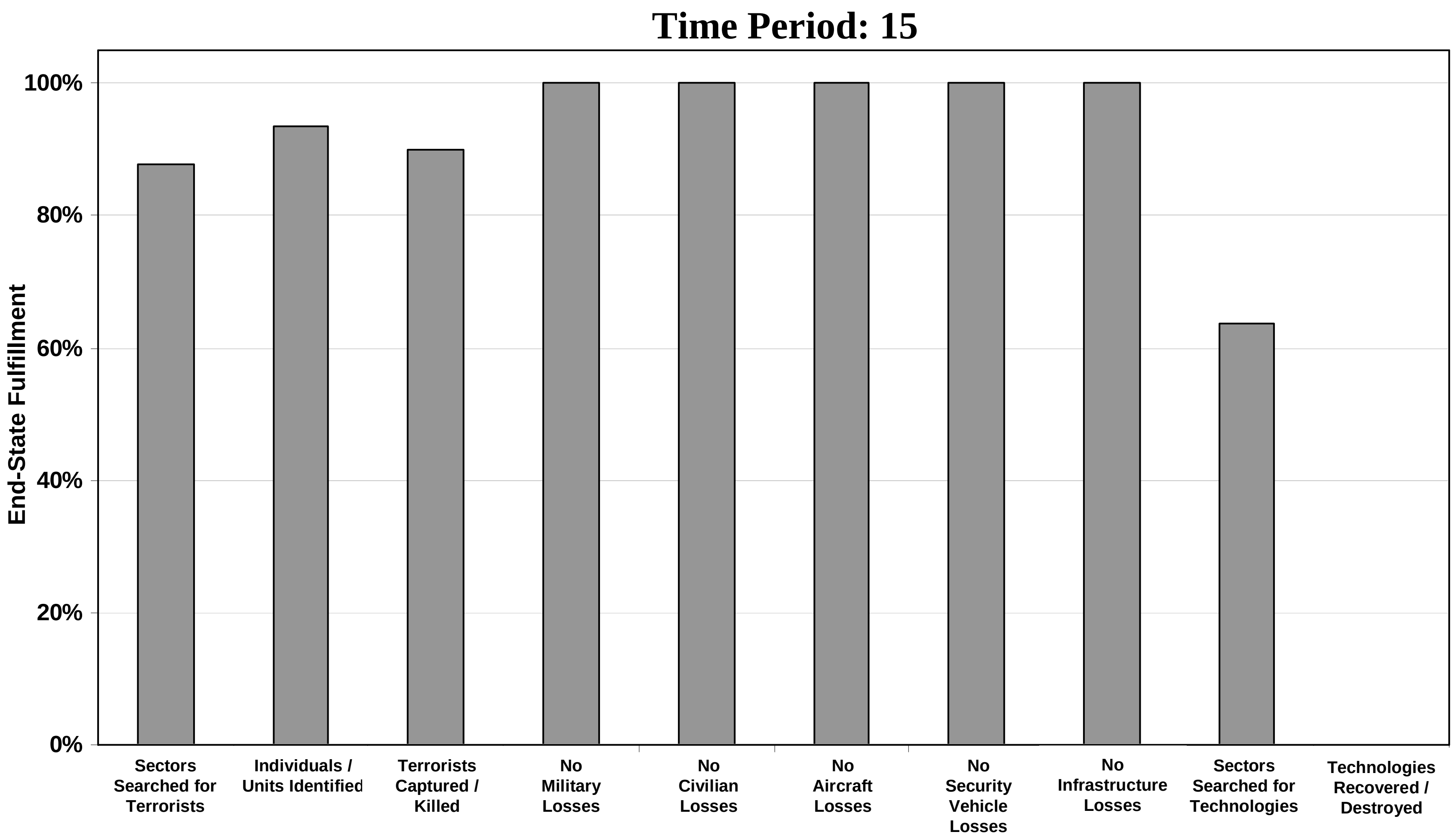

Time Period: 10
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed


Time Period: 11
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed

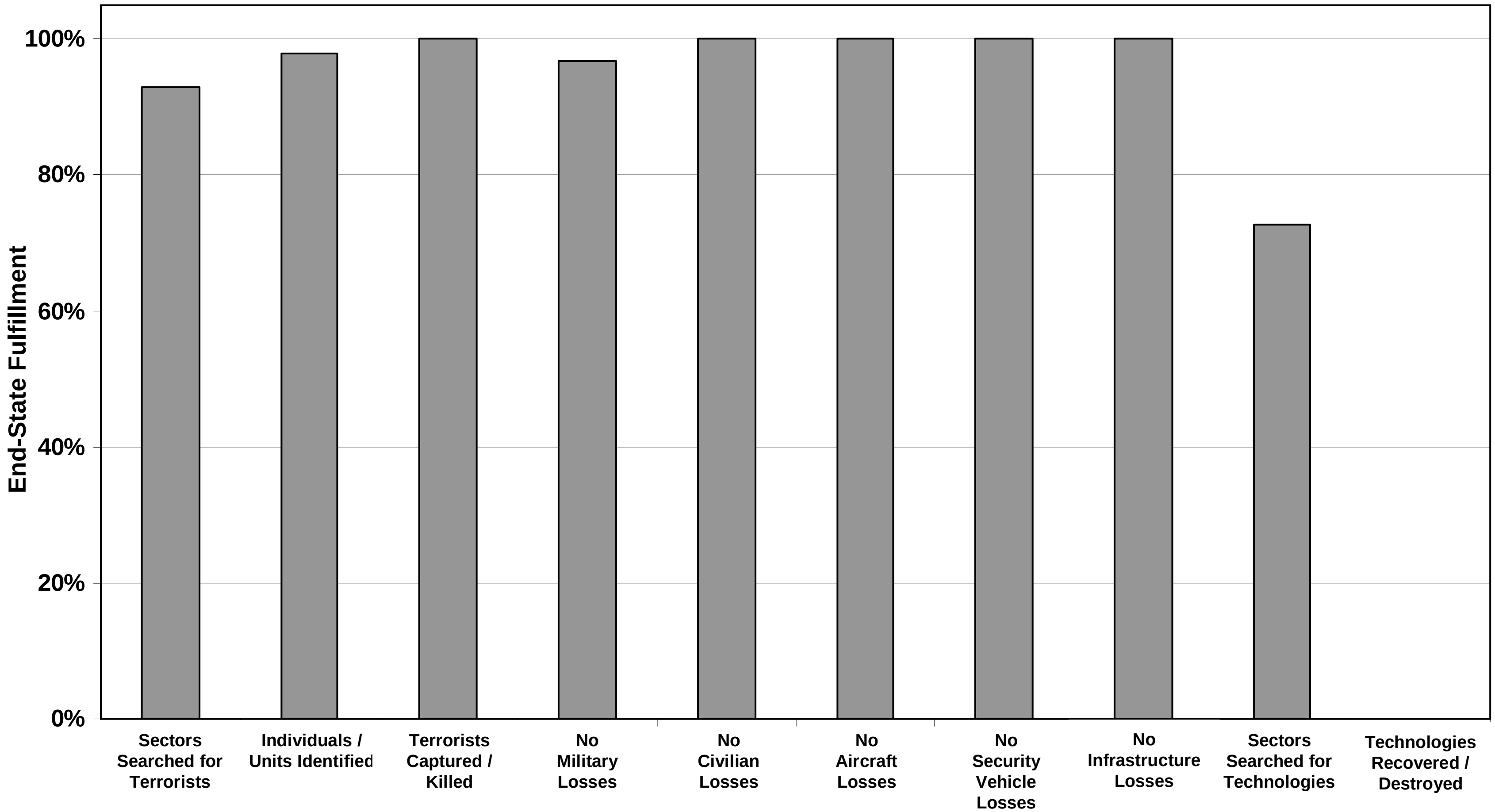

Time Period: 12
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed

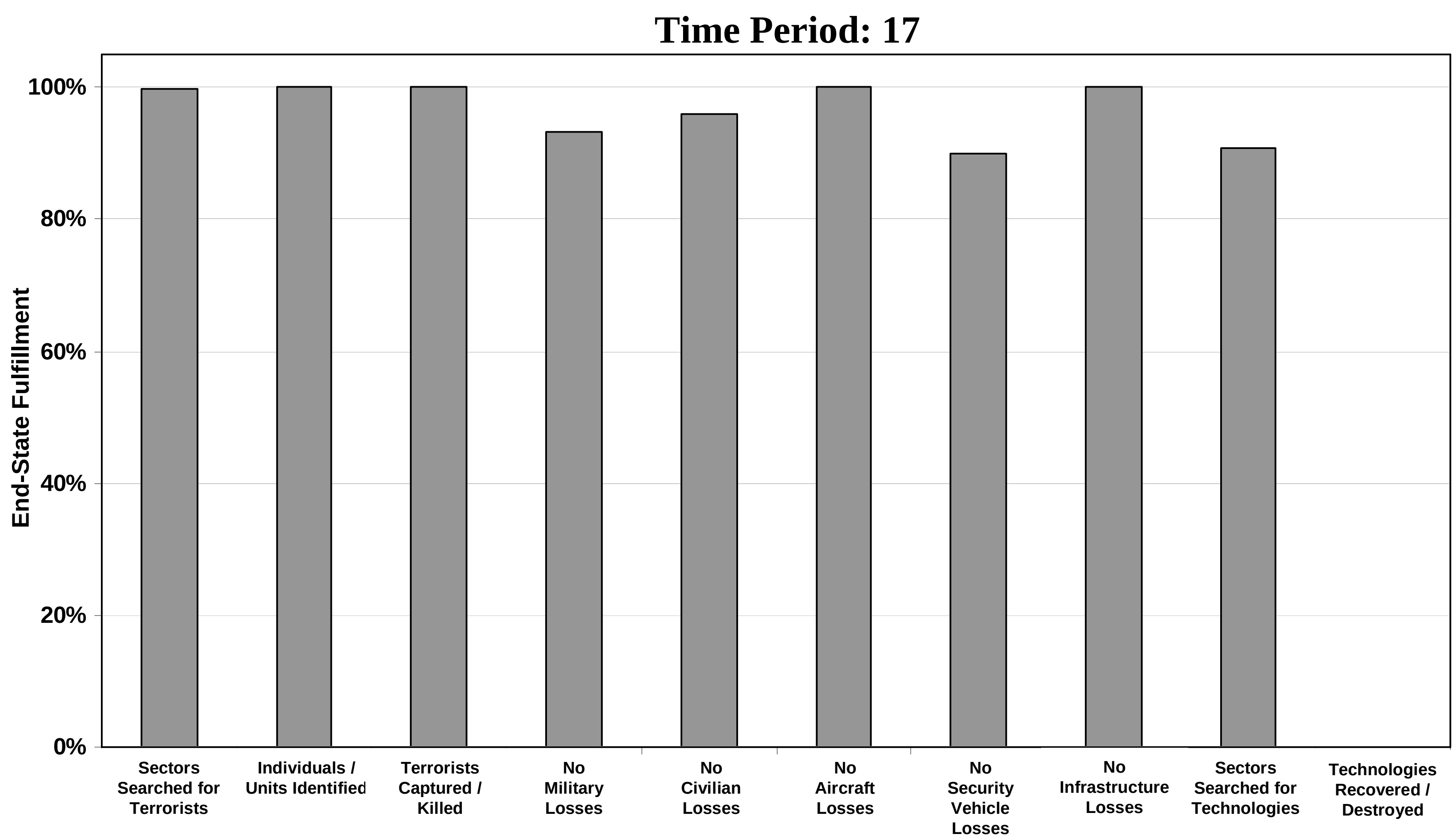

Time Period: 13
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed

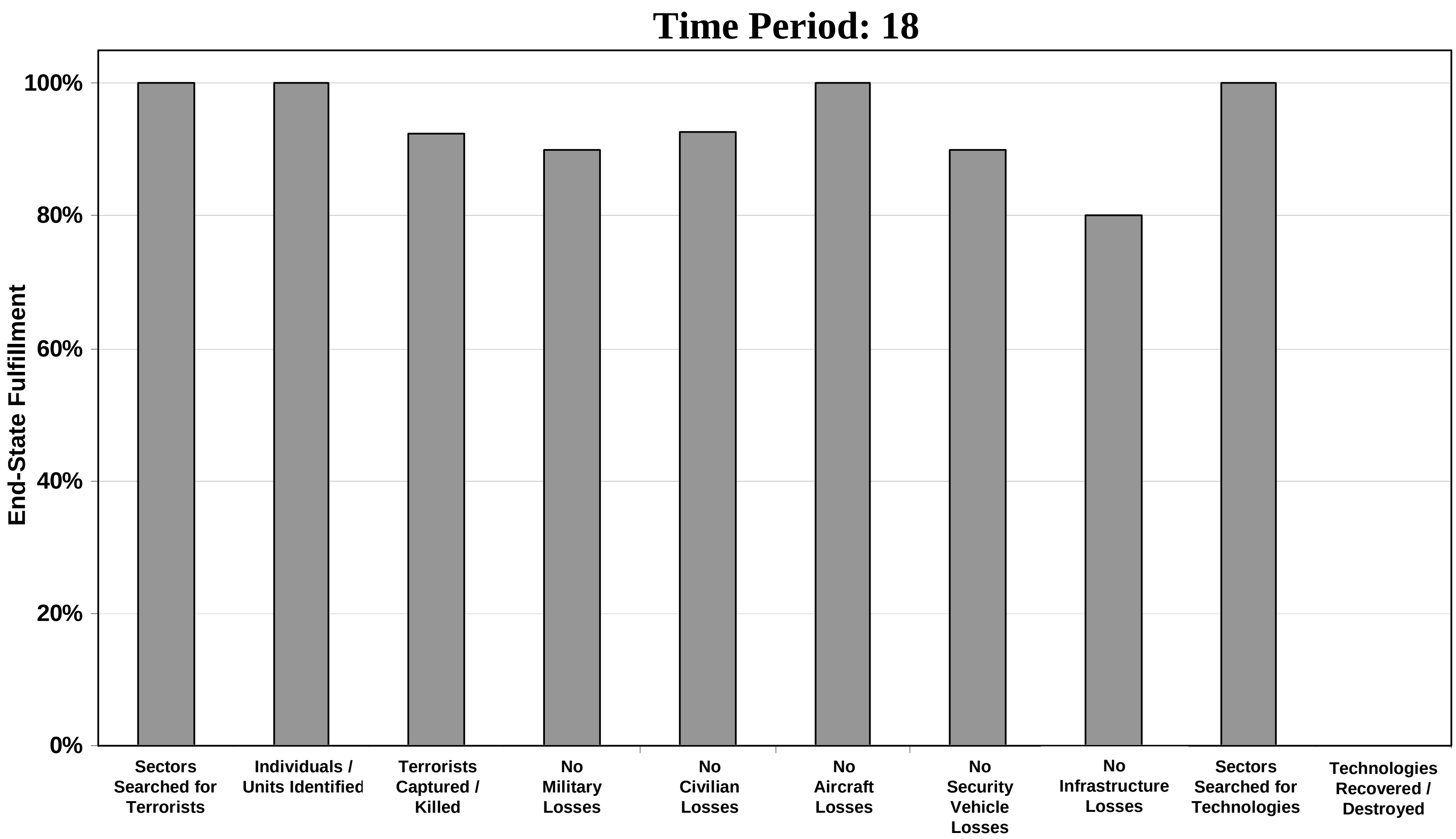

Time Period: 14
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed


Time Period: 15
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed

Time Period: 16

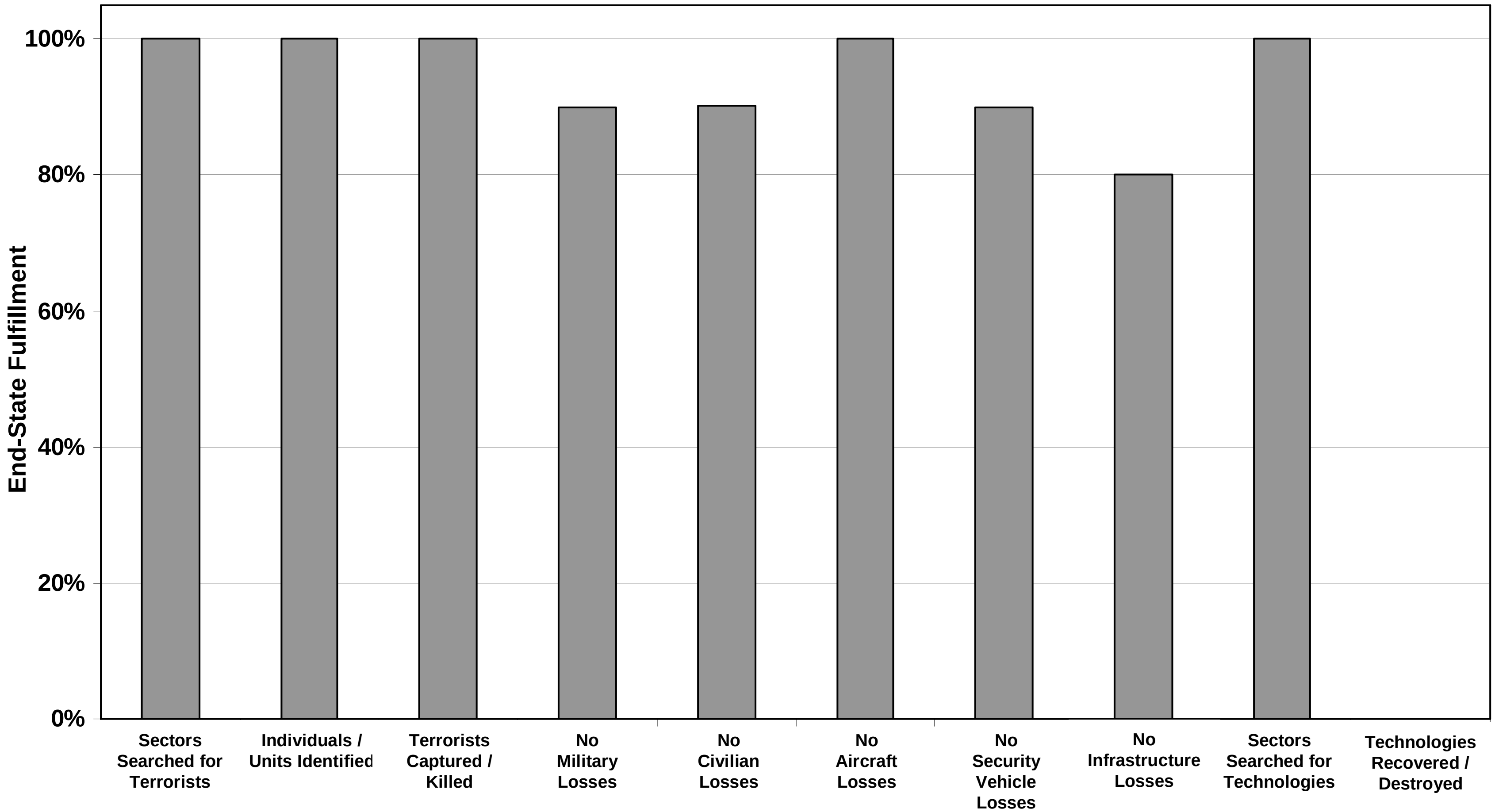


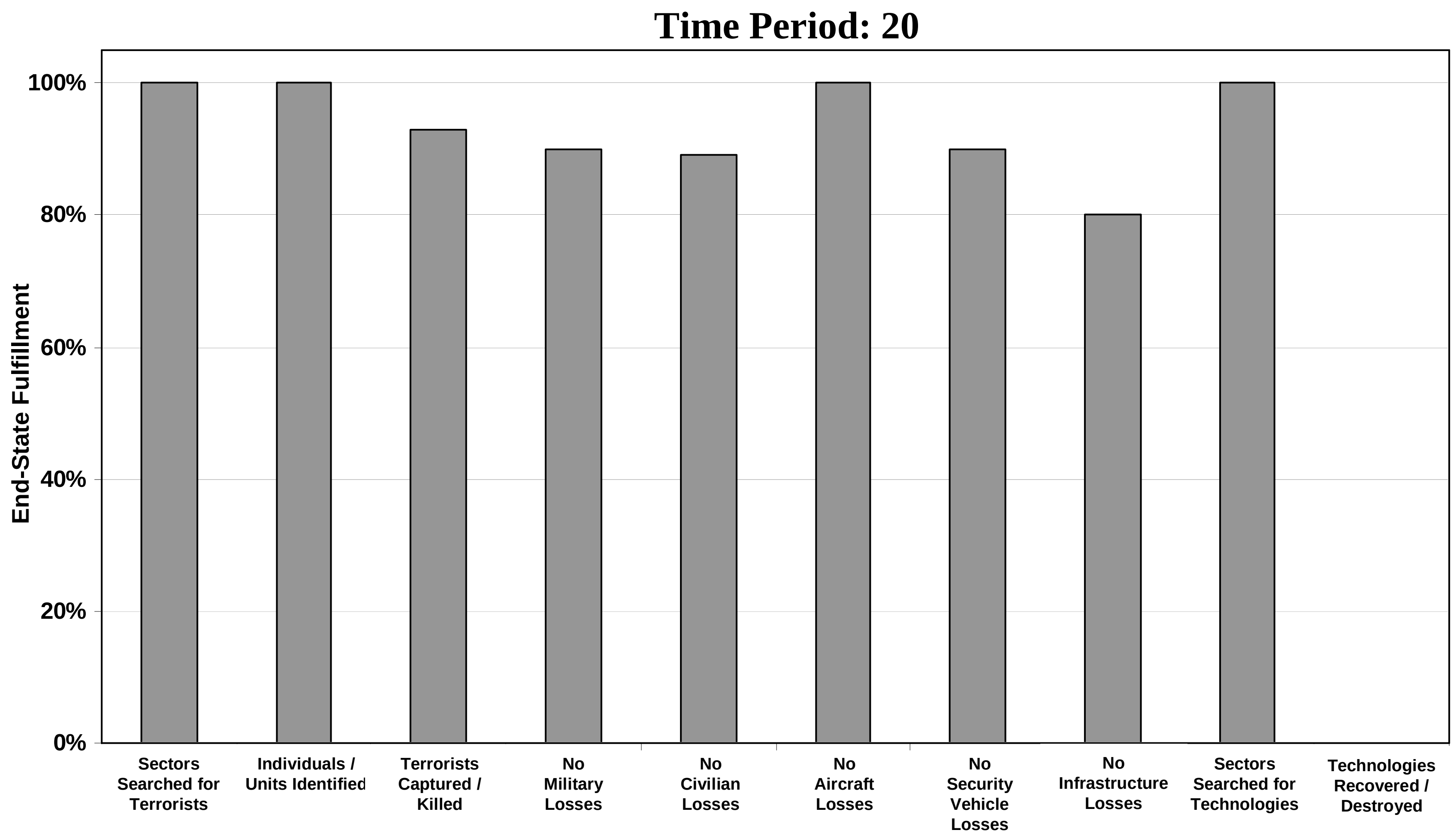

Time Period: 17
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed



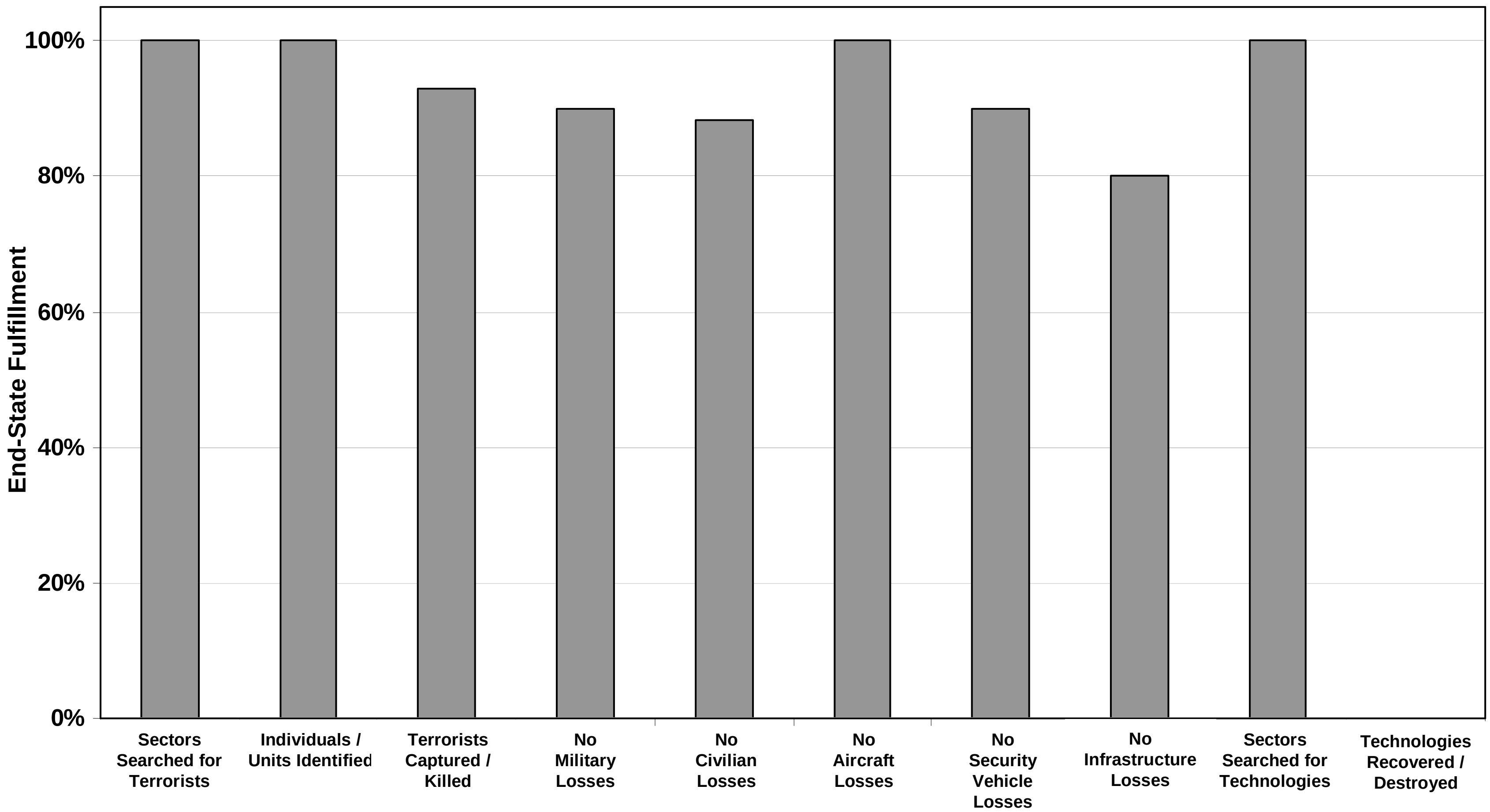
Time Period: 18
100%
80%
60%
40%
20%
0%
End-State Fulfillment
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed

Time Period: 19

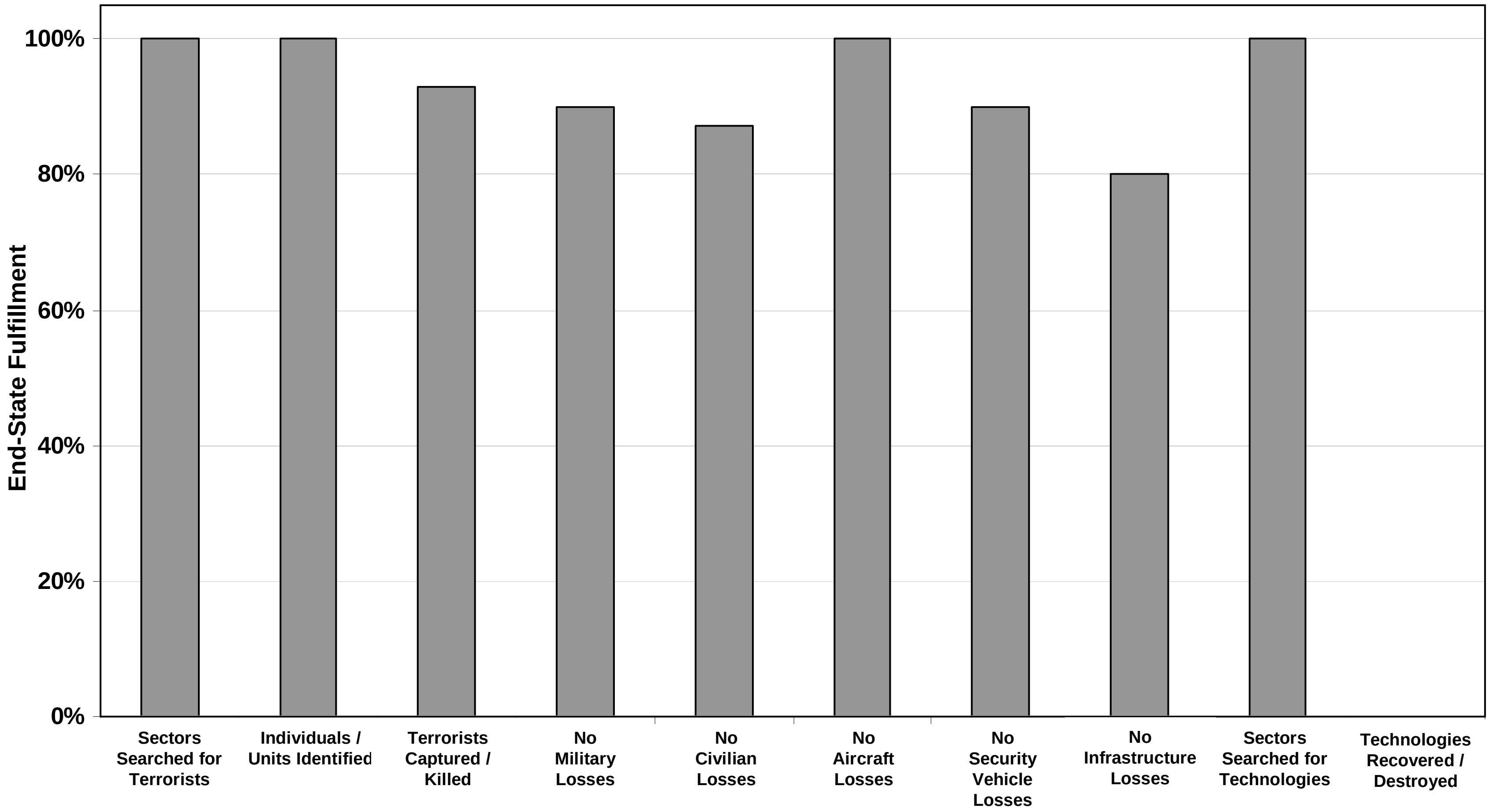


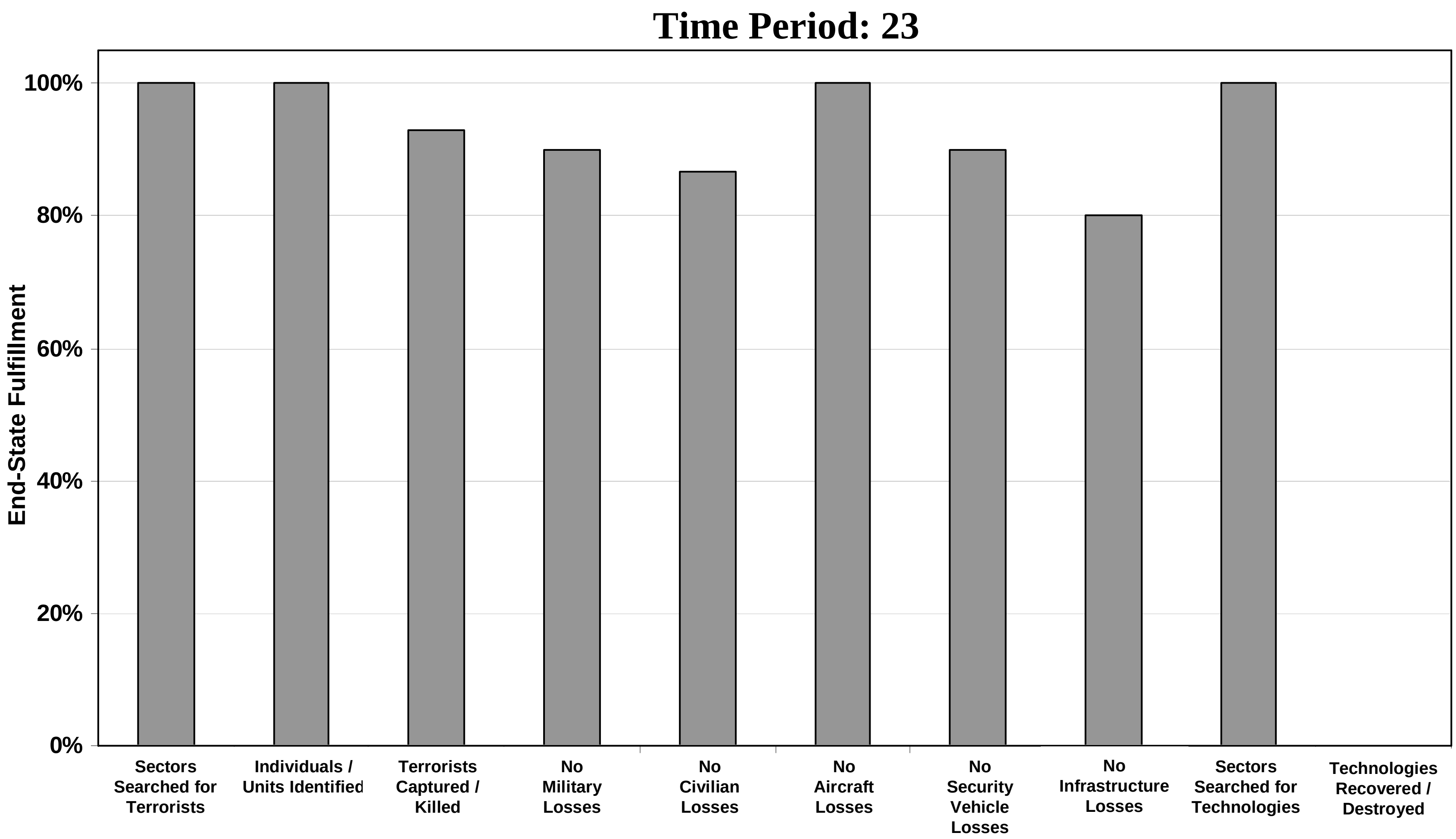

Time Period: 20
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed

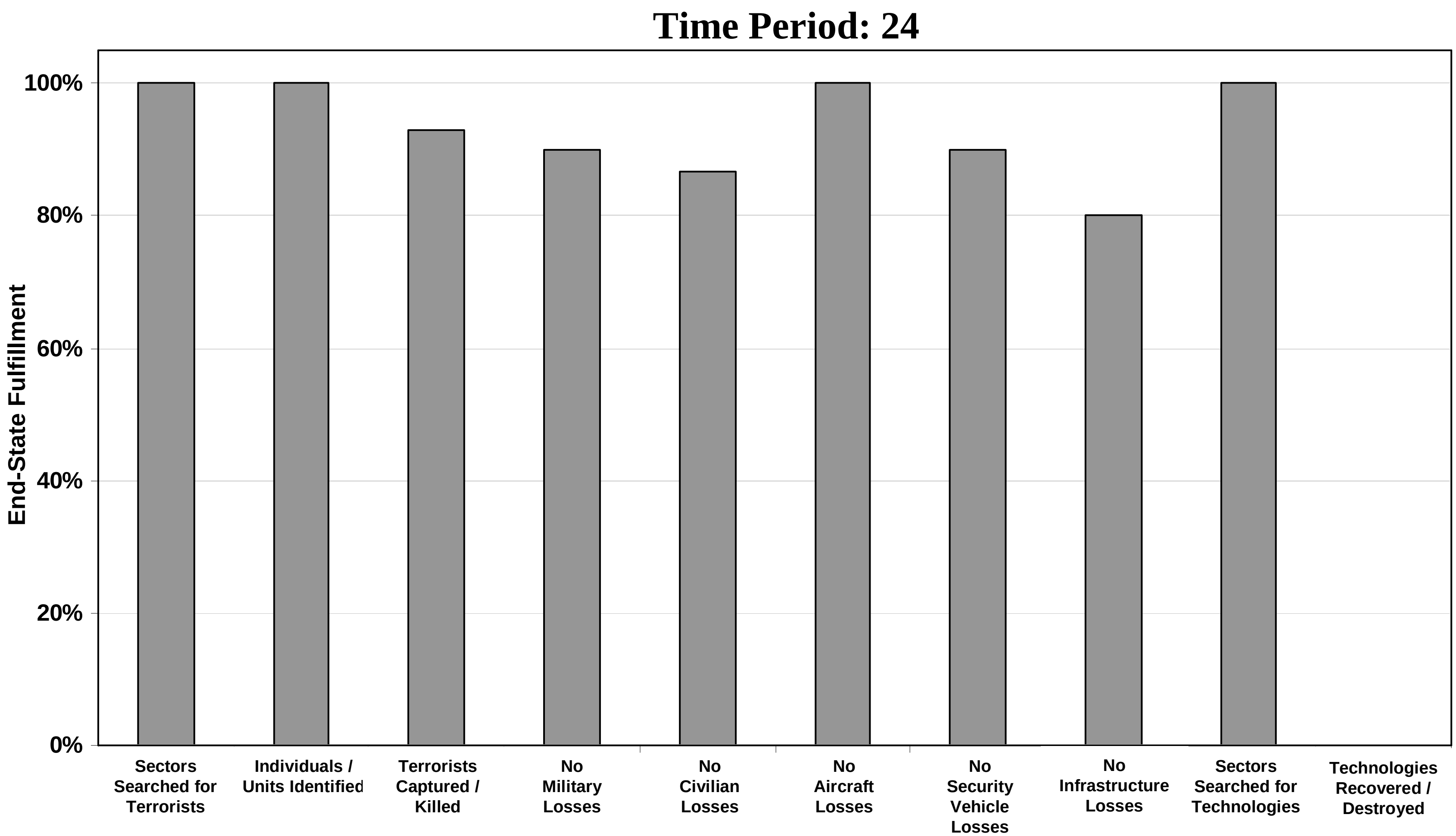
Time Period: 21

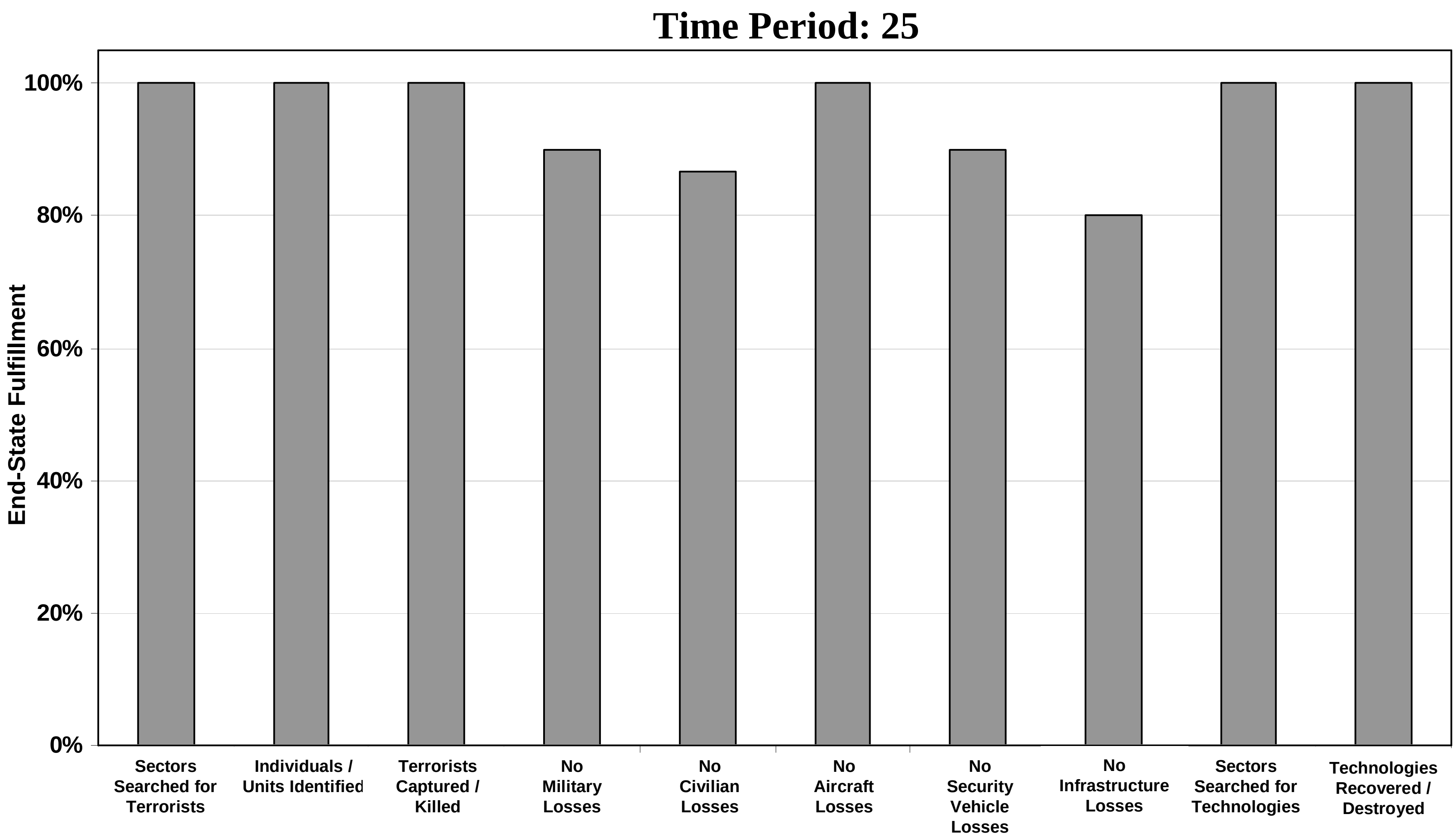


Time Period: 22




Time Period: 23
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed


Time Period: 24
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed


Time Period: 25
End-State Fulfillment
100%
80%
60%
40%
20%
0%
Sectors Searched for Terrorists
Individuals / Units Identified
Terrorists Captured / Killed
No Military Losses
No Civilian Losses
No Aircraft Losses
No Security Vehicle Losses
No Infrastructure Losses
Sectors Searched for Technologies
Technologies Recovered / Destroyed

REFERENCES

Ackoff, Russell L., "Towards a System of Systems Concepts," *Management Science*, Vol. 17, No. 11, pp 661-671, 1971.

Allen, Mary J. and Wendy M. Yen, Introduction to Measurement Theory, Waveland Press, Prospect Heights, IL, 1979.

Antony, Patrick, Jennifer Dunn, William Farr, Donna Rhodes, Garry Roedler, Cathy Tilton, and Richard Widmann, Systems Engineering Measurement Primer, International Council on Systems Engineering, 1998.

Apostol, Tom M., Mathematical Analysis, 2nd ed., Addison, Wesley, Longman Publishing, Menlo Park, CA, 1974.

Artley, Will and Suzanne Stroh, Establishing an Integrated Performance Measurement System, The Performance-Based Management Handbook, Volume 2, Performance-Based Management Special Interest Group, U.S. Department of Energy, 2001.

Artley, Will, DJ Ellison, and Bill Kennedy, Establishing and Maintaining a Performance-Based Management Program, The Performance-Based Management Handbook, Volume 1, Performance-Based Management Special Interest Group, U.S. Department of Energy, 2001.

Avis D., K. Fukuda and S. Picozzi, "On Canonical Representations of Convex Polyhedra", *Proceedings of the First International Congress of Mathematical Software*, pp 350-360, 2002.

Balkcom, John E., Christopher D. Ittner, and David F. Larcker, "Strategic Performance Measurement: Lessons Learned and Future Directions," *Journal of Strategic Performance Measurement*, pp 22-32, April/May 1997.

Beckerman, Linda P., "Application of Complex System Science to Systems Engineering," *Systems Engineering*, Vol. 3, No. 2, pp 96-102, 2000.

Beebe, R. P., Military Decision from the Viewpoint of Game Theory, Naval War College, Newport, RI, 1957.

Benoit, E. and L. Foulloy, "Towards Fuzzy Nominal Scales," *Measurement*, Vol. 34, pp 49-55, 2003.

Blanchard, Benjamin S., System Engineering Management, John Wiley & Sons, New York, 1991.

Bouthonnier, Vincent and Alexander H. Levis, "Effectiveness Analysis of C^3 Systems," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-14, No. 1, pp 48-54, 1984.

Bowman, Christopher W., Operational Assessment—The Achilles Heel Of Effects-Based Operations?, Naval War College, Newport, RI, May 2002.

Brakel, J. van, "Norms and Facts in Measurement," *Measurement*, Vol. 2, No. 1, pp 45-51, Jan – Mar 1984.

Briand, Lionel C., Sandro Morasca, and Victor R. Basili, "Property-Based Software Engineering Measurement," *IEEE Transactions on Software Engineering*, Vol. 22, No. 1, pp 68-86, January 1996.

Brown, Mark G., Keeping Score: Using the Right Metrics to Drive World-Class Performance, Productivity Press, New York, 1996.

Bueno De Mesquita, Bruce and David Lalman, War and Reason, Yale Univeristy Press, New Haven, 1992.

Bueno de Mesquita, Bruce, "The War Trap Revisited: A Revised Expected Utility Model," *American Political Science Review*, Vol. 79, pp 156-177, 1985.

Bueno de Mesquita, Bruce, The War Trap, Yale University Press, New Haven, 1981.

Bullock, Richard K., Hierarchical Interactive Theater Model (HITM): An Investigation into the Relationship Between Strategic Effects and OODA Loops, Masters Thesis, Air Force Institute of Technology, 2000.

Bulmer, Martin, "Social Measurement: What Stands in Its Way?," *Social Research*, Vol. 68, No. 2, pp 455-480, Summer, 2001.

Burke, Eve M., Quality Function Deployment from an Operations Research Perspective, Masters Thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, 1999.

Burke, Eve, Jack M. Kloeber, Jr., and Richard F. Deckro, "Using and Abusing QFD Scores", *Quality Engineering*, Vol. 15, No. 1, pp 9-21, 2003.

Bush, Robert R., Eugene Galanter, and R. Duncan Luce, "Characterization and Classification of Choice Experiments," Handbook of Mathematical Psychology: Volume 1, R. Duncan Luce, Robert R. Bush, and Eugene Galanter eds., John Wiley & Sons, New York, pp 77-102, 1963.

Byrnes, John, “Background Paper on System Dynamics as a Tool for Effects Based Operations (EBO) Analysis,” Analyzing Effects-based Operations Workshop, McLean, VA, 2002.

Campbell, Norman R., Foundations of Science: The Philosophy of Theory and Experiment, Dover Publications, New York, 1957 (original work published in 1920).

Campi, John P., “Breaking the Paradigm of Performance Measurement,” Handbook for Productivity Measurement and Improvement, William F. Christopher and Carl G. Thor eds., Productivity Press, Portland, OR, 1993.

Cantwell, Gregory L., Can Two Person Zero Sum Game Theory Improve Military Decision-Making Course Of Action Selection?, United States Army Command and General Staff College, Fort Leavenworth, KS, 2003.

Caws, Peter, “Definition and Measurement in Physics,” Measurement: Definitions and Theories, C. West Churchman and Philburn Ratoosh eds., John Wiley & Sons, New York, pp 3-17, 1959.

Champagne, Lance E., Development Approaches Coupled with Verification and Validation Methodologies for Agent-based Mission-level Analytical Combat Simulations, Dissertation, Air Force Institute of Technology, OH, 2003.

Cioffi-Revilla, Claudio, “Mathematical Contributions to the Scientific Understanding of War,” *Mathematical Computer Modeling*, Vol. 12, No. 4/5, pp 561-575, 1989.

Clausewitz, Carl von, On War, Michael Howard and Peter Paret eds/trans., Princeton University Press, Princeton, NJ, 1976.

Clemen, Robert T., Making Hard Decisions: An Introduction to Decision Analysis, 2nd ed, Duxbury Press, Pacific Grove, CA, 1996.

Cohn, Donlad L., Measure Theory, Birkhauser, Boston, 1980.

Cristianini, Nello and John Shawe-Taylor, An Introduction to Support Vector Machines, Cambridge University Press, Cambridge, 2000.

Cropley, D. H., “Towards Formulating a Semiotic Theory of Measurement Information – Part 1: Fundamental Concepts and Measurement Theory,” *Measurement*, Vol. 24, pp 237-248, 1998.

Cyranski, John F., “Measurement, Theory, and Information,” *Information and Control*, Vol. 41, pp 275-304, 1979.

Da Rocha, Alexandre Sergio, "Effects-Based Operations: A Military Application of Pragmatical Analysis," *Air & Space Power Journal*, Maxwell Air Force Base, AL, Fall, 2005.

Davis, Paul K., Effects-Based Operations. A Grand Challenge for the Analytical Community, RAND Corp, Santa Monica, CA, 2001.

Deckro, Richard F., and John E. Hebert, "Polynomial Goal Programming: A Procedure for Modeling Preference Trade-offs," *Journal of Operations Management*, Vol. 7, No. 4, pp 149-164, 1988.

De Toni, A. and S. Tonchia, "Performance Measurement Systems: Models, Characteristics, and Measures", *International Journal of Operations & Production Management*, Vol. 21, No. 1/2, pp 46-70, 2001.

Defense Systems Management College (DSMC), "Systems Acquisition Manager's Guide for the Use of Models and Simulation," September 1994.

DeGregorio, Edward A., "Integrating Effects-based and Attrition-based Modeling," *2004 Command and Control Research and Technology Symposium*, San Diego, CA, 2004.

Deptula, David A., "Firing for Effects," *AIR FORCE Magazine*, pp 46-53, April 2001.

Deptula, David A., Effects-based Operations: Change in the Nature of Warfare, Aerospace Education Foundation, Arlington, VA, 2001.

Dillon, William R. and Matthew Goldstein, Multivariate Analysis, John Wiley & Sons, New York, 1984.

Doyle, Michael P., A Value Function Approach to Information Operations MOE's: A Preliminary Study, CMSA Technical Report 97-04, Air Force Institute of Technology, OH, 1997.

Dresher, Melvin, Games of Strategy: Theory and Applications, Prentice-Hall, Englewood Cliffs, NJ, 1961.

Eccles, Robert, "The Performance Measurement Manifesto," *Harvard Business Review*, Vol. 69, No. 1, pp 131-138, 1991.

Evans, James R., "An Exploratory Study of Performance Measurement Systems and Relationships with Performance Results," *Journal of Operations Management*, Vol. 22, pp 219-232, 2004.

Fenton, Norman, "Software Measurement: A Necessary Scientific Basis," *IEEE Transactions on Software Engineering*, Vol. 20, No. 3, pp 199-206, March, 1994.

Feuchter, Christopher A., Air Force Analyst's Handbook: On Understanding the Nature of Analysis, Office of Aerospace Studies, Kirtland Air Force Base, NM, January 2000.

Finkelstein, L. and M. S. Leaning, "A Review of the Fundamental Concepts of Measurement," *Measurement*, Vol. 2, No. 1, pp 25-34, Jan-Mar 1984.

Finkelstein, Ludwik, "Widely, Strongly and Weakly Defined Measurement," *Measurement*, Vol. 34, pp 39-48, 2003.

Ford, Gary, "Lecture Notes on Engineering Measurement for Software Engineers," Carnegie Mellon University, CMU/SEI-93-EM-9, 1993.

Forrester, Jay W., "Dynamical Models of Economic Systems and Industrial Organizations," *Systems Dynamics Review*, Vol. 19, pp 331-345, 2003.

Friedman, Avner, Differential Games, John Wiley & Sons, New York, 1971.

Fudenberg, Drew and Jean Tirole, Game Theory, MIT Press, Cambridge, 1993.

Gallagher, Mark A., Wesley D. True, and Philip S. Whiteman, "Precisely Defining Effects for EBO," *PHALANX*, Vol. 37, No.4, pp 9-15, December, 2004.

Gartner, Scott Sigmond, Strategic Assessment in War, Yale University Press, New Haven, 1997.

Geisler, Eliezer, The Metrics of Science and Technology, Quorum Books, Westport, CT, 2000.

Geller, Daniel S. and J. David Singer, Nations at War: A Scientific Study of International Conflict, Cambridge University Pres, Cambridge, 1998.

Gill, Roderic, "An Integrated Social Fabric Matrix/System Dynamics Approach to Policy Analysis," *System Dynamics Review*, Vol. 12, No. 3, pp 167-181, Fall 1996.

Gillespie, John V., "Why Mathematical Models?", Mathematical Models in International Conflict, Dina A. Zinnes and John V. Gillespie eds., Praeger Publishers, New York, 1976.

Glenn, Kevin B., "The Challenge of Assessing Effects-Based Operations in Air War," *Air & Space Power Chronicles*, (online), April 2002.

Grassmann, Winfried Karl and Jean-Paul Tremblay, Logic and Discrete Mathematics: A Computer Science Perspective, Prentice Hall, Upper Saddle River, NJ, 1996.

Grewal, Mohinder S. and Angus P. Andrews, Kalman Filtering: Theory and Practice, Prentice Hall, Englewood Cliffs, New Jersey, 1993.

Hamilton, Thomas and Richard Mesic, A Simple Game-theoretic Approach to Suppression of Enemy Defenses and other time Critical Target Analysis, RAND Project Air Force, DB-385-AF, CA, 2004.

Hamner, Kenneth L. and Charles A. La Fleur, An Exploratory Survey of Methods Used to Develop Measures of Performance, Masters Thesis, Air Force Institute of Technology, Wright-Patterson Air Force Base, OH, 1993.

Harbour, Jerry L., The Basics of Performance Measurement, Quality Resources, New York, 1997.

Haywood, O.G., "Military Decision and Game Theory," *Journal of the Operations Research Society of America*, Vol. 2, No. 4, pp 365-385, 1954.

Henningsen, Jacqueline R. and Dean Cash, Analyzing Effects-Based Operations, Workshop Report, Military Operations Research Society, January 2003.

Hill, Raymond, Gregory A. McIntyre, Thomas R. Tighe, and Richard K. Bullock, "Some Experiments with Agent-based Combat Models," *Military Operations Research*, Vol. 8, No. 3, pp 17-28, 2003.

Ho, Joshua, "The Advent of a New Way of War: Theory and Practice of Effects Based Operations," Working Paper, Institute of Defence and Strategic Studies, Singapore, 2003.

HPS Simulations (HPS), Point of Attack 2, www.hpssims.com, Santa Clara, CA, 2006.

Ilachinski, Andrew, Irreducible Semi-Autonomous Adaptive Combat (ISAAC): An Artificial-life Approach to Land Warfare, CRM 97-61.10, Center for Naval Analyses, 1997.

Isaacs, Rufus, Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization, John Wiley & Sons, New York, 1965.

Ittner, Christopher D. and David F. Larcker, "Innovations in Performance Measurement: Trends and Research Implications," *Journal of Management Accounting Research*, Vol. 10, pp 205-238, 1998.

Jordan, Gretchen, Steven Prevette, and Steve Woodward, Analyzing, Reviewing, and Reporting Performance Data, The Performance-Based Management Handbook, Volume 5, Performance-Based Management Special Interest Group, U.S. Department of Energy, 2001.

Kalman, Rudolf E., "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the American Society of Mechanical Engineers, Journal of Basic Engineering*, Vol. 82 (Series D), pp 35-45, 1960.

Kaplan, Edward L., Mathematical Programming and Games, John Wiley & Sons, New York, 1982.

Kaplan, R. and D. Norton, "The balanced scorecard - measures that drive performance," *Harvard Business Review*, Vol. 70, No. 1, pp 71-79, 1991.

Kaplan, R. and D. Norton, "Using the Balanced Scorecard as a Strategic Management System," *Harvard Business Review*, Vol. 74, No. 1, pp 75-85, 1996.

Kaydos, Will, Operational Performance Measurement: Increasing Total Productivity, St. Lucie Press, Boca Raton, 1999.

Keeney, Ralph L., Value Focused Thinking: A Path to Creative Decision Making, Harvard University Press, Cambridge, 1992.

Keeney, Ralph L. and Howard Raiffa, Decisions with Mutliple Objectives: Preference and Value Trade-offs, Cambridge University Press, 1993.

Khalil, Hassan K., Nonlinear Systems, 3rd Edition. Prentice-Hall, Upper Saddle River, NJ, 2002.

Kircher, Paul, "Measurements and Managerial Decisions," Measurement: Definitions and Theories, C. West Churchman and Philburn Ratoosh eds., John Wiley & Sons, New York, pp 18-63, 1959.

Kirkwood, Craig W., Strategic Decision Making, Wadsworth Publishing, 1997.

Kitchenham, B., S.L. Pfleeger, and N.E. Fenton, "Towards a Framework for Software Measurement Validation", *IEEE Transactions on Software Engineering*, Vol. 21, No. 12, pp 929-944, 1995.

Krantz, David, R. Duncan Luce, Patrick Suppes, and Amos Tversky, Foundations of Measurement: Additive and Polynomial Representations, Vol. 1, Academic Press, New York, 1971.

Larimer, Larry and Pedro Habic, "A Means of Measuring Mission Accomplishment as an MOE", *72nd Military Operations Research Society Symposium*, Montery, CA, June 2004.

Lazarus, David B., "Effects-Based Operations and Counterterrorism," *Air & Space Power Journal*, Maxwell Air Force Base, AL, Fall, 2005.

Lebas, Michel and Ken Euske, "A Conceptual and Operational Delineation of Performance," Business Performance Measurement: Theory and Practice, A. Neely ed., Cambridge University Press, Cambridge, 2002.

Lebas, Michel J., "Performance Measurement and Performance Management," *International Journal of Production Economics*, Vol. 41, pp 23-45, 1995.

Leonard, Marcia, Air Force Material Command: A Survey of Performance Measures, Masters Thesis, Air Force Institute of Technology, Wright-Patterson Air Force Base, OH, 2004.

Levis, Alexander H., "Effects-based Operations", AIAA-IC2S Presentation, December 2001.

Love, Robert F., James G. Morris, and George O. Wesolowsky, Facilities Location: Models & Methods, North-Holland, New York, 1988.

Lowe, Donald and Simon Ng, "Effects-based operations: language, meaning and the effects-based approach", *2004 Command and Control Research and Technology Symposium*, 2004.

Luce, R. Duncan and L. Narens, "Classification of Real Measurement Representations by Scale Type," *Measurement*, Vol. 2, No. 1, pp 39-44, Jan-Mar 1984.

Luce, R. Duncan and Howard Raiffa, Games and Decisions: Introduction and Critical Survey, Dover Publications, New York, 1957.

Luce, R. Duncan and John W. Tukey, "Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement," *Journal of Mathematical Psychology*, Vol. 1, pp 1-27, 1964.

Mann, Edward C., Gary Endersby, and Thomas R. Searle, Thinking Effects Effects-Based Methodology for Joint Operations, Air University Press, Maxwell Air Force Base, AL, October 2002.

Maoz, Zeev, "Decision-Theoretic and Game-Theoretic Models of International Conflict", Dynamic Models of International Conflict, Urs Luterbacher and Michael Ward eds., Lynne Rienner Publishers, Boulder, CO, 1985.

Margenau, Henry, "Philosophical Problems Concerning the Meaning of Measurement in Physics," Measurement: Definitions and Theories, C. West Churchman and Philburn Ratoosh eds., John Wiley & Sons, New York, pp 163-176, 1959.

Mari, Luca, “The Meaning of ‘Quantity’ in Measurement,” *Measurement*, Vol. 17, No. 2, pp 127-138, 1996.

Marlow, W. H., Mathematics for Operations Research, Dover Publications, New York, 1978.

Maybeck, Peter S., Stochastic Models, Estimation, and Control: Volume 1, Academic Press, New York, 1979.

Maybeck, Peter S., Stochastic Models, Estimation, and Control: Volume 2, Academic Press, New York, 1982.

McCrabb, Maris, “Effects-based Coalition Operations: Belief, Framing and Mechanism,” *Second International Conference on Knowledge Systems for Coalition Operations*, Tolouse, France, April 2002.

McCrabb, Maris, Explaining "Effects": A Theory for an Effects-based Approach To Planning, Executing and Assessing Operations, Manuscript, DMM Ventures, August 2001.

Meilinger, Phillip S., “Air Strategy: Targeting For Effect,” *Aerospace Power Journal*, pp 48-61, Winter 1999.

Melnyk, Steven A., Douglas M. Stewart, and Morgan Swink, “Metrics And Performance Measurement In Operations Management: Dealing With The Metrics Maze,” *Journal of Operations Management*, Vol. 22, pp 209–217, 2004.

Meyer, Christopher, “How the Right Measures Help Teams Excel,” *Harvard Business Review*, Vol. 72, No. 3, pp 95-97, May/June 1994.

Meyer, Marshall W., “Dilemmas of Performance Measurement,” *Journal of Strategic Performance Measurement*, pp 33-42, April/May 1997.

Meyer, Marshall W., Rethinking Performance Measurement: Beyond the Balanced Scorecard, Cambridge University Press, Cambridge, 2002.

Mitchell, Joel, “Measurement: A Beginner’s Guide,” *Journal of Applied Measurement*, Vol. 4, No. 4, pp 298-308, 2003.

Mitchell, Tom M., Machine Learning, The McGraw-Hill Companies, New York, 1997.

Morse, Philip M., and George E. Kimball, Methods of Operations Research, Dover Publications, Mineola, NY, 2003.

Murray, William S., “A Will To Measure,” *Parameters*, pp 134-47, Autumn 2001.

Myerson, Roger B., Game Theory: Analysis of Conflict, Harvard University Press, Cambridge, 1991.

Narens, Louis and R. Duncan Luce, “Measurement: The Theory of Numerical Assignment,” *Psychological Bulletin*, Vol. 99, No. 2, pp 166-180, 1986.

Narens, Louis and R. Duncan Luce, “The Algebra of Measurement,” *Journal of Pure and Applied Algebra*, Vol. 8, pp 197-233, 1976.

Narens, Louis, “A Meaningful Justification for the Representational Theory of Measurement,” *Journal of Mathematical Psychology*, Vol. 46, pp 746-768, 2002.

Narens, Louis, Abstract Measurement Theory, MIT Press, Cambridge, MA, 1985.

Nash, John F., Non-Cooperative Games, Dissertation, Princeton University, 1950.

Nash, Stephen G. and Ariela Sofer, Linear and Nonlinear Programming, McGraw Hill, 1996.

Neely, Andy, Huw Richards, John Mills, Ken Platts, and Mike Bourne, “Designing Performance Measures: A structured Approach,” *International Journal of Operations & Production Management*, Vol. 17, No. 11, pp 1131-1152, 1997.

Neumann, John von and Oskar Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, Princeton, 1944.

Niederée, Reinhard, “What do Numbers Measure? A new Approach to Fundamental Measurement,” *Mathematical Social Sciences*, Vol. 24, pp 237-276, 1992.

Owen, Guillermo, Game Theory, W.B. Saunders Company, Philadelphia, 1968.

Park, Robert E., Wolfhart B. Goethert, and William A. Florac, Goal-Driven Software Meaurement—A Guidebook, Hankbook (CME/SEI-96-HB-002), Carnegie Mellon University, 1996.

Perret, Geoffrey, There's a War to Be Won: The United States Army in World War II, Random House, New York, 1991.

Pfanzagl, Johann, Theory of Measurement, Physica-Verlag, Wurzburg, 1971.

Pinker, Aron, Aryeh H. Samuel, and Robert Batcher, “On Measures of Effectiveness,” *PHALANX*, pp 8-12, December 1995.

Poels, G. and G. Dedene, “Formal Software Measurement for Object-Oriented Business Models,” *Proceedings of the 7th European Software Control and Metrics Conference (ESCOM'96)*, Vol. 1, Wilmslow, UK, May 15-17, pp 115-134, 1996.

Poels, G. and G. Dedene, “Distance-based software measurement: necessary and sufficient properties for software measures,” *Information and Software Technology*, Vol. 42, pp 35–46, 2000.

Potter, Ronald W., The Art of Measurement: Theory and Practice, Prentice Hall, Upper Saddle River, NJ, 2000.

Rapoport, Anatol, “Mathematical Methods in Theories of International Relations: Expectations, Caveats, and Opportunities”, Mathematical Models in International Conflict, Dina A. Zinnes and John V. Gillespie eds., Praeger Publishers, New York, 1976.

Roberts, F. S., “On the Theory of Meaningfulness of Ordinal Comparisons in Measurement,” *Measurement*, Vol. 2, No.1, pp 35-38, Jan-Mar 1984.

Roberts, Fred S., Encyclopedia of Mathematics and its Applications Vol. 7: Measurement Theory with Applications to Decisionmaking, Utility, and the Social Sciences, Gian-Carlo Rota ed., Addison-Wesley Publishing, Reading, MA, 1979.

Roche, James G. and Barry D. Watts, “Choosing Analytic Measures,” *The Journal of Strategic Studies*, Vol. 14, pp 165-209, June 1991.

Rosen, Julie A. and Wayne L. Smith, “Influence Net Modeling with Causal Strengths: An Evolutionary Approach”, *Proceedings of the Command and Control Research and Technology Symposium*, Naval Post Graduate School, Monterey CA, June 25-28, 1996.

Ruckle, William H., Modern Analysis, PWS-KENT Publishing, Boston, 1991.

Rumsey, Kirk H., Developing Departmental Measures of Performance in a White-collar Environment, Maters Thesis, Air Force Institute of Technology, Wright-Patterson Air Force Base, OH, 1990.

Russell, S. and P. Norvig, Artificial Intelligence, 2nd ed., Prentice Hall, Upper Saddle River, NJ, 2003.

Sarle, Warren S., “Measurement theory: Frequently asked questions,” *Disseminations of the International Statistical Applications Institute*, 4th edition, Wichita: ACG Press, pp 61-66, 1995.

Schelling, Thomas C., Arms and Influence, New Haven and London: Yale University Press, 1966.

Schmitt, John F., “Command and (Out of) Control: The Military Implications of Complexity Theory,” Complexity, Global Politics, and National Security, David S. Alberts and Thomas J. Czerwinski eds., National Defense University, 1999.

Schwager, K. Walter, “The Representational Theory of Measurement: An Assessment,” *Psychological Bulletin*, Vol. 110, No. 3, pp 618-626, 1991.

Scott, Dana and Patrick Suppes, “Foundational Aspects of Theories of Measurement,” *The Journal of Symbolic Logic*, Vol. 23, No. 2, pp 113-128, June 1958.

Shubik, M., “Game Theory: The Language of Strategy?,” Mathematics of Conflict, M. Shubik ed., North-Holland, Amsterdam, 1983.

Sink, D. Scott, “The Role of Measurement in Achieving World Class Quality and Productivity Management,” *Industrial Engineering*, Vol. 23, No. 6, pp 23-28, June 1991.

Sink, D. Scott, Productivity Management: Planning, Measurement and Evaluation, Control and Improvement, John Wiley & Sons, New York, 1985.

Smith, Edward A., Effects Based Operations: Applying Network Centric Warfare in Peace, Crisis, and War, DoD Command and Control Research Program, November 2002.

Snodgrass, Anthony W., Leontief Metamodeling for Military Strategic Effects, Masters Thesis, Air Force Institute of Technology, 2000.

Snodgrass, Anthony W., Mark A. Gallagher, and Gregory A. McIntyre, “Modeling Military Strategic Effects with an Input-Output Metamodel,” *Military Operations Research*, Vol. 9, No. 1, pp 19-32, 2004.

Sproles, Noel, “Coming to Grips with Measures of Effectiveness,” *Systems Engineering*, Vol. 3, pp 50-58, 2000.

Sproles, Noel, “Formulating Measures of Effectiveness,” *Systems Engineering*, Vol. 5, pp 253–263, 2002.

Sproles, Noel, “Identifying Success Through Measures,” *PHALANX*, Vol. 30, No. 4, pp 16-31, December 1997.

Sproles, Noel, “Meaning, Mapping, Measures, and Test & Evaluation,” *ITEA Journal*, pp 34-39, September/October 1996.

Stevens, S. S., "Measurement, Psychophysics, and Utility," Measurement: Definitions and Theories, C. West Churchman and Philburn Ratoosh eds., John Wiley & Sons, New York, pp 18-63, 1959.

Stevens, S. S., "On the Theory of Scales of Measurement," *Science*, Vol. 103, No. 2684, pp 677-680, June 1946.

Suppes, Patrick, David M. Krantz, R. Duncan Luce, and Amos Tversky, Foundations of Measurement: Geometrical, Threshold, and Probabilistic Representations, Vol. 2, Academic Press, San Diego, 1989.

Suppes, Patrick and Joseph L. Zinnes, "Basic Measurement Theory," Handbook of Mathematical Psychology: Volume 1, R. Duncan Luce, Robert R. Bush, and Eugene Galanter eds., John Wiley & Sons, New York, pp 1-76, 1963.

Sydenham, Peter H., "Relationship between Measurement, Knowledge, and Advancement," *Measurement*, Vol. 34, pp 3-16, 2003.

Tamiz, Mehrdad, Dylan Jones, and Carlos Romero, "Goal Programming for Decision Making: An overview of the current state-of-the-art," *European Journal of Operations Research*, Vol. 111, pp 569-581, 1998.

Timmerman, Thomas J., Assessing Effects-Based Operations Crucial Feedback For The Air Component Commander, Air Command And Staff College, Maxwell Air Force Base, Al, April 2003.

Torgerson, Warren S., Theory and Methods of Scaling, John Wiley & Sons, New York, 1958.

Tufte, Edward R., The Visual Display of Quantitative Information, Graphics Press, Cheshire, Connecticut, 1983.

Tufte, Edward R., Visual Explanations: Images and Quantities, Evidence and Narrative, Graphics Press, Cheshire, Connecticut, 1997.

United States Air Force (USAF), Air Force Doctrine Document 1, HQ Air Force Doctrine Center, September 1997.

United States Air Force (USAF), Air Force Doctrine Document 2-1.2, HQ Air Force Doctrine Center, September 2003.

United States Air Force (USAF), Air Force Doctrine Document 2-5, HQ Air Force Doctrine Center, January 2005.

United States Air Force Office of Scientific Research (AFOSR), Small Business Technology Transfer Contract: F49620-99-C-0061, Arlington, VA, 2001.

United States Joint Forces Command (USJFC), Doctrinal Implications of the Standing Joint Force Headquarters (SJFHQ), The Joint Warfighting Center Joint Doctrine Series Pamphlet 3, 2003.

United States Joint Forces Command (USJFC), Standing Joint Force Headquarters Prototype, Concept of Employment, June 2003.

United States Joint Forces Command (USJFC), Commander's Handbook for an Effects-Based Approach to Joint Operations, Joint Warfighting Center, Joint Concept Development and Experimentation Directorate, Standing Joint Force Headquarters, February 2006.

Wacker, John G., "A Theory of Formal Conceptual Definitions: Developing Theory-Building Measurement Instruments," *Journal of Operations Management*, Vol. 22, pp 629-650, 2004.

Wang, Yingxu, "The Measurement Theory for Software Engineering," *Proceedings from the 2003 Canadian Conference on Electrical and Computer Engineering*, Montreal, pp 1321-1324, May 2003.

Warden, John A., "The Enemy as a System," *Airpower Journal*, pp 40-55, Spring 1995.

Weise, K. and W. Woger, "A Bayesian Theory of Measurement Uncertainty," *Measurement Science Technology*, Vol. 3, pp 1-11, 1992.

Welch, Greg and Gary Bishop, An Introduction to the Kalman Filter, TR 95-041, Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, April, 2004.

Whittemore, David M., Modeling Strategic Effects In Wargames, Air Command And Staff College, Maxwell Air Force Base, AL, 1999.

Wilbur, Ann et al., Metrics Guidebook For Integrated Systems And Product Development, International Council on Systems Engineering, INCOSE-TP-1995-002-01, 1995.

Williams, Brett T., Effects-Based Operations: Theory, Application and The Role Of Airpower, U.S. Army War College, Carlisle Barracks, Pennsylvania, 2002.

Winston, Wayne L., Operations Research: Applications and Algorithms, 3rd ed, International Thompson Publishing, 1994.

Zarchan, Paul and Howard Musoff, Fundamentals of Kalman Filtering: A practical Approach, 2nd edition, American Institute of Aeronautics and Astronautics, Reston, Virgina, 2005.

Zuse, Horst, A Framework of Software Measurement, de Gruyter, Berlin, 1998.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 074-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
14-09-2006	Doctoral Dissertation	Sep 2003 – Sep 2006

4. TITLE AND SUBTITLE	**5a. CONTRACT NUMBER**
Theory of Effectiveness Measurement	**5b. GRANT NUMBER**
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	**5d. PROJECT NUMBER**
Bullock, Richard, K., Major, USAF	**5e. TASK NUMBER**
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way WPAFB OH 45433-7765	AFIT / DS / ENS / 06-01
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) HQ USAF/A9 ATTN: Jacqueline R. Henningsen, PhD, SES, DAF 1777 North Kent St Rosslyn VA 22209	**10. SPONSOR/MONITOR'S ACRONYM(S)** **11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Effectiveness measures provide decision makers feedback on the impact of deliberate actions and affect critical issues such as allocation of scarce resources, as well as whether to maintain or change existing strategy. Currently, however, there is no formal foundation for formulating effectiveness measures. This research presents a new framework for effectiveness measurement from both a theoretical and practical view. First, accepted effects-based principles, as well as fundamental measurement concepts are combined into a general, domain independent, effectiveness measurement methodology. This is accomplished by defining effectiveness measurement as the difference, or conceptual distance from a given system state to some reference system state (e.g. desired end-state). Then, by developing system attribute measures such that they yield a system state-space that can be characterized as a metric space, differences in system states relative to the reference state can be gauged over time, yielding a generalized, axiomatic definition of effectiveness measurement. The effectiveness measurement framework is then extended to mitigate the influence of measurement error and uncertainty by employing Kalman filtering techniques. Finally, the pragmatic nature of the approach is illustrated by measuring the effectiveness of a notional, security force response strategy in a scenario involving a terrorist attack on a United States Air Force base.

15. SUBJECT TERMS
effectiveness, measurement, measurement theory,
system measurement, measurement frameworks, measurement guidelines,
decision analysis, game theory, competitive strategy development

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					Richard F. Deckro, DBA, (ENS)
REPORT	ABSTRACT	c. THIS PAGE			**19b. TELEPHONE NUMBER** *(Include area code)*
U	U	U	UU	188	(937) 255-6565, ext 4314; e-mail: Richard.Deckro@afit.edu

Standard Form 298 (Rev: 8-98)
Prescribed by ANSI Std. Z39-18