Naval Training Systems Center

MEASUREMENT METHODS AND METRICS FOR AIRCREW ASSESSMENT DURING CLOSE-IN AIR-TO-AIR COMBAT

S.T. Breidenbach

A.P. Ciavarelli

R. Sievers

DTIC

MAY 16 1986

CENTER OF EXCELLENCE FOR SIMULATION AND TRAINING TECHNOLOGY

NAVAL TRAINING SYSTEMS CENTER

ORLANDO FLORIDA

MEASUREMENT METHODS AND METRICS FOR AIRCREW ASSESSMENT DURING CLOSE-IN AIR-TO-AIR COMBAT

S.T. Breidenbach
A.P. Ciavarelli
R. Sievers

Cubic Corporation
Cubic Defense Systems
9333 Balboa Ave.
San Diego, CA 92123

LCDR M. Lilienthal
Naval Training Systems Center
Orlando, FL 32813

Final Report - September 1985

DTIC
ELECTE
MAY 1 6 1986

A

(ref: Box #18)

Tactical Aircrew Combat Training System (TACTS); Air Combat Maneuvering Instrumentation (ACMI); performance measurement; training feedback; training effectiveness

EXECUTIVE SUMMARY

The research and development program reported here was made possible through sponsorship by two NAVAIR departments, Human Factors and Training Technology (AIR 330J), and Range Instrumentation (AIR 630). The ultimate aim of this research activity is to improve the training effectiveness of instrumented ranges, such as the Navy's Tactical Aircrew Combat Training System (TACTS), by incorporating instructional capabilities.

Planned improvements to the TACTS call for development of an Instructional Support System (ISS) consisting of new range training capabilities that include:

1. **Instructional presentations (tutorials and simulated demonstrations)** – designed to teach air combat tactics and weapon employment.

2. **Improved training feedback displays** – designed to provide instructionally relevant data in operationally useful graphic formats to be used during aircrew debriefings.

3. **Diagnostic performance assessment methods** – designed to provide analysis and diagnostic review of performance against established, user-generated training objectives and proficiency standards.

Research presented in this report deals specifically with the performance assessment component of the Instructional Support Subsystem, and it should be noted that future integration of recommended measurement methods, and performance metrics developed, is planned in order to complete the entire instructional system package for TACTS. The main objectives of this study were to:

1. Review and summarize research completed over the past few years in air combat performance assessment, with particular focus on metrics and displays developed for assessing maneuvering and energy-related tasks.

2. Recommend the most appropriate measures and measurement methods necessary to incorporate reliable and valid performance assessment capabilities as part of an overall ISS development program.

Both of these primary objectives were satisfied during the course of the study. Measures and measurement methodology were reviewed from the standpoint of their utility in assessing aircrew training progress and proficiency, within the context of current tactics and weapons availability.

Our evaluation of measurement methods and their application was conducted in a systems framework designed to identify relevant task dimensions and operating conditions. In addition, some attention was given to the need and means to test reliability and validity of measures. The authors believe that greater emphasis is required in these areas to bring training measurement development more in line with professional test construction quality standards.

It is the authors' contention that any measurement program must progress through stages which, at a minimum, call for a clear statement of the purpose of measurement, and provide systematic testing of reliability, validity, and application of measures.

Procedures for reliability and validity testing are reviewed for those who may be unfamiliar with psychometric methods, and to provide a point of departure for a recommended measurement validation phase of research.

Additionally, if measures used in training are to have any utility to operational users, they must be relevant to the particular training program. This can be accomplished by involving users in the process of measurement development, and by emphasizing design of diagnostic graphic formats for presentation of performance data. In brief, we want measures that are scientifically credible and trustworthy, yet satisfy user requirements for meaningful feedback of training results on specific operational training tasks.

A major contribution of this study was the development of an improved maneuvering index of performance effectiveness for air-to-air combat. Improvements to metrics reviewed were considered essential based on our extensive literature review which revealed serious deficiencies in previously developed metrics and measurement methods.

Our analysis of extant metrics for measurement of maneuvering performance effectiveness indicated that the particular approaches reviewed either could not accommodate current tactical environments and weapon capabilities (e.g. most of the metrics reviewed are limited to rear-hemisphere weapons), or measurement outputs yielded unacceptable "truncated" performance scales.

The metric developed here is capable of measuring effectiveness of maneuvering for air combat within visual range, in which both opponents are equipped with modern, all-aspect weapons. Also, considerable attention was given during development of the algorithm to insure that metric outputs produce continuous, equal-appearing interval scale properties with improved sensitivity in reflecting dynamic performance variations.

Plans for validation of the metric and procedures for incorporating maneuvering measures into a more comprehensive task-oriented measurement framework are discussed. This technical discussion includes a review and analysis of measures and training aids used for assessment of energy maneuverability performance.

Several new energy metrics were identified in the literature, and their potential application to air-to-air performance assessment was discussed. In general, these specific metrics address fundamental changes in "energy use" concepts stemming primarily from introduction of high-performance aircraft and weapon systems. But none of the new energy metrics has been tested, and some energy metrics require input of aircraft performance data that are not presently available.

The authors, therefore, refrained from directly incorporating any new energy metrics into the proposed maneuvering effectiveness algorithm at this time, pending further evaluation and availability of aircraft test data.

Use of the more commonly understood energy metrics, such as specific energy (E_S) and the first derivative of specific energy (P_S), is recommended at this time, but their application deserves a more meaningful display. Suggestions for improving display formats used to assess maneuvering effectiveness and energy maneuverability are presented, and a phased prototype development program is recommended.

Finally, the report is comprehensive in its treatment and necessarily lengthy. For this reason, the authors have intentionally organized the report into separate, but related, topic areas. While we do hope that most would choose to read the entire report, it is possible to read any of the major report sections with little loss of continuity.

Accession For

NTIS GRA&I
DTIC TAB
Unannounced
Justification

Dist

A-1

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES AND FIGURES

LIST OF TABLES AND FIGURES (CONTINUED)

LIST OF TABLES AND FIGURES (CONTINUED)

1.0 INTRODUCTION

1.1 PROBLEM

Maintaining a high-performance aircraft in an optimum energy profile is an extremely difficult skill to teach, learn and assess. Th's is particularly true while the pilot attempts to maneuver his aircraft to attain a position advantage and missile fire opportunity against an adversary aircraft. The difficulty in acquiring energy maneuverability skills may be due, in part, to the fact that energy management can not be viewed either as an absolute or linear concept. Instead, an ideal energy state is transient at any moment of an air-to-air engagement and is dependent upon the tactical or position advantage of the fighter relative to the adversary.

Energy may be defined as the sum of the potential and kinetic energy of an aircraft. To maintain an optimum aircraft energy package, the pilot must judiciously use his energy resources. This often involves split-second tradeoffs of speed (kinetic) for altitude (potential) and vice versa. Energy tradeoffs are made by the pilot with the specific objective to maneuver the aircraft into position to launch a missile and destroy the target. A key to skilled pilot performance is to learn to control the interplay between energy and maneuvers which are important to winning the fight.

Because of the interrelationship between position advantage and energy, both have been a subject of study by researchers. Much research emphasis during the past 10 years has been directed to assessing the relative position advantage of one aircraft to another. The work originated out of the test and evaluation field and was undertaken with the aim to develop a global criterion of air combat. It was thought that a global criterion could be used to evaluate the effectiveness of a training system. These largely fragmented efforts produced a number of techniques which attempted to assess a pilot's position advantage performance during close-in maneuvers. Unfortunately, little or no effort was made to assess the validity and reliability of the measures. Since measures developed were often not task-based, they often lacked diagnostic value as training feedback.

On the energy side, researchers have emphasized the development of training aids as opposed to developing measures for assessment. Specifically, Pruitt, Moroney and Lau (1980) developed a prototype Energy Maneuverability Display (EMD) under the sponsorship of the Naval Aerospace Medical Research Laboratory. The EMD prototype was designed to aid pilots in learning energy management skills. The EMD was subsequently implemented on the Tactical Aircrew Combat Training System/Air Combat Maneuvering Instrumentation (TACTS/ACMI) by Cubic Defense Systems and has been in operation for the past several years. The intended use of the EMD, which is described later in the report, is to enhance training feedback related to energy management during TACTS/ACMI debriefs. For the reader who may be unfamiliar with TACTS/ACMI, a description of the system is presented in Appendix A. Some terms used frequently in air combat are provided in the glossary.

1.2 PURPOSE

The present study was conducted to bring to focus the fragmented research efforts performed in the position advantage area and integrate the results with available energy metrics. It was envisioned that a product of the research would be an up-to-date algorithm that would combine energy maneuverability and position advantage assessment capability. Recommendations would also be provided for presenting algorithm output in display formats that would be meaningful and diagnostic to aircrews. An additional study objective was to review and document the current operating status of the EMD and make recommendations for updating software and revising display formats to improve their instructional value.

1.3 SCOPE

± (Energy Maneuverobility Display)

The research reported here represents the development of performance algorithms and diagnostic displays for the maneuvering portion of an air-to-air combat engagement conducted within visual range. Although the metrics described in the report are specific to air-to-air combat, the methods for development, validation and display of performance metrics are generalizable to other missions such as air-to-surface and electronic combat.

to p. 1

2.0 TECHNICAL APPROACH FOR MEASUREMENT DEVELOPMENT AND VALIDATION

2.1 BACKGROUND

This report reviews and evaluates several attempts by researchers to develop measurement models for assessing the effectiveness of performance during air-to-air combat. Several approaches, and numerous performance metrics, have been developed over the past few years which purport to measure performance effectiveness. These measurement models deal primarily with the maneuvering portion of an air-to-air combat engagement. The measures used are intended for application as an aid to training on instrumented ranges, such as the Navy's TACTS, or on flight simulators, such as the Air Force Simulator for Air-To-Air Combat (SAAC).

For the most part, this report evaluates these measurement models on their own merits with respect to their ability to accurately depict the air combat tactical environment, and on the basis of their utility in training and training effectiveness evaluation. An equal, if not more important, criterion of evaluation addresses questions concerning the actual measurement purpose and properties of performance metrics in terms of whether or not the measures proposed can meet quality standards of a test instrument. These standards require adequate sampling of a specific task domain, and statistical demonstration of acceptable reliability and validity.

Test instruments designed for educational measurement, and other behavioral applications, generally follow a procedural famework that calls for systematic phases of test development and validation (Benson and Clark, 1982).

In the development of educational measures, for example, these procedures include (1) Definition of the purpose of the test instrument (2) Specification of the measurement domain in terms of psychological traits, or tasks and subject matter, (3) Definition of the target population for test administration (4) Preparation of performance objectives, (5) Composition of test items, based upon performance objectives (6) Analysis and selection of high quality test items, i.e. on the basis of item difficulty and discrimination values, and (7) Statistical testing of test score reliability and validity (Ebel, 1979; Mehrens and Lehman, 1973).

As a further precaution to insure quality and appropriate application of educational test instruments, professional agencies have established standard practices for development and validation of tests which explicitly specify acceptable levels of reliability and validity. (American Psychological Association, American Educational Research Association, and National Council on Measurement in Education, 1974).

When one surveys the great variety of performance measurement approaches taken in the training literature, a small sample of which is reviewed here, it is apparent that there are no equivalent standard practices in effect for the development, validation, and application of performance measurement systems in military training applications.

The performance measurement systems discussed in this report are no exception to this rule. Each of the mathematical models reviewed and evaluated was formulated, and sometimes applied, with little or no attempt to understand the measures employed from the standpoints of the actual purpose and application of measures obtained, the underlying task domain, and the quality of the test instrument itself, i.e. reliability and validity of measures obtained were not demonstrated.

2.2   MEASUREMENT DEVELOPMENT

The development of performance measures for use in training systems must address several key issues, which parallel those frequently encountered in design of educational test instruments, such as:

1. Purpose of measurement - Why do we measure?
2. Measurement Types - What do we measure?
3. Measurement Methods - How do we measure?
4. Measurement Application - How, and when, are measures used?
5. Measurement Reliability - Do measures used yield consistent results?
6. Measurement Validation - Do measures work for intended purpose?
7. Methodological Considerations - What are methodological limitations and caveats?

Each of these issues is briefly discussed here, with an eye toward establishing a framework for more detailed treatment and consideration during planned measurement validation tests outlined in a later section of this report.

2.2.1   Purpose of Measurement

The effectiveness of training depends largely on the quality of the evaluation on which training decisions are based. For without the benefit of meaningful performance assessment methods, we cannot 1) determine whether instruction is meeting its intended design objectives and 2) whether or not trainees have attained the capabilities desired (Gronlund, 1976; Gagne and Briggs, 1979).

For purposes of instruction and training, the availability of measures enables us to make judicious decisions concerning the allocation of training resources, i.e. we strive not to repeat practice on tasks that are already learned, but to concentrate on tasks requiring additional practice.

Training performance measures are important because they provide essential feedback to trainees and instructors about the progress of learning, and also because they provide a quantitative data base for overall estimation of training system effectiveness, i.e. whether or not a training system is meeting its design objectives.

2.2.2   Measurement Types

Measurement in education and training is based upon a cycle of events that call for the following.

1. Specification of training objectives or intended learning outcomes.

2. Planned training activities which include a determination of an appropriate instructional strategy and a method for delivery of training.

3. Evaluation of training results in terms of measurement and assessment of actual learning outcomes.

Performance-based training objectives provide the basic building blocks for development of a particular training system, and form the basis for later specification of measures used to assess student progress and to evaluate the effectiveness of instruction (Gagne and Briggs, 1979).

In the field of education, formally constructed task taxonomies in the psychomotor (Harrow, 1972), cognitive (Bloom, 1972), and affective (Krathwhol, Bloom, and Bertram, 1972) domains exist and are usually employed for identifying task dimensions in educational test development (Benson and Clark, 1982).

But in the area of complex man-machine systems, exemplified by instrumented ranges and simulators now increasingly used in military training, no commonly accepted taxonomy of operational tasks prevails.[1]

In the absence of a standard task classification system for human skills, we are compelled to apply time consuming task analysis method. in order to arrive at an a priori listing of tasks that are presumably critical to operational mission success. (See for example, Ciavarelli, Pettigrew, and Brictson, 1981a; Vreuls and Wooldridge, 1977)

In most of the measurement approaches reviewed here, particularly the attempts at modeling air-to-air maneuvering performance, this very important step requiring precise task specification was not undertaken. As a result, measures obtained are difficult to relate to actual aircrew tasks and expected learning outcomes. This limits the application and value of these measurement models, which by themselves, are not very useful in diagnostic assessment of training progress across task areas, or for evaluating training transfer results in operational missions, such as air-to-air combat, that typically represent multidimensional task environments.

Figure 2.2.2-1 shows a simplified air combat engagement sequence as usually flown on the Navy's TACTS. Corresponding to this figure is Table 2.2.2-1 which itemizes specific task-oriented training objectives and candidate performance measures (Ciavarelli, Williams, and Pettigrew, 1981b). This figure and table illustrate the multidimensional aspects of the air combat mission, and also exemplify the need to identify measures that cover several domains, encompassing cognitive (e.g. tactical decision making), procedural (e.g. missile launch sequence), and perceptual-motor (e.g. adversary aircraft tracking) task components. Figure 2.2.2-2 illustrates the types of measures possible and their application to measurement of performance in air combat. The information provided in these figures is a useful point of departure for further development and validation of measures, discussed in more detail in a later section.

2.2.3 Measurement Methods

The way to assess learning progress during training is to build tests or other assessment methods which directly measure the human performance described in the objectives of the training program (Gagne and Briggs, 1979). In most training applications, diagnostic measures are required in order to pinpoint instructional needs, so that trainees can concentrate on skills they lack and avoid unecessary instruction.

---

[1]Several attempts at deriving human skill taxonomies have been made, but most of these have concentrated on laboratory tasks that are difficult to extrapolate to operational missions (Fleishman, 1967), or have little instructional relevance (Merrill, 1972).
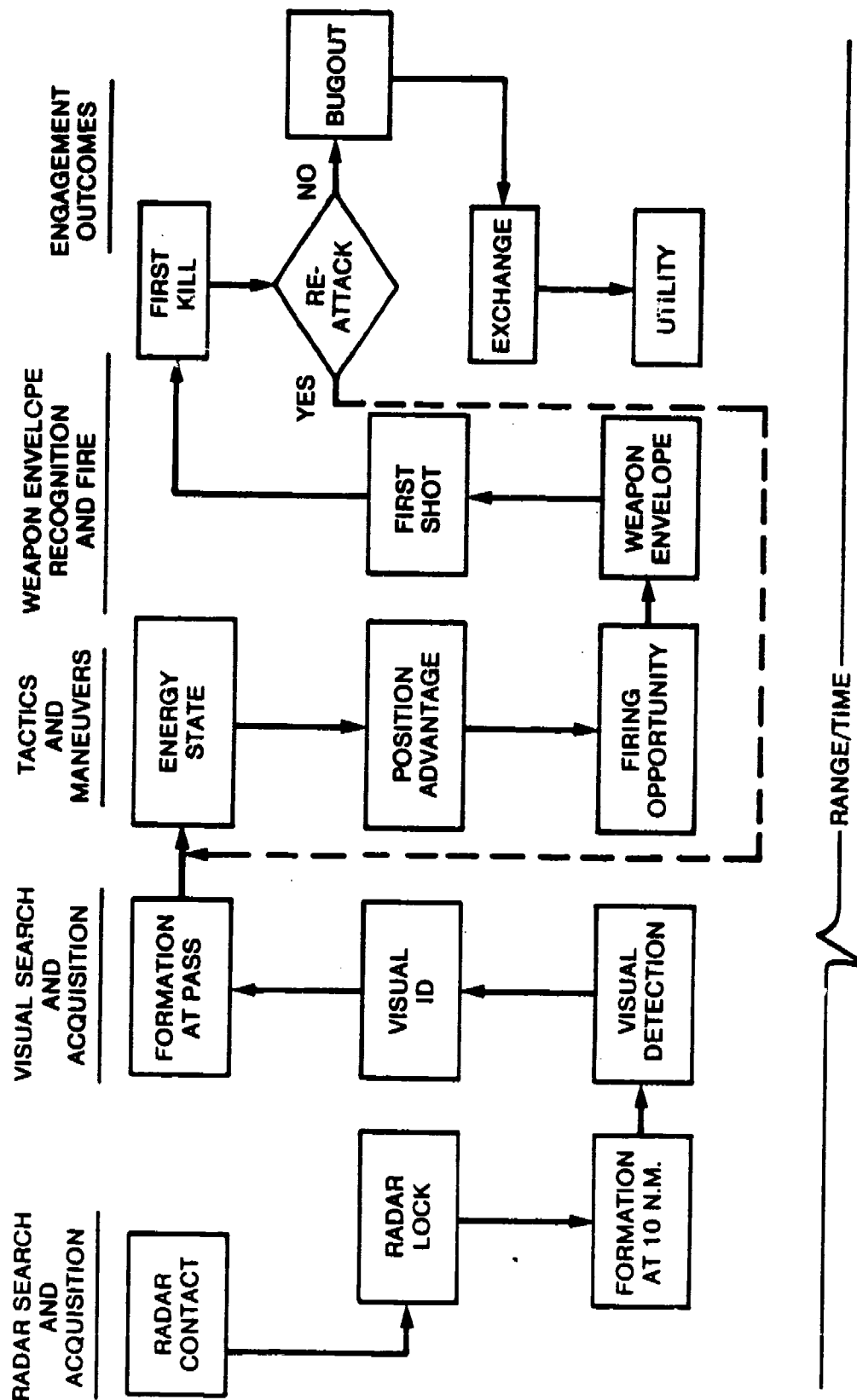
Figure 2.2.2-1 Simplified Air Combat Sequence: Shows Engagement Phases and Measurement Points

TABLE 2.2.2-1
AIR COMBAT ENGAGEMENT ANALYSIS

Training Objective	Performance Measure
1. Obtain early radar contact and lock-on	Interaircraft range and success rate (%) over engagements flown
2. Determine adversary attack formation at 10 nm.	Quantity and position of enemy aircraft
3. Obtain early visual detection of adversary aircraft	Interaircraft range and success rate (%) over engagements flown
4. Obtain early visual identification (VID) of adversary aircraft	Interaircraft range and success rate (%) over engagements flown
5. Determine attack formation at initial pass	Quantity and position of enemy aircraft
6. Maintain optimum energy state	Indicated air speed and altitude (energy package); composite energy metrics
7. Gain/maintain position advantage	% or proportion of engagement in offensive, defensive states
8. Gain firing opportunity	Time and/or % in envelope or fatal offensive state
9. Obtain first shot of engagement	Elapsed time and % first shots
10. Fire weapon in weapon envelope	Interaircraft range, angle-off-tail, pointing angle, airspeed and acceleration parameters
11. Obtain first kill of engagement	Elapsed time and % first kills
12. Execute successful re-attack	Iterate 6-11 above
13. Execute successful bugout by staying neutral, maintaining energy, and completing disengagement with no friendly loss	% neutral, indicated airspeed and altitude, % loss at bugout
14. Obtain favorable exchange rate	Ratio of fighter-to-adversary kills
15. Satisfy mission (utility) requirements	Neutralize threat aircraft and survive or minimize losses

MEASURE	DEFINITION	EXAMPLE(S)
● TIME	Elapsed time between key events, and duration between critical measurement points	● Engagement Start, Complete ● RADR, VID ● 1st Fire Opportunity ● 1st Shot Fired ● 1st Kill ● Time In Envelope or Engagement State
● STATE	Defined states based on tactical conditions, inter-aircraft position or aircraft flight dynamics	● Energy Level ● Position Advantage
● ACCURACY	Deviation or error score from prescribed envelope or performance standard	● Distance From Envelope Heart ● Deviation From Energy Ideal
● OUTCOME	Relative frequency % total success or success liklihood	● % Successful Weapon Fires ● Probability Of Kill (P_k)

Figure 2.2.2-2   Measurement Types and Application

By following this logic, educators and training specialists are turning increasingly toward the use of criterion-referenced measurement methods (Popham, 1978; Swezey, 1978; Glaser and Klaus, 1971).

Criterion-referenced measures are based on training objectives which describe learning outcomes in terms of the kind of student performance that we are willing to accept as evidence that instruction has been successful. This measurement method emphasizes determination of what an individual can do, without reference to the performance of others.[2]

The task framework (presented in Figure 2.2.2-1) implies application of a criterion-referenced measurement methodology. Using this approach relevant task dimensions are identified, together with doctrinal training performance standards to be use in proficiency evaluation.

This framework has already been applied in the development of the Performance Assessment and Appraisal System (PAAS) which provided the capability to store, retrieve, and display data in the form of diagnostic (graphic) feedback to aircrews training on the Navy's TACTS (Ciavarelli, et al., 1981b; Ciavarelli, 1982).

The PAAS allowed aircrews to assess performance against a set of proficiency standards (established by tactical experts at the Naval Fighter Weapons School) for key air combat training objectives.

A performance evaluation tool, such as PAAS, must meet a requirement for adequate sampling of tasks composing the highly multidimensional air combat mission. PAAS is deficient in this area because the system treated only discrete task components, i.e. radar contact, visual identification, missile fire, and engagement outcomes, and did not address some important continuous task operations. For example, PAAS did not include maneuvering effectiveness and energy maneuverability tasks and measures.

One purpose of the review of mathematical measurement models, and energy metrics, reported here was to select measures appropriate for assessing performance during the maneuvering portion of an air combat engagement, and thereby complete this particular measurement framework for air-to-air training evaluation.

2.2.4 Measurement Application

The purpose and application of measurement in training systems varies over the duration of instruction, as illustrated in Figure 2.2.4-1. For example, performance tests may be used early in instruction to assess the capabilities of students in undertaking planned instruction, and to appropriately assign them to a training program suitable to their particular readiness. Measures taken during initial instructional development are used to suggest modification and improvement to instruction design and delivery. Finally, measures taken after instructional delivery are applied to assess student progress and overall training effectiveness. (See Figure 2.2.4-1).

---

[2] As distinguished from norm-referenced measurement that interprets an individual's score in terms of a comparison between his performance and the performance of other members of a group (i.e. with respect to a group average).

PLACE IN SEQUENCE	PURPOSE	GENERIC CATEGORY	APPLICATION
PRIOR	TEST ENTRY SKILL LEVEL	SELECTION	• SCREEN APPLICANTS • ASSIGN STUDENTS
DURING	TEST DEVELOPMENT OF INSTRUCTION	FORMATIVE EVALUATION	• JUDGE PROGRESS & PROFICIENCY • DETERMINE LEARNING STRATEGY • TEST CURRICULA DESIGN & SEQUENCE
AFTER	TEST OUTCOMES OF INSTRUCTION	SUMMATIVE EVALUATION	• EVALUATE EFFECTIVENESS • ESTIMATE COST-TO-BENEFITS • MEASURE ATTITUDE/MOTIVATION • DETERMINE FINAL SKILL LEVEL AND READINESS

Figure 2.2.4-1 Measurement Application During Various Phases of Training

In the application considered here, we are primarily concerned with measures used to diagnostically assess performance during, and after, the course of training as it is undertaken on instrumented ranges and simulators.

The PAAS, once again, serves as an example. Figures 2.2.4-2 through 2.2.4-6 show possible application of menu-driven graphic displays for providing operational aircrews, and their instructors, with the means for assessing performance on critical air-to-air combat tasks. Figure 2.2.4-2, presents the "diagnostic assessment" menu for selection of a particular performance summary graph, and Figures 2.2.4-3, 4, 5, and 6 show, respectively, hypothetical performance graphics for radar contact, visual identification (VID), missile fire accuracy, and missile fire success rates. Using PAAS, operational aircrews are able to review performance results following training, and to determine proficiency levels against established standards, e.g. missile fire accuracy requirements. PAAS also provides (not shown here) air-to-air final engagement outcome scores (win, loss, draw) which can be applied to estimate overall performance effectiveness for air combat mission training.

2.2.5 Measurement Selection

In previous sections of this report we have drawn several parallel relationships between the construction of educational tests and performance assessment methods used in training systems. In brief, performance measures used in training (like test items and scores) must be consistent with performance objectives and must meet acceptable standards of reliability and validity.

A very important part of measurement development and validation begins with the formulation of a statement of purpose of the intended measurement instrument, which includes a specification of the domain to be measured, i.e. content area or constructs considered, and identification of the target group for which the instrument is intended. This type of early measurement planning helps us to select appropriate procedures for later reliability and validity testing. Initially, measures can be identified through task analysis procedures, and later verified for their relevance to overall mission success through correlation with a terminal measure of performance (Roscoe, 1980).

In the case of air combat measures, for example, antecedent event scores obtained on critical tasks composing the air combat mission, such as radar contact, visual identification, first-shot opportunity, and missile fire accuracy, can be correlated with final engagement win/loss outcomes (Ciavarelli, 1982). Figures 2.2.5-1 and 2.2.5-2 show results from analysis of about one hundred air-to-air engagements flown on the Navy TACTS, and illustrate this point. Figure 2.2.5-1 shows the empirically obt. ned relative frequency probabilities in an event tree format. This figure illustrates contingent relationships between various antecedent event scores and final engagement outcomes. For example, analysis of this figure into "best case" and "worst case" event outcomes indicates a .69 chance of obtaining a favorable win outcome for the best case, versus a .14 chance for the worst case. Closer examination of Figure 2.2.5-1 shows that the most significant event related to final engagement outcome is "1st missile shot." For instances in which fighters obtained the first shot opportunity, the probabilities of winning the engagement were .69, .56, .73, and .30. In cases for which adversaries obtained the first shot opportunity, the probabilities of a fighter win dropped to 0, 0, 0, .14. (See Figure 2.2.5-1 for illustration of these event relationships and outcomes.) By way of summary, when comparing results of early task performance on visual identification, first-fire opportunity etc., the results summarized in Figure 2.2.5-1 demonstrate that more favorable outcomes are attained by aircrews who have made early positive identifications, and have taken a first weapon fire opportunity.

DIAGNOSTIC ASSESSMENT MENU

Training Objective

Graphic Presentation

RADAR	Interaircraft range and success rate (%) over engagements flown
VISUAL	Interaircraft range and success rate (%) over engagements flown
ENERGY	Energy maneuverability score distributions
MANEUVERS	OFFENSIVE, NEUTRAL, and DEFENSIVE engagement state frequency distributions.
WEAPON	WEAPON FIRE position polar plots and success rate (%)
OUTCOMES	Fighter versus Adversary exchange rates
HELP	

Figure 2.2.4-2 Menu Used to Select Diagnostic Performance Graphs

Figure 2.2.4-3   Graphic Format Used To Review Radar Contact
Information For A Single Training Detachment

13

Figure 2.2.4-4 Graphic Format Used To Review Visual Identification Information For A Single Training Detachment

MISSION TYPE : 2V2   SQUADRON : VF
ADVERSARY AIRCRAFT : ALL   ADVERSARY SQUADRON : ALL
DETACHMENT DATES : 22JUN85–02JUL85

AIM MISSILE FIRE PLOT
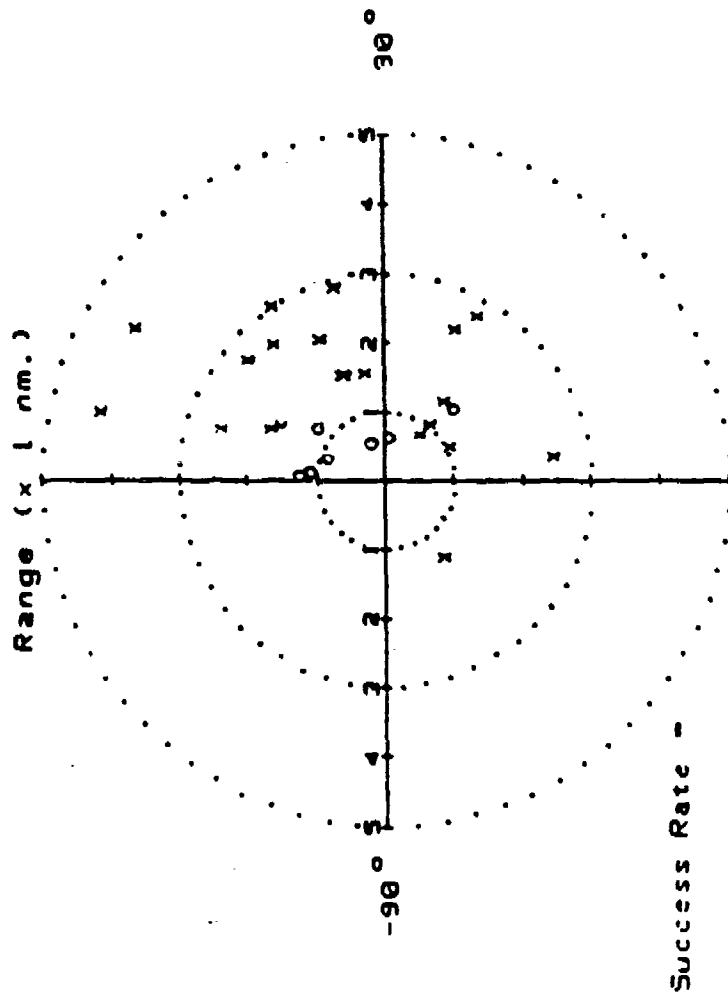
Figure 2.2.4-5   Graphic Format Used To Review Aim Missile Fire Accuracy
(Envelope Recognition) For A Single Training Detachment

WEAPON FIRE SUMMARY
(Engagements = 12)
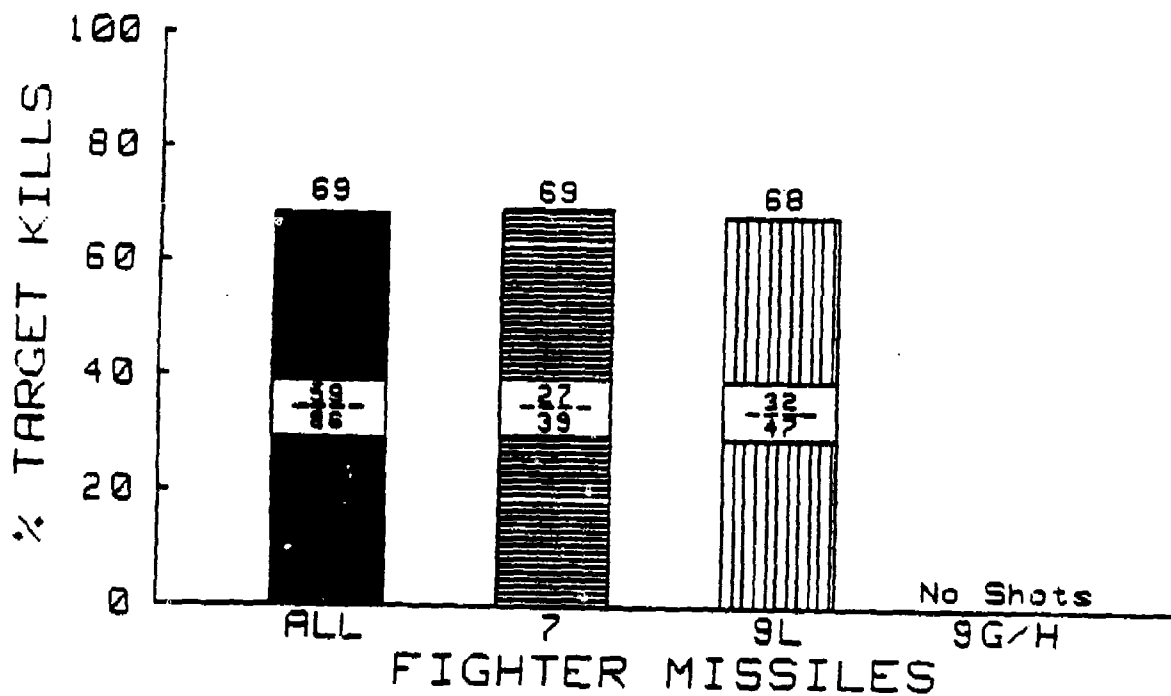
(FROM KRESS AND BREIDENBACH, 1983)

Figure 2.2.4-6 Missile Fire Success Rates
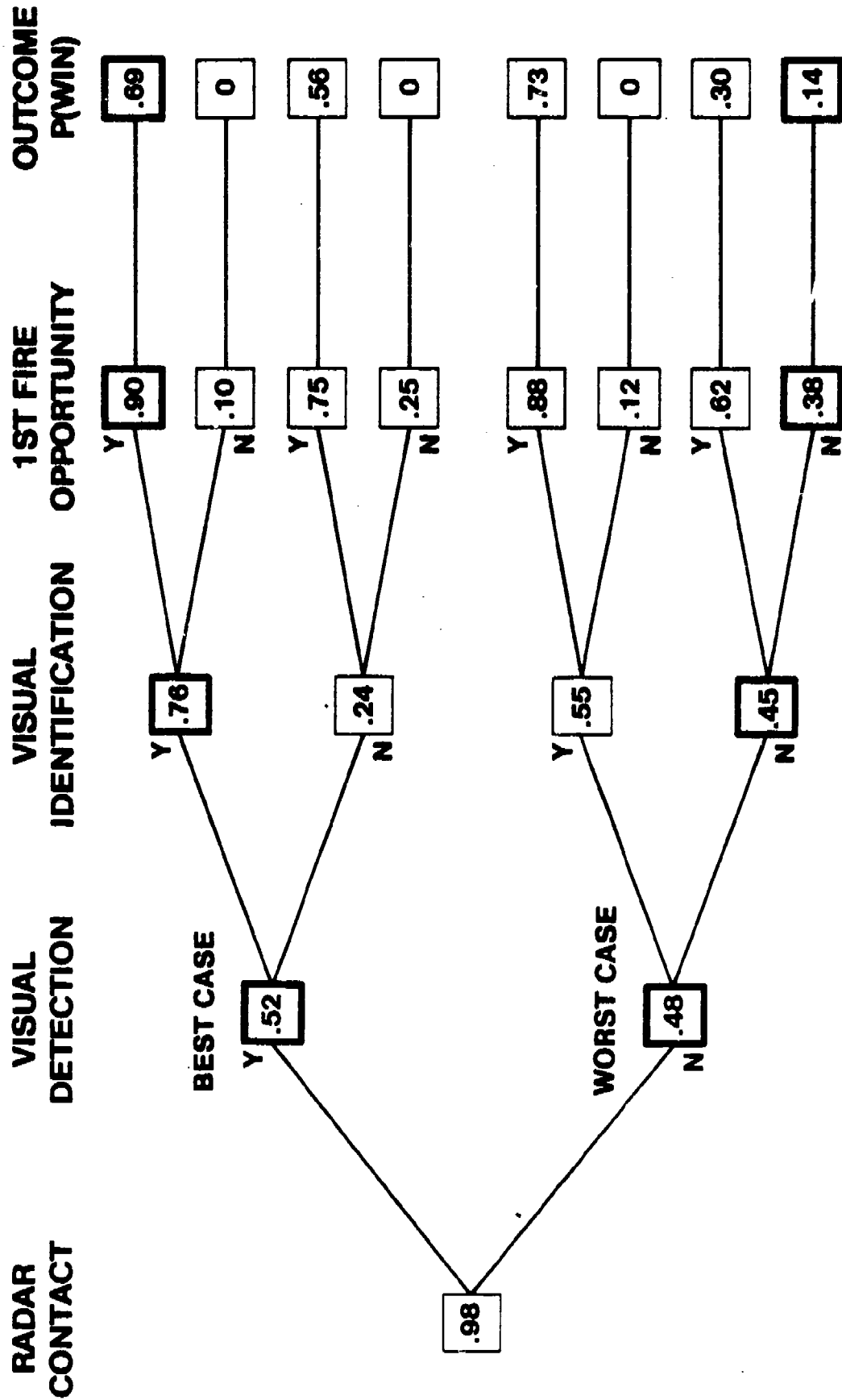
Figure 2.2.5-1   ACM Event Outcome Probability Tree

	1	2	3	4	5	6	7	8
1. RADAR CONTACT (YES, NO)	—	.24	.05	.20	.04	.07	.15	.16
2. VISUAL DETECT. (RANGE)		—	.32	.10	.07	−.17	.10	.13
3. VISUAL ID (RANGE)			—	.27	.14	−.14	.35	.34
4. 1ST SHOT				—	.10	−.08	.58	.64
5. WEAPON RANGE (ERROR)					—	.15	−.18	−.18
6. WEAPON POINTING (ERROR)						—	−.22	−.22
7. EXCHANGE							—	.98
8. OUTCOME SCORE								—

Figure 2.2.5-2 Intercorrelation Matrix

These results were supported during correlation analysis, presented in Figure 2.2.5-2, which shows relationships between antecedent task scores and final engagement outcomes. Significant correlations ($p < .05$) were established between engagement outcome scores and such early task measures as visual identification range, weapon fire accuracy, and first-fire opportunity. Measures thus selected subsequently define a meaningful test set that can then be evaluated further for reliability and validity.

2.2.6 Reliability Testing

Reliability is defined as the consistency of measurement over time, or precision of measurement.[3] Several practical methods for determining reliability are briefly described below (Allen and Yen, 1979; Benson and Clark, 1982):

1. Test-Retest - Administered by giving the same test to the same group (under the same conditions) at two different times, then correlating the scores using the product moment correlation coefficient.

2. Equivalent Forms - Give form 1 immediately followed by equivalent form 2 of same test; correlate the two scores using product moment correlation coefficient.

3. Internal Consistency - For tests with dichotomously scored items, use Kuder-Richardson formula 20; for all others use the coefficient alpha formula. (See Allen and Yen, 1979).

Of these, the test-retest method is most appropriate for testing reliability in heterogeneous test situations composed of numerous task dimensions, and is therefore applicable to the multidimensional framework used in air combat.

An example of this particular methods' application might be to have the same group of aircrews fly duplicate trials under controlled conditions on a flight simulator, using a fixed scenario and a mechanized, intelligent adversary, and correlating scoring results between successive test administrations.

To offset any possible contamination from learning effects, reliability testing should be conducted after aircrews have reached asymptotic levels of performance, i.e. are on the high end of a learning curve.

2.2.7 Test Validation Concepts

Test reliability is a necessary, but not a sufficient standard for test/measurement development. An equally important requirement for judging the quality of a test instrument entails demonstration of the validity for a given application.[4]

---

[3] Reliability ($\rho_{xx'}$), based on classical true score theory, is defined as (1) the squared correlation between observed scores and true scores, $\rho_{xx'} = \rho^2_{xT}$, or (2) the ratio of true-score variance to observed-score variance $\rho_{xx'} = \sigma^2_T / \sigma^2_x$ But these theoretical notions are most frequently represented by, test-retest, equivalent forms, and internal consistency reliability methods, because true scores cannot be empirically determined. (Allen and Yen, 1979 p. 73)

[4] It should be noted that reliability ($\rho_{xx'}$) limits validity (ρ_{xy}) because true score theory assumes that test scores can not correlate higher with any other scores than they can correlate with their own true score values, $\rho_{xy} \leq \sqrt{\rho_{xx'}}$ (Allen and Yen, 1979, p. 98).

A test is valid if it measures what it purports to measure (Allen and Yen, 1979). Another way of stating this is that a test must meet its intended purpose. Standards for test construction (APA, et al., 1974) discuss several kinds of "interdependent" validity which are briefly discussed below:

1. <u>Criterion - related</u> - Apply when one wishes to infer from a test score an individual's most probable standing on some other variable called a criterion. Predictive validity statements indicate the extent to which an individual's future level on the criterion can be predicted from knowledge of prior test performance. Concurrent validity reflects only the status quo of a particular time, i.e. both predictor test data and criterion measures are collected at the same time (APA, et al., 1974, p. 26).

Criterion-related validity is important in aptitude tests which may be needed to screen and select students with the appropriate entry level skills necessary to undertake a particular training regimen.

2. <u>Content validity</u> - is required when the test user wishes to estimate how an individual performs in the universe of situations the test is intended to represent. A test is "content valid" to the extent that it shows behavior demonstrated during testing constitutes a representative sample of behaviors required in the performance domain (APA, et al., 1974, p. 28).

The development of a task analysis framework, such as the one presented in an earlier section of this report for the air combat mission, exemplifies procedures necessary to establish the content specification for adequate sampling of the performance domain.

3. <u>Construct validity</u> - A construct is a theoretical idea developed to explain and organize existing knowledge. To establish construct validity, the investigator begins by formulating hypotheses about the characteristics of those who have high scores on the test and those who have low scores (APA, et al., 1974, p. 28).

Construct validity is based on an iteration between theory building and empirical verification of specific hypotheses, or predictions based on theoretical precepts. Hypothesis testing can use any one, or all, of three methods, 1) known-group comparisons, 2) factor analysis, and 3) multitrait-multimethod procedures (Benson and Clark, 1982).

1. <u>Known-groups</u> procedure - requires that a particular group tested already possess the attributes or capabilities being measured. In the case of education and training measures to be used in evaluating student achievement, tests/measures should be able to discriminate scores produced by entry level students (novices) from scores attained by those who have completed training (experts), or who have otherwise acquired the high levels of skill required to perform the operational mission. Measurement sensitivity may be tested in the air combat training application, for example, by comparing performance scores obtained on training ranges and simulators between the experienced aircrews (e.g. pilots under instruction) and more highly experienced aircrews (e.g. flight instructors).

2. <u>Factor analysis</u> methods – require the investigator to hypothesize the nature and number of factors underlying a particular measurement scale. The term factor refers to a theoretical variable derived by analyzing intercorrelations of test scores. (Allen and Yen, 1979). This procedure has proven to be particularly useful in the measurement field for reducing a large set of measures to a manageable number of factors, represented by highly intercorrelated measures. Factorial validity is established on the basis that factors so composed from a correlation matrix meet an expected, or theoretical, factor structure. A predictable factor structure provides evidence about the validity of the constructs originally hypothesized. Factor analysis techniques may be applied to identify, and appropriately combine, highly correlated measures into unitary (orthogonal) categories of specific task dimensions. For example, it is possible that energy metrics and maneuver effectiveness measures can be combined to form a unitary factor related to "air combat maneuvering precision" required to attain a position advantage and a weapon fire opportunity.

3. <u>Multitrait-multimethod procedure</u> – developed by Campbell and Fiske (1959), is used to determine the extent to which tests designed to measure the same capability, with different methods, are correlated. Measures are presumed valid if measures of the same capability correlate higher with each other than they do with measures of different capabilities using different methods, i.e. measures that correlate with other measures of the same construct are said to have "convergent validity." In the air combat measurement framework presented earlier, for example, we would expect objective and subjective (e.g. instructor ratings) of the same task dimensions to have convergent validity, expressed by significantly high correlation coeficients.

2.2.8 Methodological Considerations

2.2.8.1 Measurement Specification. A common mistake made in the development of performance measurement systems is the assumption that the main problem involves instrumentation, data reduction, and analysis (Roscoe, 1980). In actuality, numerous difficulties and technical hurdles face a researcher who attempts to develop, validate, and apply performance measures in military training systems. Some of the potential problem areas are summarized as follows (Adapted from Blaiwes, Puig, and Regan, 1973):

1. It is difficult to accurately define training objectives so that they are easily understood in terms of relevant task dimensions, desired measures, and required performance levels.

To accomplish this, a considerable investment must be made to analyze mission task requirements and to evaluate the relationships of subtask performance to overall mission success outcomes. Few have been willing to make the necessary investment to develop a comprehensive task/measurement framework, an essential foundation for performance measurement system development. As a result, many of the performance measurement systems, including those reviewed here, are one-dimensional views of highly multidimensional operational missions.

2. It is difficult to determine what to measure, when to measure, and how to interpret results of measures obtained, especially in situations such as tactical decision making where complex team interactions confound individual measures.

Methods are needed for isolating individual and team components of performance and for assessing their relative contribution to overall mission accomplishment. Both

individual performance standards (e.g. weapon launch success) and team performance standards (e.g. engagement outcomes) need to be established.

3.   Problems arise during measurement development because operational training situations are usually not amenable to experimental control. Measures are usually only available on a "not to interfere" basis.

This makes it difficult to attain the degree of control necessary for reliability and validity testing, which calls for obtaining measures under test conditions designed to limit extraneous sources of measurement variance.

Addressing this problem requires development and implementation of a performance measurement system that is highly valued and used by operational personnel, and the establishment of close working relationships with aircrews undergoing training. The use of criterion-referenced measures (based on aircrew inputs related to doctrinal training standards), and emphasis on diagnostic feedback instead of "global evaluation" criteria, has proven to be a manageable way to gain acceptance and support in operational training situations (Ciavarelli, 1982).

2.2.8.2 **Reliability Testing Problems.** Test specialists have accepted the reliability coefficient as an important indicator of the trustworthiness of a test instrument. Since the reliability of a measurement device reflects both the precision or consistency of measures, and the degree of measurement error, the magnitude of the reliability coefficient is used to estimate the extent to which we can generalize from one test application to another (Nunnally, 1975).

In cases where important decisions must be made about individuals on the basis of their test scores, e.g. assignment of students to different training treatments, our confidence in the test instrument must be very high. A reliability coefficient of .90 may still not be acceptable in some psychological test applications, e.g. personnel selection. Yet, such high values in testing, even under carefully controlled test administrations, are seldom attained (Bittner, Carter, Kennedy, Harbeson and Krause, 1984).

Of course, this situation is compounded in field test applications, such as military training, where control over practice effects, fatigue, and environmental conditions are difficult at best. Therefore, it would be highly unrealistic to expect reliability coefficients of such magnitude in applied settings.

This section discusses some of the more important methodological and practical limitations to testing measurement reliability in applied settings.

2.2.8.2.1 **Test-Retest** - reliability methods appear to be the most practical to apply in field settings, but there are some problems inherent in this procedure. (Allen and Yen, 1974):

1.   The most serious problem with the test-retest method is the potential for carry-over effects between test administrations. Practice effects, fatigue, and changes in test conditions, influence test scores and may result in underestimation or overestimation of actual test reliability.

2.   Different lengths of time between testing can affect the reliability estimate in different ways, sometimes underestimating and sometimes overestimating reliability.

3. Repeated measures on many tasks frequently show practice effects. At some point certain task measures stabilize, i.e, the mean and variance remains constant over repeated trials. But other tasks vary considerably in their stability characteristics. This lack of stability is indicated by the presence of a "super diagonal" in a correlation matrix between trials (Jones, Kennedy, and Bittner, 1981). In other words, the correlation between trials decreases with separation (i.e., adjacent trials have higher correlations). This finding is thought to reflect changes in relative skill composition during learning acquisition. The inherent instability of certain tasks measures, particularly during initial learning acquisition, may seriously limit the use of stability-based, test-retest methods, used for estimating reliability.

2.2.8.2.2 Internal Consistency Reliability - is estimated from a single test application and therefore avoids the problems associated with carry-over effects. However, as with test-retest methods, internal consistency methods also have limitations (Allen and Yen, 1979):

1. The most commonly applied methods yield a split-half reliability estimate that is based on the assumption that test items are homogeneous, i.e they measure the same trait or capability. Thus, both coefficient alpha and Kuder Richardson formulas can only be used for a homogeneous test.

2. Internal consistency estimates, based on split-half methods, do not yield accurate results if assumptions of parallel tests, or τ equivalence cannot be met.[5] For example, split-half methods are not appropriate for a speed test in which all examinees have achieved mastery and can obtain high scores if given enough time to complete the test.

Finally, all of the reliability estimation methods reflect a major weakness of true score theory (Weiss and Davison, 1981). Reliability estimates are highly sample dependent. For example, the magnitude of the reliability coefficient depends to a great extent on the distribution, or spread, of scores in the group of individuals tested. Typically, heterogeneous groups, by virtue of their obtained score variability, demonstrate higher reliability estimates.

Weiss and Davison recommend using the standard error of measurement (SE)[6] because SE is a useful index of test precision that can be used to define limits around an observed score within which we could expect the true score to fall, i.e. $X \pm Z_{SE}$, where X = obtained score, and Z = critical value of a normal score deviate (Allen and Yen, 1979).

In practical applications, particularly with the use of criterion-referenced testing, the SE can be applied to estimate the probability of misclassifying an examinee for a specific criterion level.

---

[5]Parallel tests - true scores and error scores are equivalent
τ equivalent - true scores are equal except for an additive constant.

[6]SE is based on reliability, as the formula $SE = S\sqrt{1 - r_{xx}}$ shows; where, S is the sample standard deviation.

Other reliability methods (Livingston 1971a, 1971b; Lovett, 1977) have been specifically developed for application with criterion testing approaches, since these measures typically yield restricted score ranges. These methods are based on the assumption that classical true score theory can be used to test reliability using an analysis of variance (ANOVA) framework. Using this approach, deviations from mean score (commonly used in normative measurement) are replaced by deviations from established criterion values. Reliability, then, as in ANOVA can be interpreted in terms of mean squared deviations from a criterion value (Swezey, 1978). Such methods may prove to be of greater utility in practical situations, where restriction of range is likely, given for example the highly homogeneous populations of students and use of mastery level, criterion-referenced test methods. But problems associated with gathering necessary reliability test data in military training environments still represents a significant limitation.

In spite of these limitations, attempts can and should be made to estimate reliability through selection and application of the most appropriate method. Decisions based on test scores related to training progress and proficiency, military readiness, and training effectiveness are important enough to warrant expenditure of time and resources necessary to have some degree of assurance that measures used in making such decisions are reliable and trustworthy.

As indicated in a previous section of this report, perhaps the best vehicle for reliability testing would be a flight simulator since conditions for testing can be more readily controlled. Therefore, plans should be incorporated in a measurement development program to accommodate simulator-based reliability testing of any proposed performance measurement system.

2.2.8.3 **Validity Testing Problems.** Over the past few years, some methods of validity testing have been sharply criticized. A brief review of these criticisms and recommendations for establishing a construct validation approach is presented below. This review is presented as a point of departure for development of a validation strategy for air combat performance measurement, presented in a following section.

1. Criterion-related validity. Up until the mid-1950s most validity testing was reported in terms of the accuracy of prediction between a test and some specific criterion measure (Cronbach, 1971). Dependence on criterion-related validity, however, has been criticized over the past few years by testing specialists. Some of the key criticisms are reflected in the test standards manual (APA, et al., 1974, p. 27) and include: (1) test conditions are never the same from sample to sample; (2) procedures assume that the criterion measure itself is a valid measure; (3) the sample used in validation may not be representative of the population; and (4) it is difficult, at times, to obtain an adequate sample size.

2. Content validity. Tenopyr (1975) was critical of test standards (APA, et al., 1974) for not providing adequate guidance to compare the kinds of validity. The distinction between content and construct, according to Tenopyr, has resulted in the most confusion. Tenopyr concludes that the term "content validity" should refer to inferences about test construction, whereas the term "construct validity" refers to inferences about test scores. The controversy surrounding the issue of content validity was more recently reviewed by Fitzpatrick (1983). Following an extensive critique of the content validity

literature, Fitzpatrick concludes that no adequate means for defining and quantifying content validity was found.

3.   <u>Construct validity</u>. Following publication of Cronbach and Meehl's (1955) classic article on construct validity, emphasis shifted to understanding the meaning of test scores in terms of underlying psychological processes related to obtained score variance. Test validation methods centering on constructs established a comprehensive statistical methodology designed to provide empirical evidence that test scores do in fact represent hypothesized capabilities. Cronbach and Meehl called for a test validation approach which examined, "the entire body of evidence offered," in order to determine the meaning cf test scores. Validation was to be established not on the basis of a single "validation study" but by building a mountain of evidence that supported predictions, while eliminating alternative hypotheses. These authors discuss use of a "nomological net" as a set of interlocking laws or formal theoretical principles that tie constructs and observable properties together in an integrated framework. The validation process is one on which empirical evidence is gathered through systematic experimentation which support hypothetical predictions, or that eliminates competing hypotheses. Determination of construct validity is a means to refine the relationships specified in a nomological net. Cronbach (1971) called for test validation based on the need to understand and interpret the meaning of test scores... "to explain a test score, you must have some sort of theory about the causes of test performance and their implications" (p. 443). He extends his point of view to educational measures as well, since instructors also need to have "some conception about acceptable performance," i.e., in terms of proposed standards for evaluation, type of measurement scale applied, and possible outcomes.

Messick (1975) maintains that all test developers and users must be able to answer at least two questions regarding test application, (p. 962): (1) Is the test any good as a measure of characteristics it is interpreted to assess?; and (2) Should the test be used for the proposed purpose? Neither criterion-related nor content validation techniques are adequate options in arriving at an answer to these questions. Even in practical educational settings it is important for the test developer and user to be able to determine whether or not the test is a "good measure," and to decide on its appropriate application. Construct validity information provides the body of evidence for test developers, test administrators and instructors to interpret the meaning, and value, of test results.

Improvements in statistical methodology are emerging to assist in establishing a set of coherent construct validation procedures. For example, Hocevar-Page and Hocevar (1982), propose use of confirmatory factor analysis (CFA) as a statistical methodology to evaluate factorial validity. Using CFA requires that a theoretical factor structure be furnished "a priori" during validation testing. The CFA technique avoids some of the pitfalls of exploratory factor analysis, such as indeterminancy of rotation. CFA also allows investigators to test alternate theoretical models for the best fit to empirical data as another useful tool in theory guided research.

Educational researchers are becoming increasingly aware of the need to develop and test "causal models" through path analysis methods (Pedhazur, 1982). Using this approach, a model that describes hypothetical relationships among key variables is constructed, diagrammed in the form of path network of variables, and tested through regression analysis. Hierarchical regression solutions are used to statistically confirm hypothetical relationships specified in the a priori model.

Both CFA and causal model methods require the investigators to be guided in their research by a theoretical, or conceptual, framework. Application of these methods helps to satisfy requirements specified by key test specialists (Messick, 1975) that all measures should be construct-referenced.

2.2.9 Strategy for Measurement Validation

Following recommendations covered in the test validation literature cited previously, a planned program of measurement testing should be developed to establish the construct validity of air combat measures. A strategy which calls for several stages of test instrument validation is recommended and includes the following procedures:

1. Tests of measurement framework - Procedures include testing of metrics developed to measure the maneuvering performance effectiveness. Initially, this is accomplished by correlating summarized outputs of metrics against a parametric set of data extracted from TACTS mission tapes. This procedure will be used to establish the relevance of part-task scoring metrics to final engagement outcomes.

    Subsequently, a more sophisticated correlation study should also be undertaken in order to demonstrate that the entire measurement set, including radar contact, visual identification, first fire opportunity, maneuvering effectiveness metrics, and weapon fire accuracy scores are meaningfully related to outcome scores.

    A possible approach to this more comprehensive treatment can be accomplished through causal analysis methods using path network correlation models. For example, a path network, based on findings reported earlier (Ciavarelli, 1982), can be tested using path analysis methods (Pedhazur, 1982).

    Verification of a hypothetical path structure provides one level of construct validity, in that a theoretical framework, established on the basis of tactical doctrine, can be empirically validated.

2. Tests of skill discrimination - Another step in the validation process calls for demonstration that measures used, in fact, are sensitive to individual differences and discriminate among aircrews of various skill levels. Statistical discrimination tests run between experienced and inexperienced aircrews can provide an experimental paradigm for validating the usefulness of measures for determining learning acquisition and skill retention levels.

3. Tests of User Acceptance and Training Utility - Finally, validated measures need to be implemented in user adaptable formats that provide diagnostic information essential to aircrews undergoing instruction. The benefit of and value to training can be assessed in terms of their direct utilization by aircrews to improve training, and/or through collection and analysis of attitude survey data.

3.0   REVIEW OF AIR COMBAT MEASUREMENT MODELS

Numerous air-to-air combat performance measurement models have been developed over the years. Despite the fact that researchers often had similar objectives, various technical approaches were used resulting in some diverse measurement models. This section reviews the principal measurement models that have been developed during the past 10 years. Before describing the models, the review method and the basis for comparing the relative merits of the models are presented.

3.1   REVIEW METHODS AND EVALUATION CRITERIA

After collecting pertinent documentation for several performance measurement models developed for air-to-air combat, an analytical approach was used to review and assess their relative merits. The models were evaluated on the basis of the following criteria:

    1.   Utility of measures
    2.   Appropriateness
    3.   Completeness
    4.   Scale Properties

The utility of measures involves a basic question: Are measures generated by a model meaningful and useful to aircrews? If a measurement tool does not provide feedback to aircrews which is task-based and diagnostic, chances are that it will not be useful or accepted by aircrews.

The appropriateness criteria concerns the extent to which the assumptions underlying a model are up-to-date by incorporating the latest developments in aircraft, weapon systems and tactics. For example, the model must be able to handle high-performance aircraft which launch all-aspect weapons.

A model's completeness considers how well the model adequately samples the task domains which comprise the air-to-air combat mission.

Scores generated by a measurement model should exhibit certain desirable scale properties. For instance, the scale should maintain, at a minimum, ordinal relationships with consistent rankings from low to high and be sensitive to performance differences. Preferably, the scores should also reflect the magnitude of performance differences. The distance between each point on a scale should be of equal length or interval. Other important factors to consider are the validity and reliability of the models. Unfortunately, since limited or no validation work has been attempted with the models, these factors could not be evaluated.

The remainder of this section describes several of the performance measurement models which have been developed for air-to-air combat and evaluates each based upon the criteria described above.

3.2 PERFORMANCE ASSESSMENT AND APPRAISAL SYSTEM

The PAAS was developed under the sponsorship of the Naval Training Systems Center. PAAS is essentially a stand-alone data base management and graphics system developed on a desk-top micro computer. The prototype system is designed to provide diagnostic feedback to aircrews after TACTS/ACMI engagements. This is accomplished by selectively retrieving performance measures and displaying them in meaningful graphic formats. Descriptions of the system and its measurement framework are presented in Ciavarelli, et al. (1981b) and Ciavarelli (1982).

A distinguishing feature of the PAAS model is the underlying measurement framework upon which it is based. This framework was developed from a task analysis of the air-to-air mission as flown on TACTS/ACMI. The framework highlights the mission phases and measurement points during the course of the air combat engagement. (See Figure 2.2.2-1 for framework illustration.)

The PAAS measurement framework is conceptually the most complete of the measurement models to be discussed because it encompasses the entire spectrum of tasks from initial radar contact to combat disengagement, or bug out. Limitations imposed by manual data input enabled the prototype to present only discrete performance measures during a preliminary field test with aircrews. However, the designer of the measurement framework envisioned that continuous measures related to tactics and maneuvers and energy management would ultimately be included. A sample of measures included in the PAAS measurement framework is illustrated in Table 2.2.2-1 and Figure 2.2.2-2.

The task-based nature of PAAS performance measures make them intuitively appealing, especially when presented in simplified graphic formats. This appeal was confirmed by limited presentations which were well received by aircrews during the preliminary field tests. While performance measures may be appealing and useful as feedback, they must also be valid and reliable as an assessment instrument to evaluate the progress, effectiveness or transferability of training. PAAS researchers attempted some preliminary validation work and found that many of their measures correlated with end-game outcomes. These early findings, reported in Section 2.0, provided a useful measurement set with empirically established relevance to end-game outcomes.

The remainder of this section presents the merits of other performance measurement models. The models focus mainly on depicting position advantage information during the maneuvering portion of the air combat engagement.

3.3 BACKGROUND FOR POSITION ADVANTAGE MODELS

Numerous methods have been developed to evaluate air combat performance and in particular maneuvering performance. Perhaps the simplest and most direct measure which can be applied to multiplane engagements is the kill/loss ratio. Although this metric may be calculated for various engagement scenarios and then applied as a predictive measure in similar engagements, it sheds little light onto the relative importance of the individual aspects of the engagements. More specific metrics have been developed, which, while correlating with the overall kill/loss ratio, address segments of the engagement which are a little further removed from the final outcome of battle. Among these are the single value metrics, time to envelope for radar and heat-guided missiles, time to first valid shot, number of valid and invalid shots, and number of missed opportunities (Robinson, Drinnon, Eubanks and Eddowes,

1981). However, these are still summary metrics and do not reveal how an engagement evolves with time.

To describe this evolution with time, researchers have developed what are called position advantage models. The following four position advantage models are discussed below:

1. Maneuver Conversion Model (Oberle, 1974)
2. Performance Index (Simpson, 1976)
3. All-Aspect Maneuvering Index (McGuinness, Forbes and Rhoads, 1983)
4. TACSPACE (Wooldridge, Kelly, Obermayer, Vreuls, Nelson and Nerman, 1982)

A common thread among the first three researchers was their purpose of measurement. Each desired an ultimate or global criterion of air combat which could be used for test and evaluation purposes. Although their approaches differed, all these researchers used instantaneous values of interaircraft parameters as a basis to represent maneuvering performance of aircraft engaged in air-to-air combat.

Three parameters commonly employed are the angle off tail (AOT), antenna train angle (ATA) and interaircraft range (R). The first parameter, AOT, is the angle (expressed in degrees) between a line extending out the tail of the target aircraft along its center line and the line of sight between the attacking and target aircraft. The ATA parameter is the angle between a line extending out the nose of the attacking aircraft along its center line and the sight line between the two aircraft. A third parameter, R, is the range between the attacking and target aircraft. The geometric representation of these interaircraft parameters is illustrated in Figure 3.3-1. All three parameters are normally computed and displayed on both TACTS/ACMI ranges and air-to-air combat flight simulators.

Since each of the models produce metrics which vary between any two instants in time, the dynamics of the engagement can be studied. These continuous data provide combat trends and enable one to identify sections of the engagement which are particularly critical to the outcome of the engagement. Knowing which sections are most critical provides the potential to optimize training by emphasizing techniques and tactics pertinent to these critical areas.

Two of the models (Oberle, 1974 and Wooldridge, et al. 1982) are state space models. Briefly, a state space model continuously measures an aircraft with respect to several variables and assigns specific location in a state space based upon ranges of values. For example, one cell of a state space may be defined by airspeed ranging between 250 and 300 knots, interaircraft range between 10,000 and 12,000 feet, and so on. Other cells within the space may be defined subsequently by different ranges of the same variables. As an aircraft is continuously measured during an engagement, it may occupy one or more of the cells defined in the state space. The states that an aircraft occupies can be examined and statistics calculated. While the underlying distribution of variables measured may be continuous, the coalescing of ranges of values into cells results in a discrete model with a finite number of states. In general, one disadvantage of state space models is that important information may be lost by reducing the dimensionality of measures to a finite number of states.

AOT = Angle Off Tail
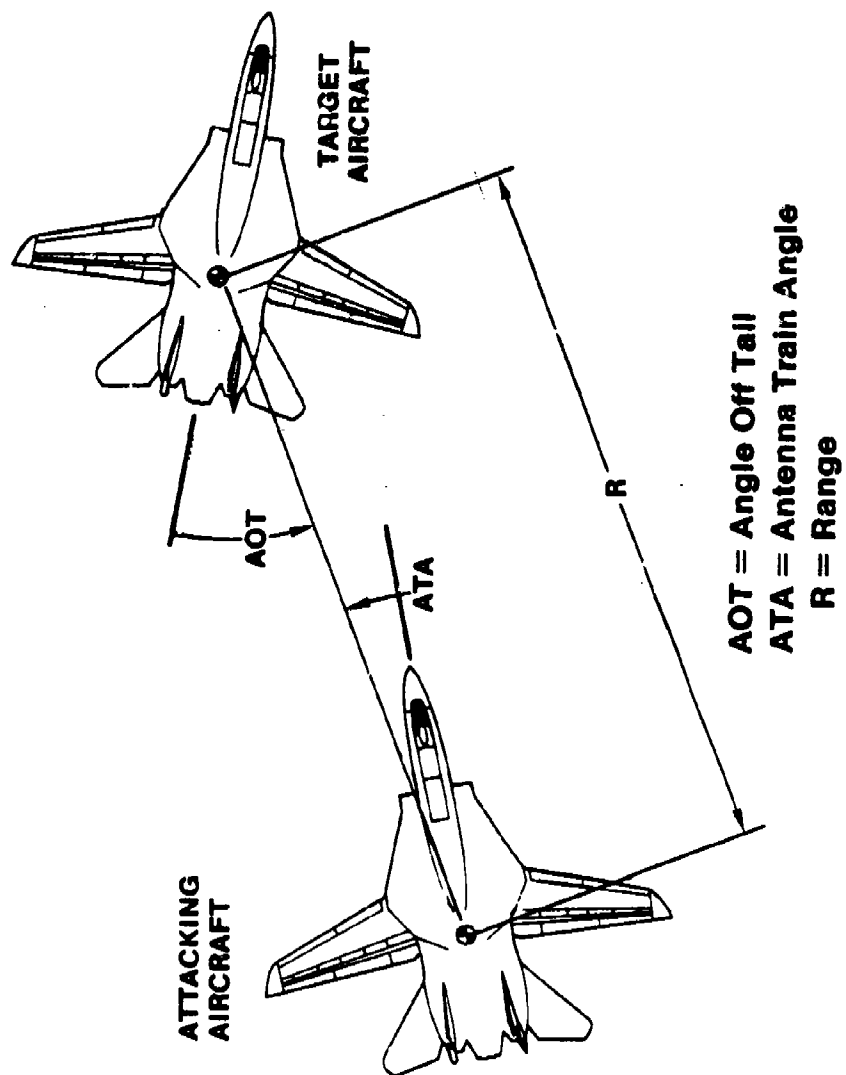ATA = Antenna Train Angle
R = Range

Figure 3.3-1 Geometric Representation of Interaircraft Parameters

The other two models (Simpson, 1976 and McGuinness, 1983) are continuous models with each being the product of several continuous functions of time. The four position advantage models will now be detailed.

3.3.1 Maneuver Conversion Model

First developed was the Oberle Maneuver Conversion (MC) model (Oberle, 1974). This was an attempt to put intuitive, qualitative statements pilots made about portions of an engagement into quantitative terms. The MC model was developed to reflect, precisely define and analyze the pilot's characterizations of combat as "Offensive," "Defensive" and "Neutral." Paired aircraft were compared with respect to relative range, fuselage orientation and closing velocity and then placed in one of five states in a one-dimensional space. Engagements were then analyzed using a semi-Markov process as follows:
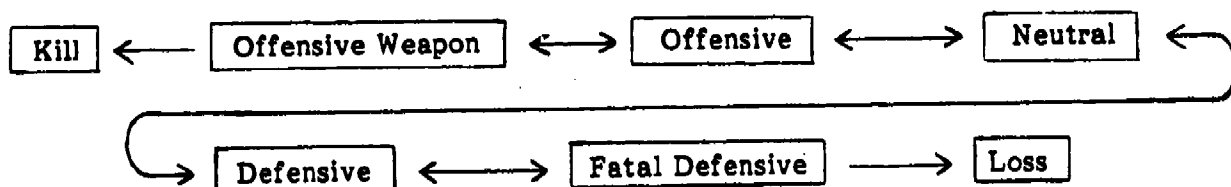
Kill and Loss are absorbing states, in other words they can be entered but not left, while the other five states can be entered from or exited to an adjacent state. If, for example, the adversary is in front of the fighter and with a relative range, look angle and closing velocity to put him in the fighter's weapons envelope, the fighter is in an Offensive Weapon state. Kill and Loss are entered by a successful simulated missile or gun attack. The other states are entered or left through maneuvers which alter the variables which determine the state.

A number of statistics can be calculated by continuously monitoring which states are being occupied by fighter and adversary. For example, a table of state conversion probabilities can be constructed resulting in a matrix as shown in Table 3.3.1-1. The probability matrix is a zero matrix except for the two diagonals adjacent to the main diagonal. This matrix provides a useful summarization of the expected sequence of events in an engagement. For example, transitions are restricted to adjacent states, and by using the probabilities of allowable transitions, the probability of a specific sequence or scenario can be calculated.

TABLE 3.3.1-1
STATE TRANSITION PROBABILITIES

Present State	Loss	Probability of Transition to Fatal Defensive	Defensive	Neutral	Offensive	Offensive Weapon	Kill
Fatal Defensive	-	0	-	0	0	0	0
Defensive	0	-	0	-	0	0	0
Neutral	0	0	-	0	-	0	0
Offensive	0	0	0	-	0	-	0
Offensive Weapon	0	0	0	0	-	0	-

Note: 0 is zero probability; - indicates non-zero probability determined empirically

Using the conversion probabilities in Table 3.3.1-1 and with the additional monitoring of time of weapon fire, the following statistics can also be calculated.

Probability of transition between states in a unit time interval.
Time in state probability distribution.
Probability a pilot will recognize a weapon opportunity.
Probability of fire out of envelope.
Probability of first-fire opportunity over adversary.
Probability of achieving and using first-fire opportunity.
Expected fraction of time in offensive or offensive weapon state.
Ratio of probability of kill to the probability of loss in a representative engagement.

Initially, the MC model assumed that the probability of switching from one state to an adjacent state depended on the present state and not on any previous states occupied. After a number of engagements on TACTS were analyzed, it was found that previous history had a significant influence on transition probabilities (Oberle, 1983). Probabilities were then calculated using the current state and either one or two immediately previous states.

An attempt was also made to expand this model to two fighter versus one adversary engagements. A fighter pair working together (section) is placed into a single state by individually comparing each member of the pair to the adversary. For example, the pair is offensive when at least one member is offensive and the other is higher than Fatal Defensive. Oberle (1983) did not attempt to adapt the MC model to more than three aircraft.

A severe limitation of the MC model is that it assumes that at least one aircraft uses only rear-hemisphere missiles. This results in symmetrical states between fighter and adversary. For example, a fighter's occupation of an Offensive Weapon state implies the adversary is in a Fatal Defensive state. Likewise, the adversary's possession of an Offensive state implies the first pilot is in a Defensive state. With forward hemisphere capable missiles, both fighter and adversary can simultaneously have a shot opportunity and thus both be Offensive. With present-day, all-aspect weapons, this tactical situation is a common occurrence, but is not handled by this model.

3.3.2 Performance Index

Historically, the next approach taken to quantify the importance of relative position in air combat maneuvering was by Simpson (1976). The specific metric developed was the Performance Index (PI). The metric consists of a continuous real-valued function of interairplane position and dynamics. This function is a product of three separate continuous real-valued functions which are as follows:.

1. The normalized direction angle function defined below which yields values ranging from -100 to +100.

$$DA_N = 100 \left[\frac{180 - (AOT + ATA)}{180} \right]$$

2. The range penalty function is a moderately complicated function involving actual, optimum, and maximum ranges for guns and missiles. Function values range from 0 to 1 as the interaircraft range goes from maximum to optimum missile ranges. The range penalty function formula and its characteristic "S" shape are presented in Figure 3.3.2-1.

3. An energy influence function was initially included as an attempt to compensate the PI for situations where the fighter and adversary have a non zero closing velocity. For example, a fighter may be in a position on the adversary's tail and near optimum range but the closing velocity is so great that an overshoot is forced. Despite its initial inclusion, the PI function calculated with and without the energy function produced essentially the same curves and was subsequently dropped from the PI (Oberle 1983). The energy influence initially included with the PI is presented in Figure 3.3.2-2.

An underlying motivation for the PI model is to have a relatively simple metric for two aircraft engagements which, when one is in an offensive position, will approximate the probability of kill if a missile or guns are fired from the current position. On the other hand, when the adversary is in an offensive position, the fighter's PI is negative and should be approximately proportional to the probability of the adversary scoring a kill. In a neutral situation PI equals zero.

The PI model appeared to be a useful and appropriate tool in the time frame that it was developed. Statistics similar to those calculated for the MC model can be calculated. A major drawback of the PI, however, is that it is not appropriate for use with weapon systems carried aboard new-generation aircraft. Specifically, the model was developed for rear-hemisphere weapons and does not handle all-aspect weapons which can be fired head-on or in the forward hemisphere.
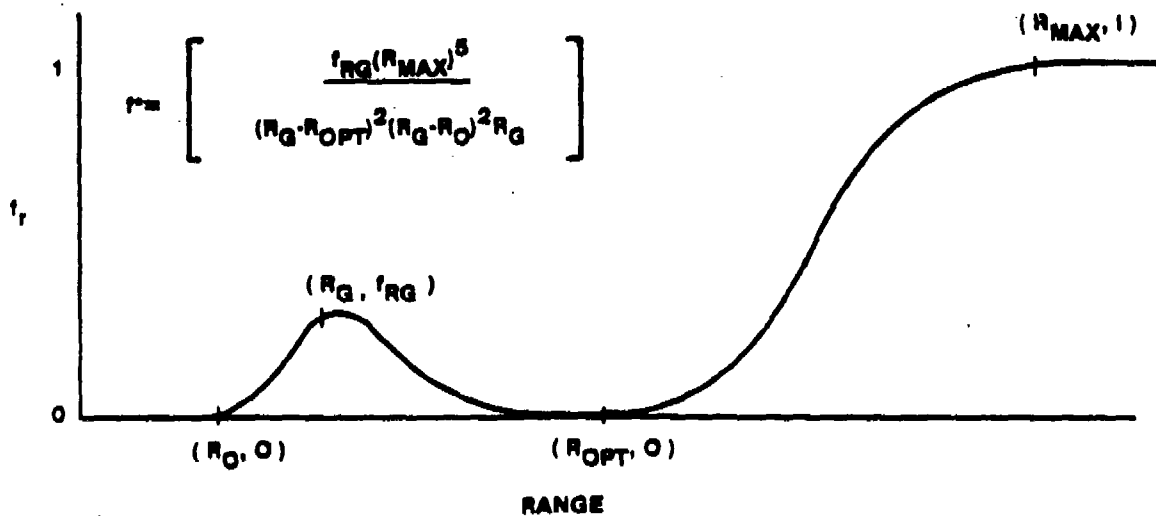
An extension of the PI to multiplane engagements (more than two aircraft) was made by weighting and combining the individual PI scores into a single composite value for "n" fighters in a section through the following equations.

$$K = \frac{\sum_{i=1}^{n} (PI_i) \mid PI_i \mid}{\left| \sum_{i=1}^{n} (PI_i) \mid PI_i \mid \right|} \qquad C = \pm C$$

$$PI_S = K \sqrt{\left| \sum_{i=1}^{n} (PI_i) \mid PI_i \mid \right|}$$

The value "C" in the above equation is a proportionality constant to establish maximum and minimum values.

$$f_r = f^\circ \frac{R}{R_{MAX}} \left[\left(\frac{R - R_{OPT}}{R_{MAX}} \right) \left(\frac{R - R_O}{R_{MAX}} \right) e^{-8} \left(\frac{R - R_G}{R_{OPT}} \right)^2 \right]^2 + \left[\frac{1}{1 + 500 \, e^{-\frac{(R - R_{OPT})}{R_{MAX}} \, 12}} \right]$$

$$f^\circ = \left[\frac{f_{RG}(R_{MAX})^5}{(R_G - R_{OPT})^2 (R_G - R_O)^2 R_G} \right]$$

(FROM SIMPSON, 1976)

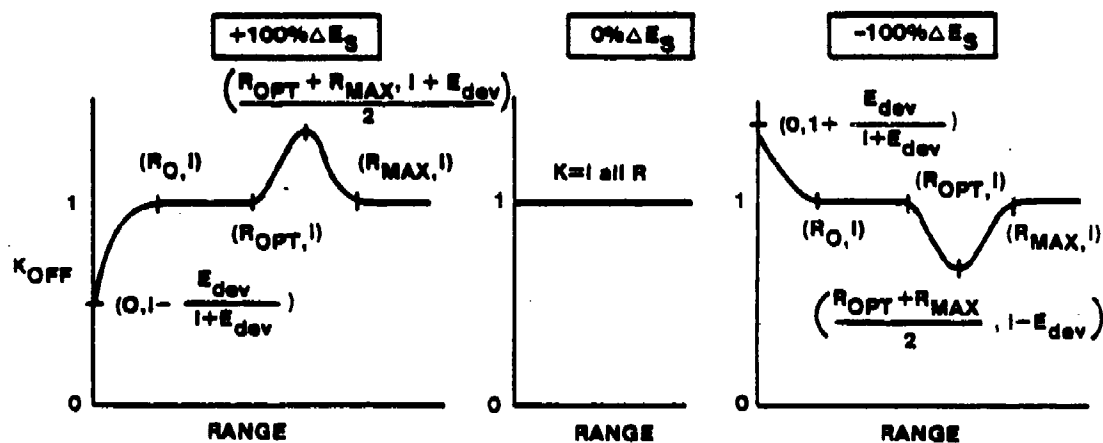**Figure 3.3.2-1 Range Performance Penalty Function**

$$PI = DA_N(I - f_r)K$$

OFFENSIVE'

$$K_{OFF} = 1 + \left[\frac{1}{1 + E_{dev}\ e^{-6.91\frac{R}{R_o}}} + E_{dev}\ e^{-4\left(\frac{2R - R_{MAX} - R_{OPT}}{R_{MAX} - R_{OPT}}\right)^2} - 1 \right] \frac{\Delta E_S}{E_S}$$

DEFENSIVE'

$$K_{DEF} = \frac{1}{K_{OFF}}$$

(FROM SIMPSON, 1976)

Figure 3.3.2-2 Energy Influence Function

A validation of this expanded model was not part of the memorandum (Simpson, 1976). In a technical report published the following year, Simpson and Oberle (1977) suggested another approach to expand the PI function, which was titled the Conversion Coefficient. However, in a still later report, Oberle (1983), both the MC & PI approaches to multiplane engagements were abandoned in favor of functions (the firing order methods) not addressing position advantage.

As a final technical note, the integration routine used for the numerical analysis of the PI and MC models (Simpson and Oberle, 1977) used equally spaced functions incorporating Simpson's rule and Newton's 3/8 rule. A more effective method would have used an adaptive quadrature algorithm. Basically, instead of equally spaced intervals as in a simple Newton-Cotes formula, shorter intervals would have been used where the function is varying rapidly and longer intervals used where the function is changing more slowly. These analysis techniques are discussed in detail in a number of numerical analysis texts; for example, Burden, Faires and Reynolds (1981).

3.3.3 All-Aspect Maneuvering Index

McGuinness, et. al. (1983) developed a third model in which the metric is called the All Aspect Maneuvering Index (AAMI) and is a continuous measure of aircraft offensiveness/defensiveness. It is the product of two separate functions and is a modification of Simpson's approach with the DA_N function modified to reflect all aspect missiles. The first function uses the fighter's orientation with respect to the adversary and is as follows:

$$100 \ ((90 - ATA)/90) \text{ for } 0 \leq ATA \leq 90 \text{ and } = 0 \text{ otherwise}$$

It differs from the DA_N function of Simpson in that it is a normalized linear function with respect to ATA and not ATA plus AOT, resulting in positive values whenever the adversary is in front of the pilot's own aircraft. If fighter and adversary are coming at each other, each may have a positive value for the ATA function. Consequently, the metric does not measure position advantage but fighter offensiveness. To obtain a position advantage measure, the AAMIs for pilot and adversary are subtracted to form what McGuinness calls Romp curves. Range and AOT are incorporated into the second function. This is calculated in two steps: first, minimum, optimum, and maximum ranges for each weapon type are selected from a table where these values are varied with respect to AOT. Second, these values are then placed in a function and a value from 0 to 1 is calculated. No further details were given in the report on the specific form of this function. The AAMI was mentioned to have been modified to incorporate closing velocity and altitude. Again, no more details were listed so the reasonability of these modifications cannot be determined. The statistics available are similar to those listed for the MC and PI models and are based on several parameters 1) time distribution of fighter and adversary AAMI values; 2) time of weapon fire; and, 3) time of kills. As with previous models, the AAMI was designed for one-on-one engagements.

Major problems with the AAMI are its lack of sensitivity to angle changes in interaircraft geometry and undesirable scale properties. These problems will be discussed in greater detail in Section 4.0.

3.3.4 TACSPACE

The final model in this review was developed by Wooldridge, et al. (1982) and is designed to maximally differentiate pilots of low and high skill level. A three-dimensional state space was developed with axis variables of ATA, AOT, and Range. The space was divided into small cells corresponding with specific intervals of the axis variables. These cells were then grouped into offensive, neutral, and defensive subspaces based on ATA and AOT. A discriminant score was calculated over location using the parameters collected in a flight simulator: air speed, turn rate, G, ATA, closing rate, throttle position, roll rate, and lateral velocity. The discriminant score is a statistically derived linear composite of the parameters which maximally discriminates between high and low skill groups. All the above parameters except throttle position would be obtainable on TACTS/ACMI. The statistics available include the time distribution of location in TACSPACE, the moments of this distribution, and the real time values of the discriminant scores.

There are two major problems with TACSPACE which are related to the appropriateness and utility of the model. First, the TACSPACE is oriented, as with previous models, to rear-hemisphere weapons and is not appropriate for use with all-aspect weapons and tactics. While TACSPACE could reasonably be adapted to the latest weapons, there remains the utility problem. Discriminant scores are useful in designing a selection instrument but their value as training feedback to aircrews is questionable.

3.3.5 Review Summary

Five air-to-air combat performance measurement models were reviewed. The PAAS model appears to be the most comprehensive because it encompasses most of the task domains of air combat. The model provides a framework of performance measures which are related to specific training objectives.

A primary deficiency of the PAAS model is a lack of adequate measures which sample aircrew tactics and maneuvers during the engagement. To fill this void, four position advantage models, which were developed as stand-alone measurement systems, were reviewed. Each model has its relative strengths and weaknesses. Primary limitations of the models are that either they do not incorporate all-aspect weapons or they exhibit undesirable scale properties. A list of the limitations of each model reviewed is presented in Table 3.3.5-1.

Based upon limitations of existing position advantage models, it was deemed appropriate that a new model be developed for inclusion in the PAAS framework. A description of the new model developed is presented in the next section.

TABLE 3.3.5-1
LIMITATIONS OF EXISTING PERFORMANCE MEASUREMENT MODELS

Model		Limitations
PAAS	o	Lacks continuous measures for maneuvering phase of engagement
Maneuver Conversion	o	Limited states collapse information reducing diagnostic feedback potential
	o	Rear-hemisphere weapons only
Performance Index	o	Limited to rear-hemisphere weapons
AAMI	o	Does not include AOT parameter
	o	Undesirable scale properties
TACSPACE	o	Model restricted to rear-hemisphere weapons
	o	Discriminant scores not meaningful to aircrews

4.0 DEVELOPMENT OF A MANEUVER INDEX

Previous discussions of the technical approach, and review of air combat measurement models have identified the need for performance measures and limitations of existing models. There are presently no continuous measures which adequately depict pilot maneuvering performance during engagements. This deficiency is particularly evident during the close-in maneuvering portion of the fight and with aircraft equipped with all-aspect (front-and rear-hemisphere) weapons. To fill this void in the air-to-air combat measurement framework, an algorithm for computing a maneuver index (MI) was developed.

An MI to accurately depict a pilot's position advantage during air combat is an important first step in developing energy maneuverability metrics. The importance of integrating energy with position advantage information is due to the fact that the optimum energy state of a fighter aircraft is highly dependent upon the relative tactical position of the fighter with respect to the adversary. If, for example, the fighter is in a defensive position, it is likely that the pilot would want to maintain a large amount of energy to permit escape. On the other hand, the pilot in an offensive state likely wants to carefully control his energy level to prevent overshooting the target. In any event, it is important to first develop an MI so that optimum energy profiles may be identified and tagged to specific position advantage states. The remainder of this section describes the basic components and computation of the MI.

4.1 BASIC COMPONENTS OF A MANEUVER INDEX

There are three basic components comprising the MI which were also used in Simpson (1976) and McGuinness, et al. (1983) models:

1. Angular geometry component (AGC)
2. Weapon range component (WRC)
3. Scale factor (SF)

The AGC quantifies the relative angular positions between two aircraft in space and can be expressed as a function of two interaircraft parameters. The two parameters, AOT and ATA, were previously defined and are illustrated in Figure 3.3-1.

Although AOT and ATA could be combined in several ways to compute the AGC of the MI, the methods developed by Simpson (1976) and McGuinness, et al. (1983) were used as initial baseline candidates, subject to evaluation. If it could be determined that either AGC computation method produced a metric with desirable scale properties, it would be used in the overall MI algorithm. Otherwise, it would be necessary to develop a new computation method.

Simpson's algorithm incorporates both the AOT and ATA parameters to quantify the angular relationship between two aircraft. On the other hand, McGuinness uses only the ATA parameter in his algorithm. These two baseline AGC computation methods are presented together in Figure 4.1-1, where a direct comparison can be made.

STUDY	COMPUTATION METHOD	COMMENTS
Simpson (1976)	$AGC = \dfrac{180 - (ATA + AOT)}{180}$	• Most Sensitive To Angular Geometry Changes • Incorporates Adversary Orientation • Generates Symmetric Scale
McGuinness (1983)	$AGC = \dfrac{90 - ATA}{90}$	• Insensitive To Angular Geometry Changes • Does Not Incorporate Adversary Orientation • Non-Symmetric And Truncated Scale Generated

Figure 4.1-1 Methods for Computing Angular Geometry Component of Maneuver Index

Simpson's and McGuinness's AGC algorithms were evaluated by computing AGC values for several relative aircraft position examples. A few relative aircraft position examples are presented in Figure 4.1-2. For each case presented in this figure, the relative position of a fighter and adversary aircraft is shown along with the associated ATA and AOT values. AGC scores computed using each algorithm are presented in Figure 4.1-3.

Clearly, the Simpson AGC algorithm produces a desirable scale of values and is superior to the McGuinness algorithm. As shown in Figure 4.1-3, the Simpson algorithm produces a scale with positive values when the fighter aircraft has a tactical position advantage and negative values when the adversary has the advantage. The algorithm produces zero values when the fighter and adversary aircraft are in "neutral" or "standoff" positions. Scale values range from a maximum of 1.00, when the fighter is in a most advantageous tactical position, to -1.00, when the fighter is in, what could be considered as, a worst possible tactical situation. Between these two extremes, the Simpson AGC algorithm produces scale values which are sensitive to changes in angular geometry. For example, as shown in Figure 4.1-3, case 1 has a higher scale value than case 2, which, in turn, is higher than case 3, and so on. These relative changes in scale value are mapped closely to relative changes in angular geometry. More importantly, the scale reflects the tactical situation in terms of fighter-adversary relative position advantage. Overall, the scale generated using Simpson's equation is symmetric and accounts for both fighter and adversary aircraft orientations.

Scale values computed using McGuinness's AGC algorithm are insensitive to changing tactical situations. For example, despite radically different tactical situations presented in cases 1-4 in Figure 4.1-3, the scale values shown are identical (1.00). This failure to accurately map scale values to tactics would result in misleading or erroneous feedback for aircrews. Since the AOT parameter was omitted from the algorithm, the resultant scale values do not account for adversary orientation and do not provide sufficient detail. The Romp curves described in Section 3.0, which are generated by subtracting fighter and adversary AAMI scores from each other, also provide insufficient detail because the summarized numbers used in the process have already lost useful information.

Due to the demonstrated superiority of Simpson's equation and the desirable scale values generated, it was found appropriate for use in computing the MI.

Although the AGC is a necessary component of the MI, it is insufficient to quantify a pilot's maneuver performance during close-in, air-to-air engagements. Another important ingredient of the MI is the WRC. Figure 4.1-4 illustrates the influence of weapon range. Despite the fact that case 1 and 2 in Figure 4.1-4 have identical AGC scores (1.00), the cases differ considerably in tactical significance. They differ because in case 1, aircraft 2 is in the heart of aircraft 1's weapon envelope, which is at the optimum range for aircraft 1 to obtain a missile kill. On the other hand, aircraft 1 in case 2 is in a good tactical position, but aircraft 2 is outside its weapon envelope, which precludes an immediate weapon kill. Due to the obvious importance of weapon range, the WRC was incorporated into the MI computation.

The third and final component of the MI is the SF. The SF is a numerical constant, which combines with the AGC and WRC, and is used to establish the origin and end points of the MI scale. The actual mechanics used to compute the individual MI components and their combination with the SF to form the MI are described next.
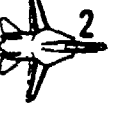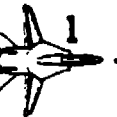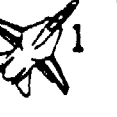
| | | | ANGULAR RELATION (DEG) FIGHTER – ADVERSARY | |
CASE	FIGHTER	ADVERSARY	ATA	AOT
1			0	0
2			0	45
3			0	90
4			0	180
5			45	135
6			90	90
7			90	180
8			135	180
9			.180	180

Figure 4.1-2 Relative Aircraft Position Examples

CASE	FIGHTER	ADVERSARY	AGC COMPUTATION METHOD	
			SIMPSON	McGUINNESS*
1			1.00	1.00
2			.75	1.00
3			.50	1.00
4			0	1.00
5			0	.50
6			0	0
7			−.50	0
8			−.75	0
9			−1.00	0

* McGUINNESS ASSIGNS AGC=0 FOR ATA>90°

Figure 4.1-3 Scoring of Angular Geometry Component for Relative Aircraft Positions

**Side View of Weapon
Launch Envelopes**

Case 1: ◣— 1

2

Case 2: ◣— 1 ◣— 2
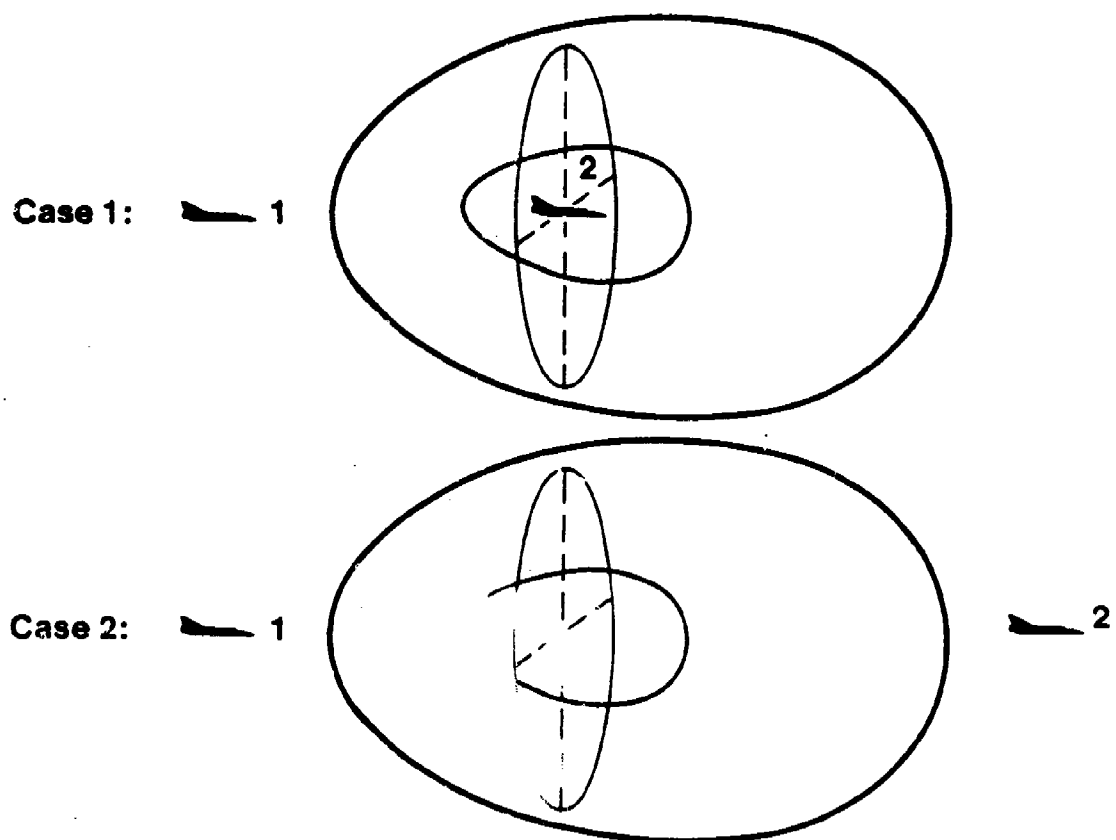
Figure 4.1-4 Requirement for Weapon Range Component of Maneuver Index

4.2 COMPUTATION OF THE MANEUVER INDEX

The first step involved in computing the maneuver index is to compute WRC scores separately for each weapon carried aboard fighter and adversary aircraft. Since probability of kill (P_k) results are provided on TACTS <u>only</u> after a missile has been launched, these values could not be used in computing the WRC. The next best option is to use standard "rule of thumb" launch boundaries that have been established for each weapon type. Figure 4.2-1 illustrates the top view of a hypothetical weapon launch envelope. For each AOT degree, minimum (MIN), optimum (OPT), and maximum (MAX) range boundaries have been established for each weapon. These values can be conveniently stored in computer memory as a look-up table for rapid retrieval.

In operation, the AOT value, at an instant in time, is used to determine MIN, OPT and MAX ranges from the look-up table. Next, interaircraft range (R parameter illustrated previously in Figure 3.3-1), is tested to determine if it is less than the MIN or greater than MAX range. If it is, then the opposing aircraft is outside the weapon envelope. If inside the weapon envelope, a normalized error score which reflects range deviation from OPT range is computed. The error score is multiplied by pi with the result expressed in radians. This result is, in turn, operated on by a cosine function which produces the WRC value for a missile. WRC values are computed for each weapon carried aboard fighter and adversary aircraft. The WRC ranges in value between 0 and 1. At the low end of the scale (WRC = 0), the opposing aircraft is outside all weapon envelopes. At the high end of the scale (WRC = 1), the opposing aircraft is at the OPT range or "heart of the envelope." The highest WRC value (optimum weapon) for the adversary aircraft is then subtracted from the fighter's highest WRC value. The resultant difference score (WRC_{Diff}) ranges in value between -1 and 1, with the sign depending upon which aircraft has the weapon advantage. A positive score indicates a fighter advantage, a negative score reflects an adversary advantage. If neither aircraft has a weapon range advantage, the difference score is 0. A summary of the WRC computation logic described above is shown in Figure 4.2-2.

The AGC is computed after the WRC difference score is obtained. The AGC score and WRC difference score are then <u>added</u> together. Unit weighting of the two quantities is employed since no empirical data are available to justify other weighting schemes. Finally, the MI is obtained by multiplying the SF to the sum of the AGC and WRC difference scores. A summary of the MI processing steps is presented in a simplified flow diagram in Figure 4.2-3. During each computation cycle outlined in the flow diagram, the optimum weapon, envelope status (in or out) and MI are obtained. This information can be displayed to range/simulator operators and aircrews in various formats.

**TOP VIEW OF WEAPON
LAUNCH ENVELOPE**
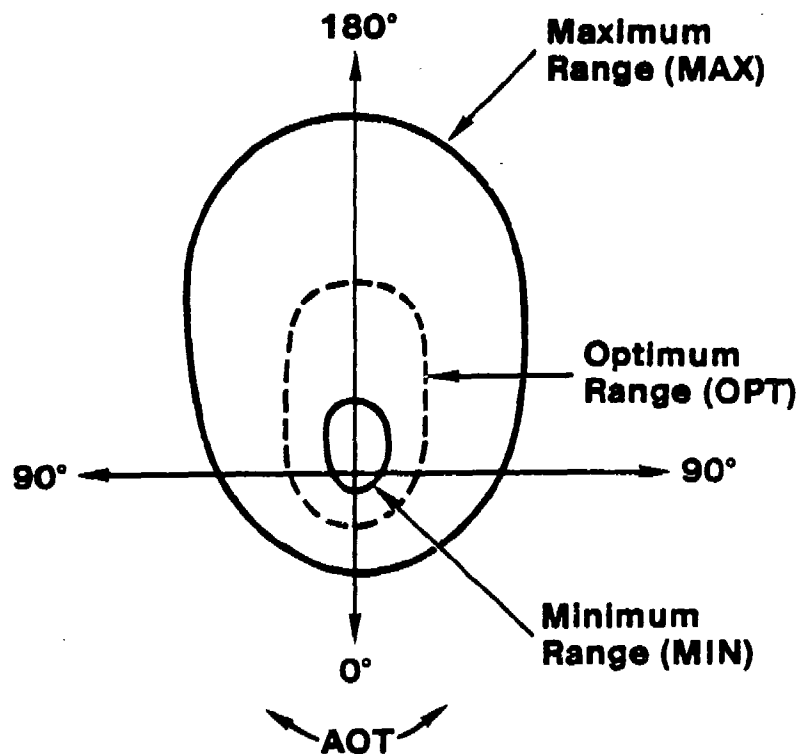
Figure 4.2-1 Illustration of Top View of Weapon Launch Envelope

WRC COMPUTATION LOGIC

1. Compute $WRC_{Fighter}$ and $WRC_{Adversary}$ separately for all weapons carried aboard respective aircraft. Use optimum weapon.

2. Based upon AOT, look-up MIN, OPT and MAX ranges of weapon envelope.

3. If R<MIN or R>MAX then WRC=0   (outside envelope)

4. If OPT≤R≤MAX then

$$\theta = \pi \left(\frac{R - OPT}{MAX - OPT} \right)$$

5. If MIN≤R<OPT then

$$\theta = \pi \left(\frac{OPT - R}{OPT - MIN} \right)$$

6. $WRC = \dfrac{\cos(\theta) + 1}{2}$

7. $WRC_{Diff} = WRC_{Fighter} - WRC_{Adversary}$

Figure 4.2-2 Summary of Computation Logic for Weapon Range Component of Maneuver Index

SCALE VALUES

$0 \leq WRC_F \leq 1$

$0 \leq WRC_A \leq 1$

$-1 \leq WRC \leq 1$

$-1 \leq AGC \leq 1$

$-2\,(SF) \leq MI \leq 2\,(SF)$

COMPUTATION

Parameters Passed: WPN, AOT, R, WRC

WRC_F – computed for each weapon. Use optimum weapon.

WRC_A – computed for each weapon. Use optimum weapon.

$WRC_{Diff} = WRC_F - WRC_A$

$$AGC = \frac{180 - (ATA + AOT)}{180}$$

$MI = SF\,(AGC + WRC_{Diff})$

START

COMPUTE WRC_F

COMPUTE WRC_A

COMPUTE WRC_{DIFF}

COMPUTE AGC

COMPUTE MI

DISPLAY MI, ENVELOPE STATUS, WEAPON
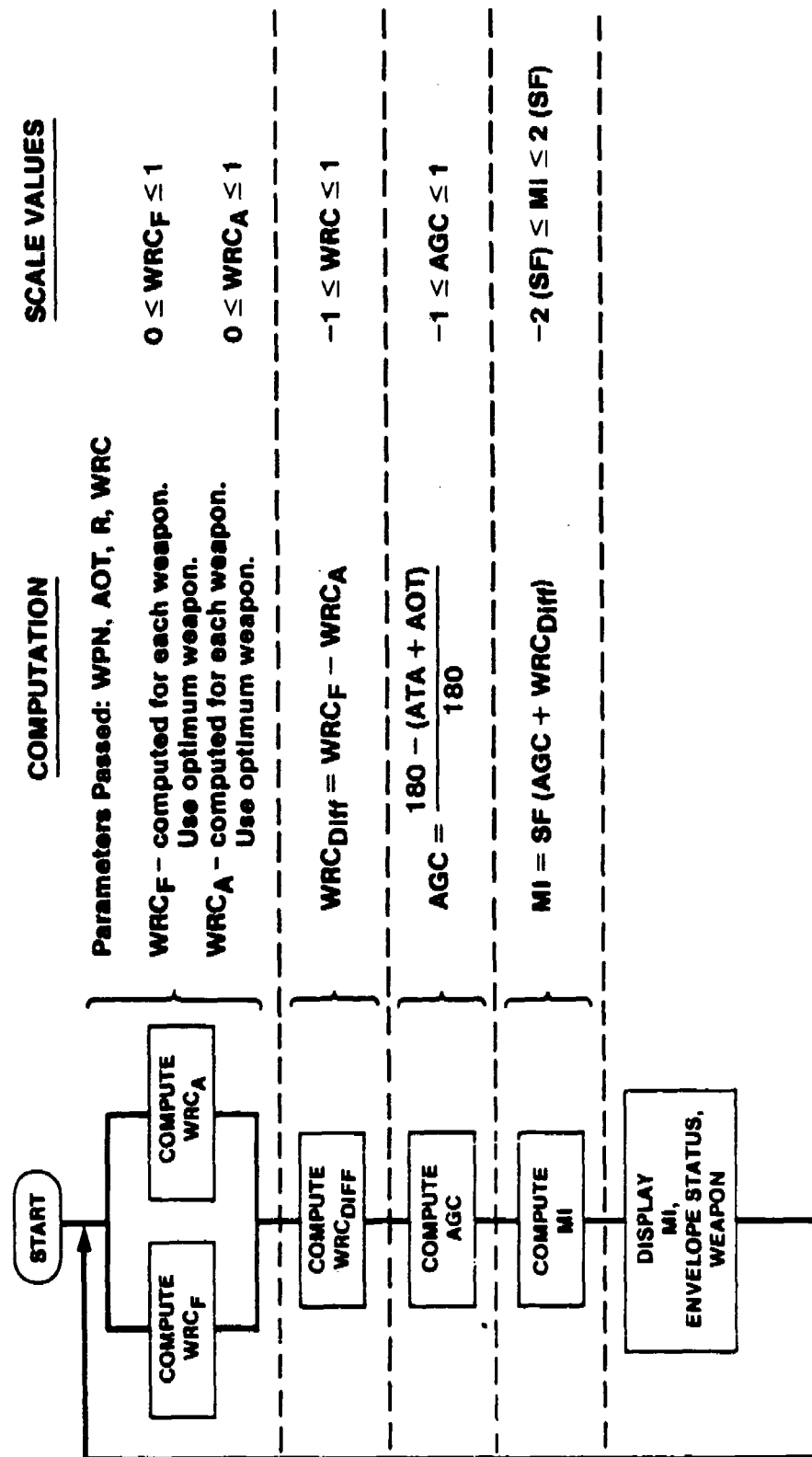
Figure 4.2-3 Simplified Flow Diagram of Maneuver Index Processing Steps

5.0 ENERGY MANEUVERABILITY CONCEPTS

5.1 HISTORICAL PERSPECTIVE

As new structural materials, propulsion systems and electronic hardware have become available, there has been a parallel evolution in the complexity and sophistication of fighter aircraft. Reflecting this increased sophistication, measures of aircraft capability and pilot performance have been altered, new metrics developed and some old measures refined. Today, Energy Maneuverability (EM) is the key concept in comparisons between fighter capabilities for visual range aerial combat. Once visual identification has been established (by fighter and adversary), the pilot, who more efficiently utilizes his energy to maximize maneuverability in getting an early shot opportunity, increases his chances of survival while decreasing his opponent's chances. The rest of this section briefly reviews the historical development of energy and maneuverability metrics and suggests some possible future directions.

Early in World War I it was observed that tactical advantages could be obtained by exchanging altitude (potential energy) for airspeed (kinetic energy) and vice versa. A pilot with an adversary behind and below could achieve an escape by diving, gaining speed and moving out of range. As combat experience was gained, a variety of basic combat maneuvers along with variations of each were developed. It also became quickly apparent that aircraft differed in their ability to perform these maneuvers. The SPAD was faster and could dive better while the Fokker triplane could climb faster and out-turn the SPAD. In World War II, Zeros were better at turning while the P-47 was better at diving. With these observations, tactics were developed which exploited the advantages of one's own aircraft. Early attempts at fighter comparisons, however, were qualitative and subjective.

To develop a more detailed and analytic approach, quantitative measures of performance were obtained. The first attempts at this produced point measures such as:

    Maximum airspeed
    Maximum altitude
    Thrust-to-weight ratio
    Wing loading
    Maximum constant energy turn rates

The first three of the above metrics describe the energy capabilities of the aircraft, the fourth specifically addresses maneuverability and the last combines energy and maneuverability. The meanings of the first two are self-evident; however, it should be pointed out that both depend not only on atmospheric conditions and current aircraft weight but also on turn rate. The third metric is concerned with the time rate of change of airspeed (acceleration), although this particular metric is of limited value with respect to more recent aircraft. Acceleration is the result of net force, which is the difference between thrust and drag. Drag depends on atmospheric conditions and aerodynamic efficiency which in turn depends on type of aircraft and current configuration such as swept or unswept wings. Thrust also varies with altitude, airspeed and position of variable geometry inlets, none of which is incorporated into the point measure, thrust to weight ratio.

Wing loading is the ratio of aircraft weight to wing surface area and is a measure of turning performance. Wing loading is inversely related to the aerodynamic load factor n_w, which is defined as the ratio of the force normal to the wind axis and gross weight. Thus as wing loading increases the load factor decreases. Turn rate can then be described with the following equation:

$$TR = \frac{1092 \sqrt{n_w{}^2 - 1}}{V}$$

where: TR is a level turn   degrees/second.
           n_w is the aerodynamic load factor   lbs/lbs (i.e., dimensionless).
           V is the airspeed   knots.

As seen from this equation, for load factors much above 1, turn rate is nearly proportional to load factor.

Up through the 1960's the approximation that wing loading was inversely proportional to load factor worked fairly well. In the 1970's this situation changed and n_w could be altered independently of wing loading. For example, wings could be swept or the wing camber changed with flaps. Additionally, with the development of thrust vectoring, turning can be affected separately from wings and flap deployment.

Despite the above limitations which have long been recognized, the single value or point measures of performance were valuable predictors of the outcome of aerial combat through WW II. Changes occurred during the 1950's with the advent of super-sonic fighters carrying guided air to air missiles. Through the early 1950's, each new fighter developed was designed to be faster than the one it replaced. With the appearance of the F-104, fighters entered the Mach 2 range. At these great speeds turning ability is greatly reduced and an aircraft must fly a much straighter path. While a higher top-end speed may be useful in quickly getting out of or closing in on gun range, it is not nearly as effective in avoiding a guided missile. In fact, since centrifugal force is proportional to the ratio of the square of the velocity over turn radius, a pilot can use an oncoming missile's speed to his own advantage. Although the missile can pull more "G's," if the pilot can sight the missile, he may be able to turn tightly and force a missile overshoot. The lack of utility of a higher maximum speed is also demonstrated by the fact that most aerial dogfights have taken place at subsonic speeds. For example, data from 1963 to 1973 in Southeast Asia indicate that combat was usually between 280 and 450 knots (Gunston and Spick 1983). Maneuverability and acceleration at subsonic speeds have since been recognized as more important than top-end speed.

Spurred by observations such as those above, a reassessment of aerial combat maneuvering was conducted. The interaction of turn rate, turn radius, airspeed, and maximum altitude was emphasized. This analysis resulted in the development of EM diagrams in the 1960's. Two basic diagrams (Figures 5.1-1 and 5.1-2) have since been adopted as standard display formats for EM descriptions (Martin, Luter, Caddy and Mueller, 1984). These are two-dimensional diagrams with either turn rate and Mach number or altitude and Mach number as the axes. As opposed to single-value measures, EM diagrams describe the variation of maximum turn rate or maximum altitude with Mach number. For example, in Figure 5.1-2 the dotted (not the longer dashed) lines mark the boundary of turn rate as a function of Mach number. At low speeds, turn rate is limited by the aerodynamic characteristics of the aircraft, in other

H-M DIAGRAM
1G

A-7E
2SW+8PYLON+FAMMO
28173 (60X FUEL)
ORG LTV RPT SEPT 78
VER AUG 83

TRUE MACH NUMBER
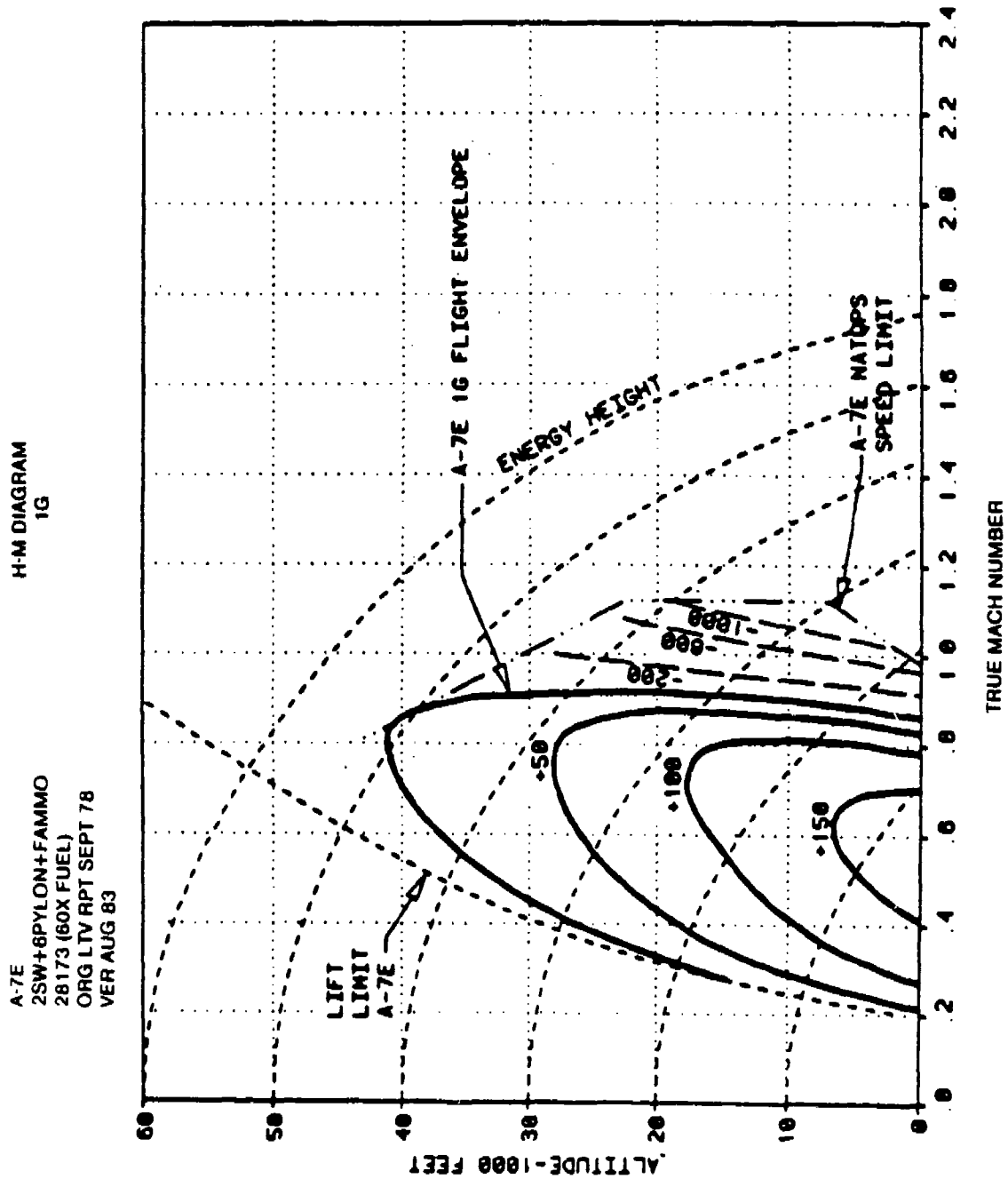
(FROM MARTIN, LUTER, CADDY AND MUELLER, 1984)

Figure 5.1-1 Altitude – Mach Diagram

Figure 5.1-2 Turn Rate – Mach Diagram

(FROM MARTIN, LUTER, CADDY AND MUELLER, 1984)

words, the lift limit. As velocity increases so does the force normal to the wings and eventually the maximum allowable load is reached. At this point, turn rate is simultaneously limited by both the lift and load limits. The velocity corresponding to this point is called the corner velocity and is the maximum instantaneous turn rate for level flight. As seen in Figure 5.1-2, the rate of energy change for this turn rate is very negative and the pilot is quickly losing speed and/or altitude. To maintain the turn rate altitude must be decreased. Above the corner velocity, in order to stay within the load limit, turn rate must be progressively decreased until the maximum speed of the aircraft is reached. At this point the turn rate boundary drops vertically to the Mach number axis. In addition, lines of constant specific excess power (P_s)[7] can be overlaid on the diagrams. The $P_s = 0$ curve represents the functional relationship between maximum sustainable turn rate and Mach number. In regions where available P_s is positive, the aircraft thrust can exceed drag resulting in either an increase in airspeed or altitude. These diagrams can also be used to compare dissimilar aircraft. Regions where P_s values differ between aircraft by 100 feet per second or more are, as a general rule, taken to indicate an advantage in maneuvering. From the EM diagrams, one can develop tactics to fly at a speed and turn rate where the opponent loses energy faster than the fighter so that eventually he will reduce his speed and can no longer turn at the fighter's rate. The fighter can then position his aircraft for a missile firing.

From Figure 5.1-2 it can also be seen that the velocity for maximum turn rate (corner velocity) occurs slightly above 400 knots so that too high a velocity is actually a disadvantage in a turning engagement. In a typical engagement scenario, a pilot would come into combat unloaded, i.e., not turning, and a bit above corner velocity. As he begins turning, energy is drained away ($P_s < 0$) and airspeed drops so that a short way into the engagement he is at corner velocity. Here turn rate is maximum, allowing the pilot to rapidly get a nose on position for a missile firing, or if need be, a quick high "G" defensive maneuver. After missile firing, the pilot goes down to a low "G" state and rapidly picks up air speed to reposition himself for another maneuver.

While the problem with just looking at maximum speed has already been pointed out, there is also a problem with just looking at maximum turn rate. If the fighter's maximum turn rate is greater than the opponent's, it would at first appear that the fighter has an advantage. However, if a pilot's P_s value for similar turn rates is much more negative than his opponent's, he may not be able to maintain near corner velocity long enough to get his nose on the opponent's aircraft. The negative P_s will cause a quick loss of airspeed and the opponent may soon have a turn rate advantage. Finally, maximum turn rate is not the only important parameter; a slow drop off in turn rate as one gets away from corner velocity is also valuable. This later desirable characteristic is not apparent with the point measure, maximum turn rate, but is in Figure 5.1-2.

EM diagrams allow pilots to look at performance throughout the flight envelope rather than at one point in it. Much more information is present in EM diagrams such as presented in Figures 5.1-1 and 5.1-2 than in the single-value measures first developed.

5.2 METRICS FOR NEW-GENERATION AIRCRAFT

While the EM diagrams discussed above are major improvements over earlier metrics, they too have their shortfalls. For example, Figure 5.1-2 compares maximum turn rate with Mach number. When this diagram was first developed, aircraft were

---

[7] P_s = Change in the sum of kinetic plus potential energy divided by time.

designed such that maximum turn rate occurred near minimum turn radius. Tactics which were developed for the aircraft to be near corner velocity at a critical segment of the engagement also came close to optimizing for minimum turn radius. Maximum turn rate no longer occurs at near minimum turn radius for some of the most recent fighters, and tactics have been developed which exploit one aircraft's turn radius advantage over another. The turn rate against Mach number diagrams cannot distinguish these differences in turn radius. Furthermore, there are maneuvers such as the barrel roll which has a maximum rate which varies with airspeed and turn rate. The standard EM diagrams do not address barrel roll, so obviously they cannot display the variation in barrel roll rate with other parameters.

Other difficulties with the current EM diagrams have been pointed out by Skow and Hamilton (1984). These are concerned with transition times between maneuver states. For example, an aircraft can be characterized at any point on Figure 5.1-2 by using an ordered triple of Mach number (MN), turn rate (TR), and P_s. Two different positions can be compared using their ordered triples (MN_1, TR_1, P_{s_1}) and (MN_2, TR_2, P_{s_2}). What is not shown is the minimum time required for a transition between these states. Characterization of these transition rates is a natural extension of the earlier EM metrics. Initially, measures such as maximum turn rate, maximum airspeed, maximum altitude, and thrust-to-weight ratio gave single values. Maximum turn rate and maximum airspeed in particular characterized the maximum energy possible. P_s is the maximum possible of the first time derivative of energy, and its placement on the standard EM diagrams reveals how it varies with other parameters. Instantaneous transition rates between ordered triples then give the second time derivative of energy. It was this second derivative which was addressed by Skow in his energy onset rate ($\Delta P_s/\Delta t$), which will be described below. Likewise, although Figure 5.1-2 displays how maximum turn rate varies with Mach number, it reveals nothing about the minimum time required for a transition between two TR-MN pairs. There are aerodynamic limits on how rapidly turn rates can be altered which depend on Mach number. As a metric for this transition time, Skow has suggested what he called the turn agility metric which will also be described below.

The energy onset rate is defined as the difference between P_s at maximum thrust, minimum drag and P_s at minimum thrust, maximum drag, divided by the time to make the transition. Depending on what is limiting and whether one is concerned with acceleration or deceleration, it is either spool-up time, spool-down time, time to deploy flaps, or time to deploy thrust reversers. An argument for the metric is that while one aircraft may have a higher maximum climb rate or maximum acceleration, another aircraft with a shorter spool-up time may be able to accelerate faster during the early part of an engagement.

The first aircraft may be the recipient of a missile before there is time to take advantage of its higher top-end acceleration. The second metric is turn agility and is equal to the turn rate divided by the minimum time required to switch from a 45° left or right-banked turn to a 45° banked turn with the opposite bank.

These two new metrics, being approximate time derivatives of previous metrics, are further refinements of measurements of air combat maneuvering. In the future they may play important roles in aircraft design and combat tactics. In the meantime, much work needs to be done with them. Values must first be generated for various aircraft, including the variation of these new metrics with other parameters. For example, the energy onset rate will vary with turn rate depending on how the inlet distortion and recovery varies with angle of attack (Skow and Hamilton, 1984).

Similarly, the time required for a 90° change in bank angle depends on maximum roll rate. As roll rate decreases with higher angles of attack, so does turn agility.

Any models developed for the variation of Skow's metrics with other parameters need to be confirmed with flight tests. After confirmation of accuracy, these new metrics must also be tested for utility on TACTS/ACMI ranges. For these metrics to be useful, it needs to be shown that differences in pilot and/or aircraft performance as measured by these metrics, correlates significantly with discrete engagement outcomes. Based on the metric definitions, it would appear that the best chance for a significant correlation for the energy onset rate metric would be in engagements with large and rapid changes in throttle position, while the turn agility metric would be expected to do best in engagements with rapid changes in bank angles.

Pilots must be able to transfer a classroom understanding of these metrics to cockpit controls during actual flight. Pragmatically, this must be accomplished within reasonable amounts of classroom, simulator and flight time. Considering the difficulty in applying the current TR-MN display to the TACTS/ACMI ranges, an immediate attempt to apply these new metrics seems premature.

Based upon the above considerations, it is recommended that a very simple metric be applied to TACTS/ACMI ranges. It is then left to the pilot to interpret this metric and decide on specific cockpit actions. The metric suggested is the specific energy (kinetic plus potential energy divided by total weight of aircraft), E_S, or possibly its first time derivative, P_S. Only after a thorough analysis of the relationship between the variation of E_S during TACTS/ACMI maneuvers and the outcome of the combat, should a metric be derived for performance with respect to energy maneuverability.

6.0 ENERGY MANEUVERABILITY DISPLAY

6.1 DESCRIPTION

The EMD was developed and implemented for use on TACTS/ACMI. Development of the EMD is detailed in Pruitt (1979) and Pruitt, Moroney and Lau (1980). The display is designed as a training aid to assist aircrews in acquiring energy maneuverability skills.

The EMD is a version of the standard turn rate versus Mach number diagram which was presented previously in Figure 5.1-2. The diagram has been used for years by aeronautical engineers to compare aircraft performance. A simplification of the standard diagram contains two nested graphs, which are referred to, collectively, as the "Maneuver Triangle." Data points which comprise the Maneuver Triangle portion of the EMD are defined and illustrated in Figure 6.1-1. The outer or larger of the two nested graphs represents maximum instantaneous turn rates achievable by an aircraft at different speeds. The side of the outer graph to the left of the apex is the aerodynamic lift limit of the aircraft. The structural and placard limits are shown along the outer graph to the right of the apex. The inner or smaller graph represents the P_S = 0 curve. Below this curve, energy may be gained, while above it, energy is lost.

Each graph is a linearized approximation over five subintervals which are defined by the points in Figure 6.1-1. Since the variables TMRPS, TCTPS, TXRPS and TMXPS are maximum sustained turn rates, they occur at P_S = 0. Connecting these points with VMR and VMX airspeeds enable linear approximations to be made for the five subintervals. Likewise, the maximum instantaneous turn rates defined by the points TMR, TCT, TXR, and TMX are connected to 0 and VMX airspeeds and provide a basis for linear approximation over five subintervals. Each of the variables defined in Figure 6.1-1 represents a vertex whose location depends at any moment on altitude, weapons load and fuel status. In operation, the two nested graphs defined above for the EMD are updated using a look-up table which assigns a value to each variable depending upon gross weight, armament and altitude.

The EMD is displayed on the TACTS/ACMI Display and Debriefing Subsystem (DDS), where graphs of two aircraft are overlaid upon each other. An illustration of EMD is presented in Figure 6.1-2. In the example, the maximum instantaneous and sustained curves for aircraft 1 are shown as solid lines. The graphs for aircraft 3 are presented as dashed lines. The numbers 1 and 3 shown with the graphs are the current turn rates and airspeeds of aircraft 1 and 3, respectively. Alphanumeric data such as velocity, G, altitude, specific excess power, acceleration and rate of altitude change are presented for each aircraft beneath the graphs.

Normally, one can analyze the graphs of two aircraft and determine an ideal speed range and set of tactics for each aircraft to exploit weaknesses of the other. However, in the example presented in Figure 6.1-2, aircraft 3 has a superior turning capability over aircraft 1 across all speed regimes. It would be wise for the pilot of aircraft 1 to avoid a turning fight against aircraft 3 because his aircraft is greatly overmatched.
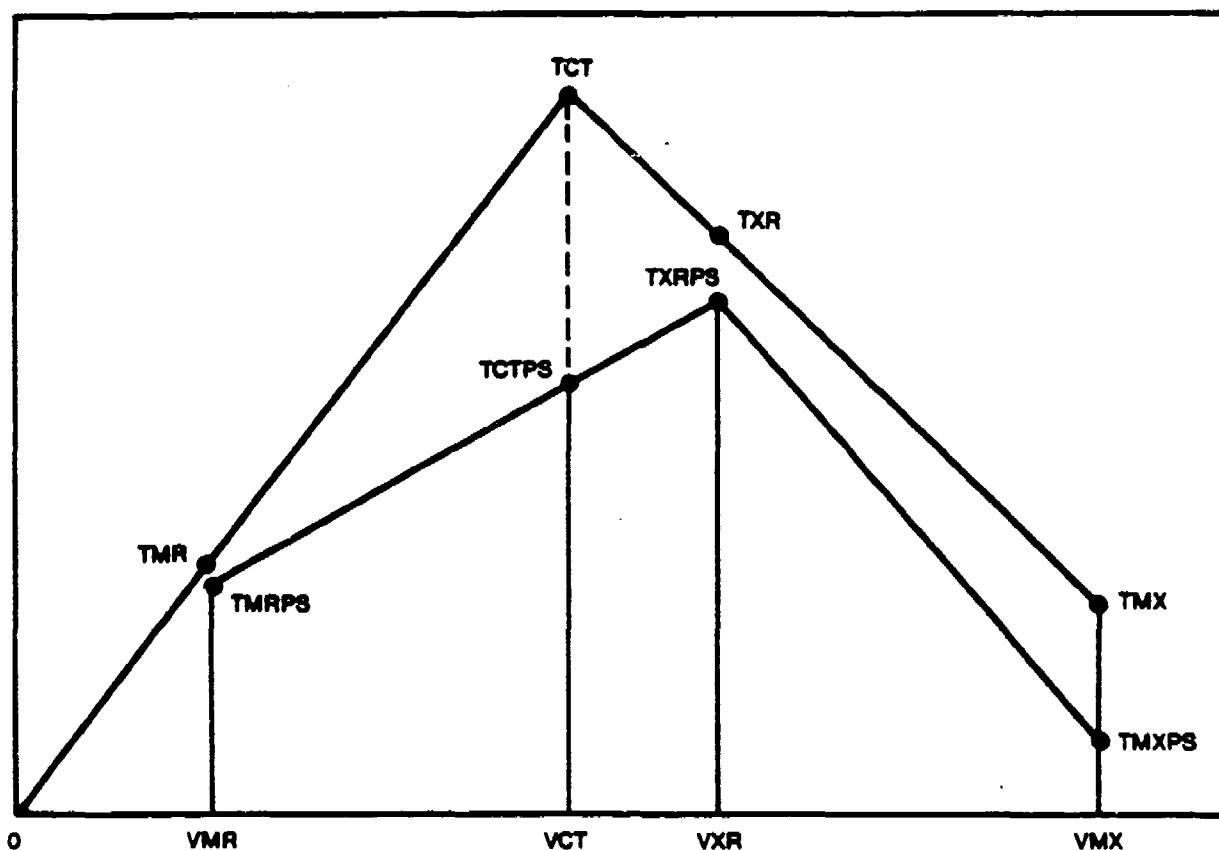
VMR = AIRSPEED AT MINIMUM SUSTAINED TURN RADIUS

VCT = CORNER VELOICTY - AIRSPEED AT MAXIMUM INSTANTANEOUS TURN RATE

VXR = AIRSPEED FOR MAXIMUM SUSTAINED TURN RATE

VMX = 750 KCAS OR MAXIMUM VELOCITY IF LESS THAN 750 KCAS

TMR = MAXIMUM LEVEL TURN RATE AT VMR

TMRPS = MAX SUSTAINED TURN AT VMR

TCT = MAXIMUM LEVEL TURN RATE AT VCT

TCTPS = MAX SUSTAINED TURN RATE AT VCT

TXR = MAXIMUM LEVEL TURN RATE AT VXR

TXRPS = MAX SUSTAINED TURN RATE AT VXR

TMX = MAX LEVEL TURN RATE AT VMX

TMXPS = MAX SUSTAINED TURN RATE AT VMX

**Figure 6.1-1 Display Data Point Definition, Maneuver Triangle
[From Pruitt, Moroney, and Lau 1980]**
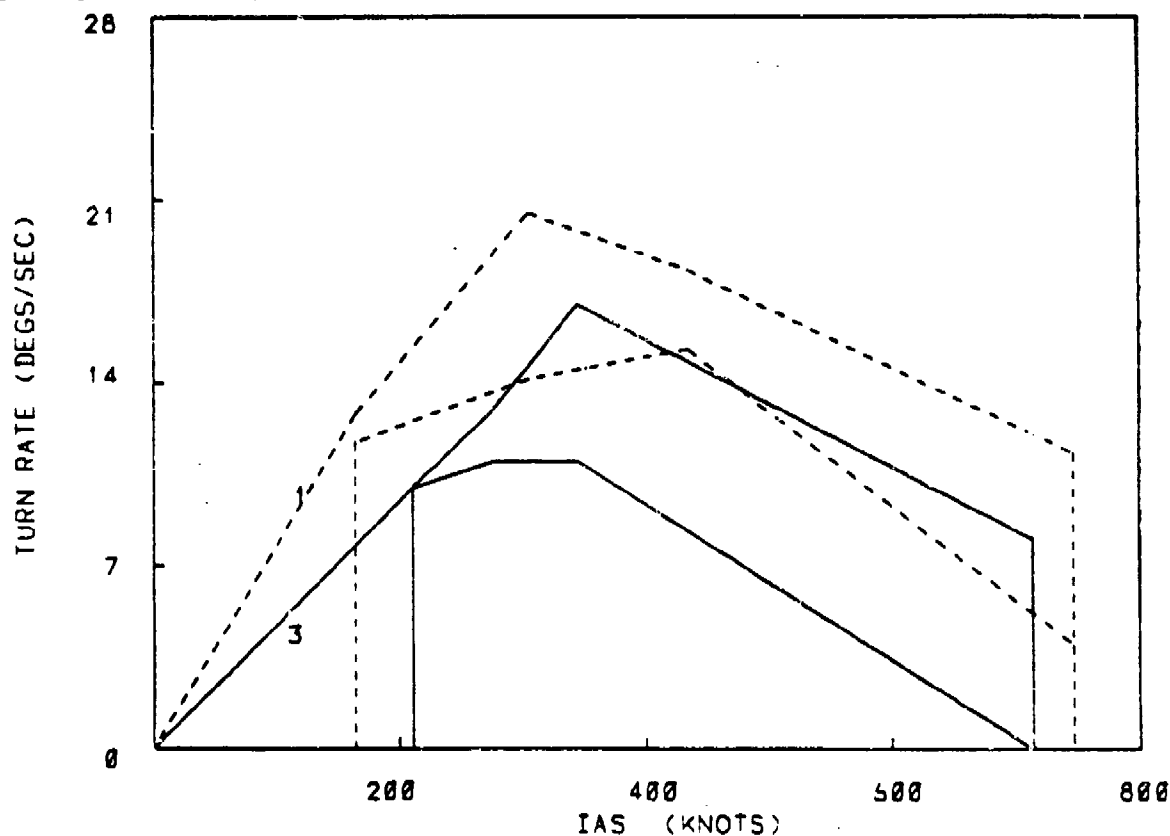
Figure 6.1-2 Illustration of Energy Maneuverability Display

6.2 OPERATIONAL STATUS AND USAGE

The EMD has received limited use on both Navy and Air Force training ranges. The EMD has been used by Navy Fighter Weapons School instructor pilots, at NAS Miramar, California, primarily to ascertain the limits of their aircraft maneuverability envelopes. The Nellis AFB, Nevada, range has completely removed the software from the system to conserve system resources.

The reason for this limited use is due to the nature of the display. The concept of EM is very useful but difficult to apply in flight. The information displayed is not correlated to specific tactical encounters. Pilots must extrapolate the EM data and relate it to the time and position of engaging aircraft, and tactical maneuver attempted.

The software for the present EMD consists of a subroutine within the Control and Computation Subsystem (CCS) that calculates turn rate, and a display driver on the DDS that interprets the CCS output. The software has been modified and migrated from Yuma to Nellis and back to Yuma over the life cycle of the program. The Yuma site has the most current updates and should be considered the field software baseline.

The existing software is poorly documented, and support of the code is dependent upon a few site programmers who have monitored its evolution. The display exhibits periodic digressions when viewed on the DDS. Sometimes the aircraft position indicator is projected outside the triangular performance envelope of the display. An example of this digression is illustrated in Figure 6.1-2 where the current turn rate and airspeed of aircraft 1 are shown outside the boundary of its aerodynamic limit. This problem can be attributed to the fact that aircraft profile data are defined for limited discrete altitudes, and approximations must be made when actual aircraft altitude is between defined altitudes. While more tabled values for altitude could be stored in memory, the increased resource demands would severely burden limited computer resources. To correct this problem, a more powerful computer with larger memory capacity would be required.

A problem with EMD software has been found in which P_S values do not agree with alphanumeric data on airspeed, time rate of change of airspeed and time rate of change of altitude. Correcting this and other potential software problems would be difficult due to the poor documentation. It is recommended that EMD software not be updated at this time due to the limited use of the display, poor documentation and upcoming changes in computer architecture for the next generation ranges. Currently, there are no plans to incorporate the EMD at the latest range at NAS Fallon, Nevada.

6.3 EVALUATION OF THE EMD

In developing the EMD, three candidate diagrams were considered:

1. Altitude versus Mach
2. "g" versus velocity
3. Turn rate versus Mach

The three candidates were reviewed by aircrews who expressed preference for using the turn rate versus Mach diagram in the EMD (Pruitt, et al., 1980).

Despite their preference for this EMD format, however, pilots seldom use the display and it has been taken off the DDS at some sites. While velocity is easy for pilots to

check, there is no turn rate display in the cockpit, and pilots have had difficulty in relating EM diagrams to the actual maneuvers they perform in the aircraft. Just knowing the range of airspeeds where there is a region of favorable P_s values does not appear to give sufficient information to lead maneuvers, resulting in a favorable position with respect to the adversary.

As a final note, a word of caution should be given on the development of tactics using EM diagrams. To obtain a position advantage, a pilot may fly in a region of the Maneuver Triangle where his aircraft has superior performance. If this region is a relatively small portion of the triangle, the pilot is quite restricted in his choice of maneuvers. This restriction may make the pilot's actions predictable, enabling enemy pilots to anticipate his maneuvers and to be prepared with counter maneuvers.

6.4 ALTERNATIVES TO THE CURRENT ENERGY MANEUVERABILITY DISPLAY

A possible alleviation to the difficulty pilots have in relating the turn rate versus Mach number diagram to cockpit controls is to overlay constant "G" and constant turn radius lines onto the graph as was shown in Figure 5.1-2. With respect to the predictability problem, it should be pointed out that it is still useful to have your adversary lose energy faster than you. Thus diagrams such as the turn rate versus Mach number graph are good to have in one's mind as long as they do not become overly restrictive. Rather than have the EM diagram on the DDS for debriefs, it may be more appropriate if it is employed in the classroom stage of instruction, where part-task simulations could be used to demonstrate EM concepts. As discussed in Section 5.0, a simplified metric such as P_s may be most appropriate for pilot feedback. Techniques to display energy and maneuver information are described in the next section.

7.0 ENHANCED DISPLAYS FOR AIRCREW TRAINING

This section presents several enhanced displays which were designed during the course of developing maneuver and energy performance metrics. The displays represent a few of a much larger group of instructional and diagnostic assessment displays which could comprise a TACTS/ACMI instructional enhancements package. Although the display presentation focuses on TACTS/ACMI enhancements, much of the information is directly applicable to the design of instructor operator stations for air combat flight simulators.

Prior to presenting display concepts, a framework for the design and use of enhanced displays is presented. The framework is based upon a simplified aircrew training model which is illustrated in Figure 7.0-1. The model shows the generic phases one must go through to train aircrews to achieve or maintain proficiency in a skill area.

The aircrew training model begins with a specification of training objectives which delineate the task elements to be trained. Tasks are identified from mission and task analysis of the aircrew job. Other important outputs produced with the training objectives are as follows:

1.   Instructional prescriptions for teaching each task
2.   How task performance is assessed
3.   What performance level determines mastery of a task

After these requirements are specified, aircrews enter the training cycle, where they are provided instruction followed by practice on the task. Aircrew performance during practice is assessed to determine if learning has occurred. Diagnostic feedback is an essential element of learning obtained during practice and assessment phases. With knowledge of prior performance in hand, each aircrew receives additional instruction and continued practice until his performance reaches the prescribed level.

The significance of the above framework is that information requirements, which may be unique to each phase of the training cycle, should drive the display design. For example, displays developed for the instruction phase could teach difficult-to-grasp principles involving spatial relationships which are not amenable to a classroom setting. To teach how a weapon envelope is distorted under diverse conditions, an aircrew could be provided a three-dimensional view of a weapon envelope on a desktop display system. The aircrew could observe, directly, envelope distortions as he manipulates the movement of a simulated aircraft with a joystick. This is one example of many that could be developed as part of a future TACTS/ACMI instructional enhancements package. The remaining section focuses on displays suitable for the practice and assessment training phases.

Traditionally, emphasis has been placed on developing displays for the practice phase of air-to-air combat training. A classic example is the TACTS/ACMI DDS. The DDS provides numerous graphic and alphanumeric displays which are monitored by the Range Training Officer (RTO) during live exercises. An illustration of the DDS is presented in Figure 7.0-2. The same display formats are also utilized by instructors and aircrews who observe engagement replays during post-mission debriefs.
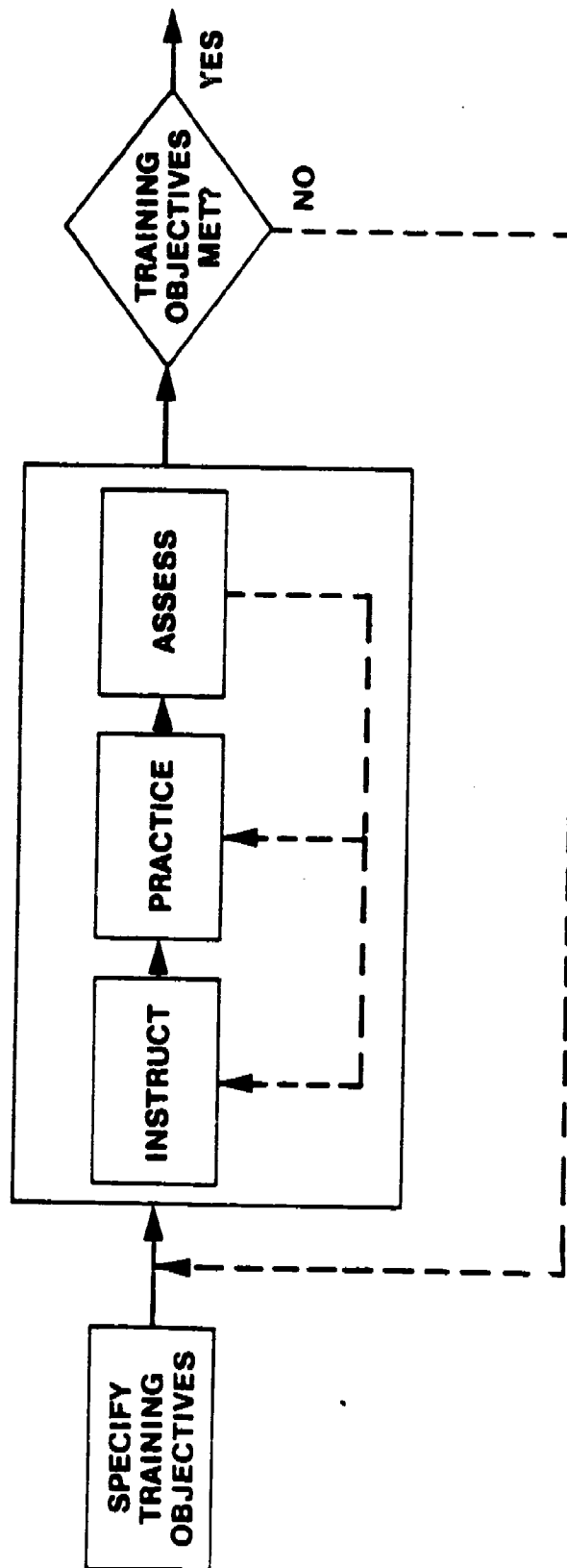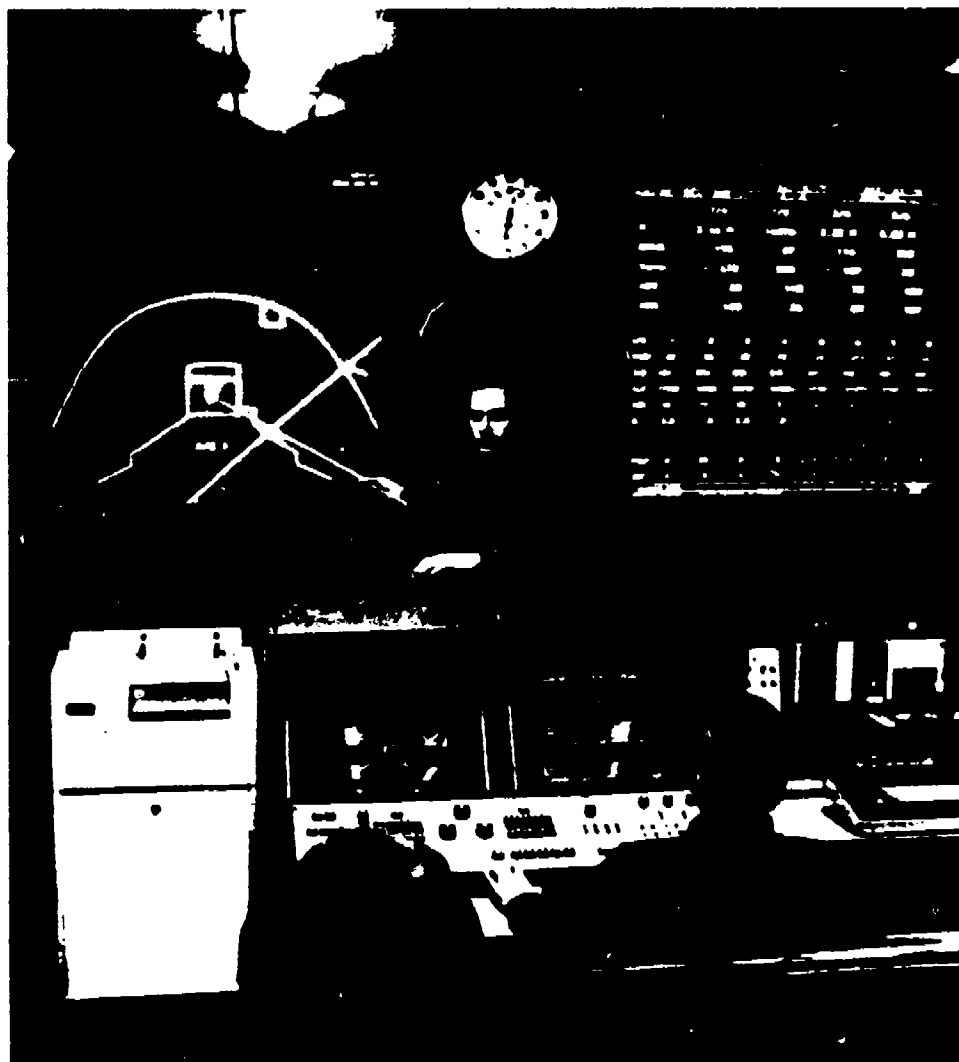
Figure 7.0-1 Aircrew Training Model

Figure 7.0-2 Display and Debriefing Subsystem: (DDS) for TACTS/ACMI

Two frequently used graphic displays on the DDS are the centroid view and pilot cockpit view displays. The centroid view enables a DDS user to view several aircraft simultaneously within the TACTS/ACMI operating area from a 'god's eye' perspective. The centroid view scene can be rotated easily for side viewing. Top and side centroid views on the DDS are presented in Figures 7.0-3 and 7.0-4, respectively. The pilot view display provides the user with the perspective of looking out the cockpit of interest at the adversary aircraft. An illustration of the pilot cockpit view display is presented in Figure 7.0-5. In the illustration, the user can observe aircraft 4 as though he were looking through the cockpit of aircraft 2.

The pilot cockpit view display provides a good pictorial rendition of events during live exercises and replays of engagements flown on TACTS/ACMI. In addition to pilot view display graphics, the DDS operator can view detailed alphanumeric data on each of several alphanumeric displays. One example of an alphanumeric display presented on the DDS is the alphanumeric flight data display (Figure 7.0-6). Although alphanumeric displays can be called up readily, users prefer graphic display formats which they view most frequently. A possible reason for this preference may be that users are required to synthesize raw data contained in alphanumeric displays. Data synthesis is a difficult task which becomes even more complicated when the user must then integrate the synthesized information with graphic information presented dynamically in the pilot view display.

A potential aid to assist in data comprehension may be to synthesize some of the most relevant alphanumeric data for the user and display the synthesized data in meaningful graphic formats. The MI, described previously, is a most suitable candidate because the algorithm synthesizes the raw data parameters AOT, ATA, and R, and produces a summarized output. Moreover, the MI output can be transcribed directly to graphic formats which are compatible with graphic scenes presented in the pilot view display. By utilizing compatible graphic formats, the authors believe that the user may be better able to integrate and comprehend the large quantity of training data generated by TACTS/ACMI and simulators.

Application of the above concepts is illustrated in the enhanced pilot cockpit view display (Figure 7.0-7). The enhancements are shown on the right side of the display. As illustrated in this figure, there are two analog bars. The bar on the left is the MI scale which indicates aircraft position advantage. The open arrowhead symbol ($\triangleright$) to the left of the MI bar, as shown in the illustration, represents the current position advantage of aircraft 2 with respect to aircraft 4. As aircraft maneuver during the course of an engagement, the arrowhead moves against the fixed MI scale. If the arrowhead symbol is at the top of the scale, aircraft 2 would have a significant position advantage. If it is at the bottom of the MI scale, the opposing aircraft (4) would have the position advantage. Neutral or standoff states would be indicated by the location of the arrowhead near the middle of the MI bar.

Located to the right of the MI bar in Figure 7.0-7 is an energy index (EI) bar. The EI may represent E_s, P_s or some other calculated energy parameter. The most desirable energy parameter to use for the EI bar should be determined through additional research and empirical testing. In any event, the EI bar should represent an energy scale with high energy depicted at the top of the bar and low energy at the bottom. The open arrowhead symbol ($\triangleright$) to the left of the EI bar would represent, as shown in Figure 7.0-7, the energy level of aircraft 2. The second symbol ($\succ$) on the left side of the EI bar would represent a desired or ideal energy level which would correspond to prescribed doctrinal standards. The current energy level of aircraft 4 would be

N-W 32

DEAD

· 0
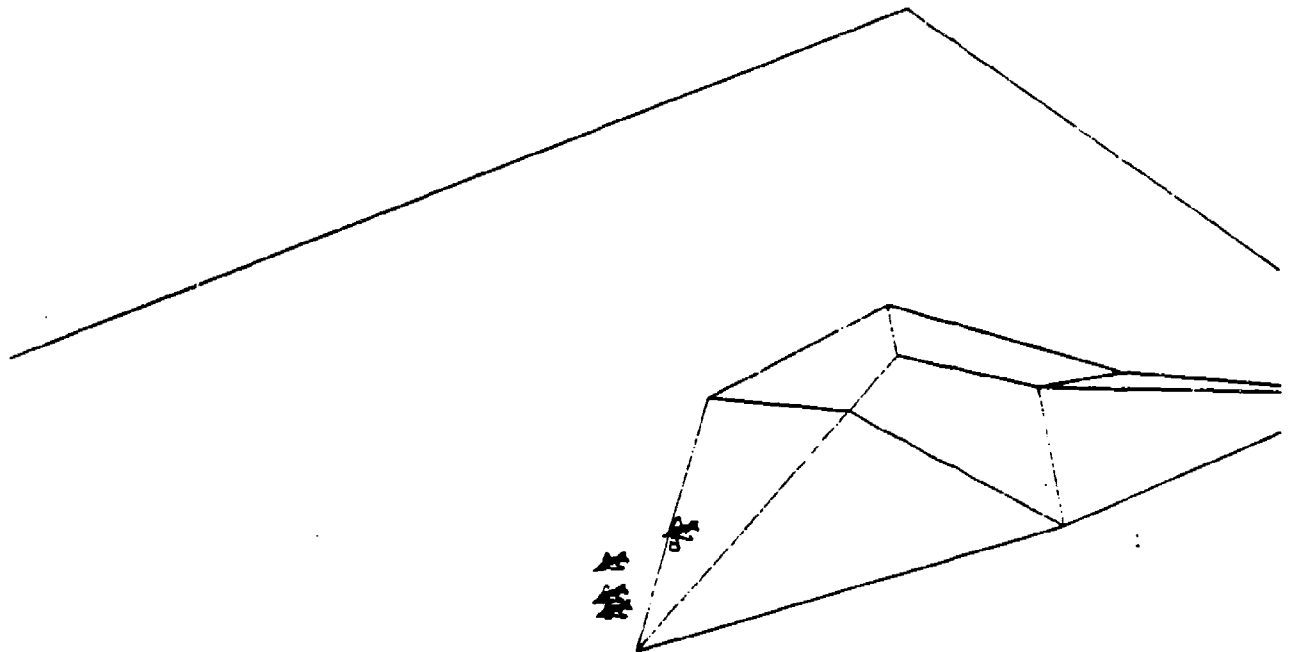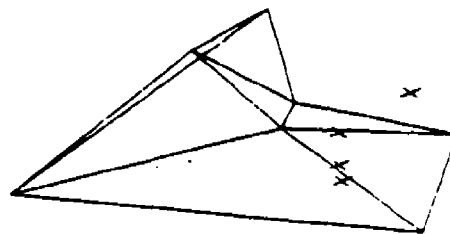
Figure 7.0-3 Display and Debriefing Subsystem: Centroid Top View

W 53   A/C 1   MERRITT   EXCESS G   + 6 3   LIVE
                                                                1250 23 60

                                                                 ACT OFF
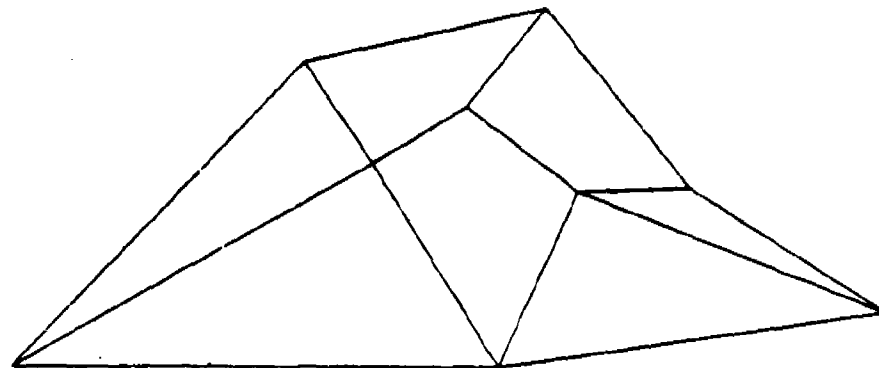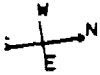
DEAD

1/   1-9L   MISS   / NO TGT /
* 2/   1-9L   MISS   / NO TGT /

'Ø
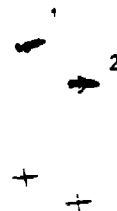
Figure 7.0-4 Display and Debriefing Subsystem: Centroid, Rotated Side View

REPLAY
1200:54:37

A/C 2
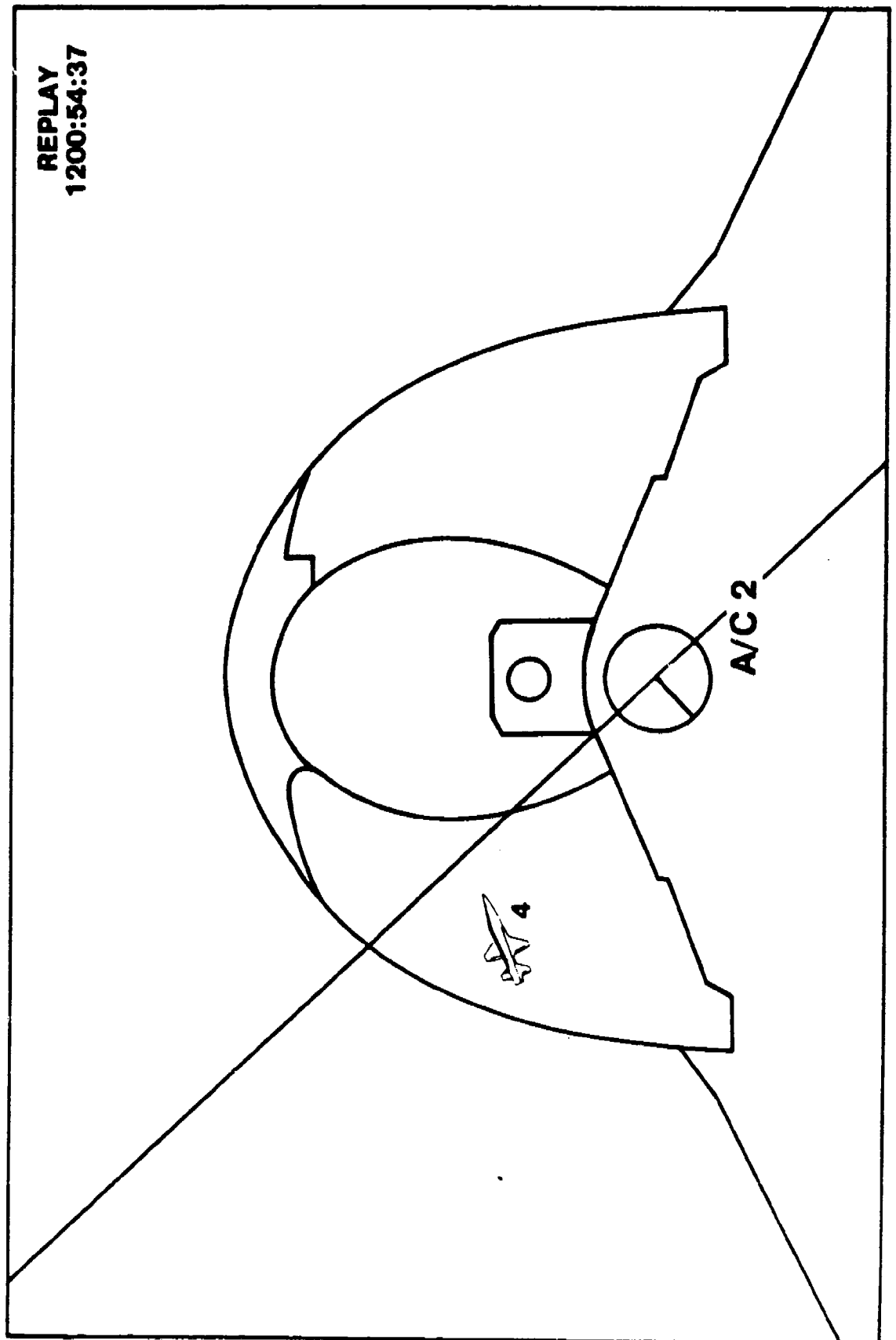
Figure 7.0-5 Display and Debriefing Subsystem: Pilot Cockpit View Display

MA41211-1297		12/11/84			1019 42 80		
CONSOLE MODE	REPLAY / ACTIVE / 6	ACM NO = 1			CONTROL CONSOLE 3A		

		2/6				6/2		
BRG		263				83		
R		1 02 M				1 02 M		
Vc(KTS)		287				287		
AOT		176				61		
ATA		119				4		

A/C	1	2	3	4	5	6	7	8
MACH	00	77	00	00	79	80	1 02	
IAS	0	347	0	0	287	355	474	
ALT	00	22239	00	00	31719	23313	18743	
ACA	0	12	0	0	1 1	1 3	1 1	
G	0	3 1	0	0	2 3	3 5	4 8	
TONE	0	0	0	0	0	0	0	
SEEKER	50	50	50	50	50	4 2	50	
RADAR	50	50	50	50	50	4 2	50	
AIM-9	4	4	4	4	4	4	4	
AIM-7	4	4	4	4	2	2	2	
FLT	1 0 1 U&P	0	1 0 1 U&P	1 0 1 U&P	0	1 3 0	0	
INT	2	1 4	6	2	2	1 4	2	

KEYBOARD INPUT							

Figure 7.0-6 Display and Debriefing Subsystem: Alphanumeric Flight Data
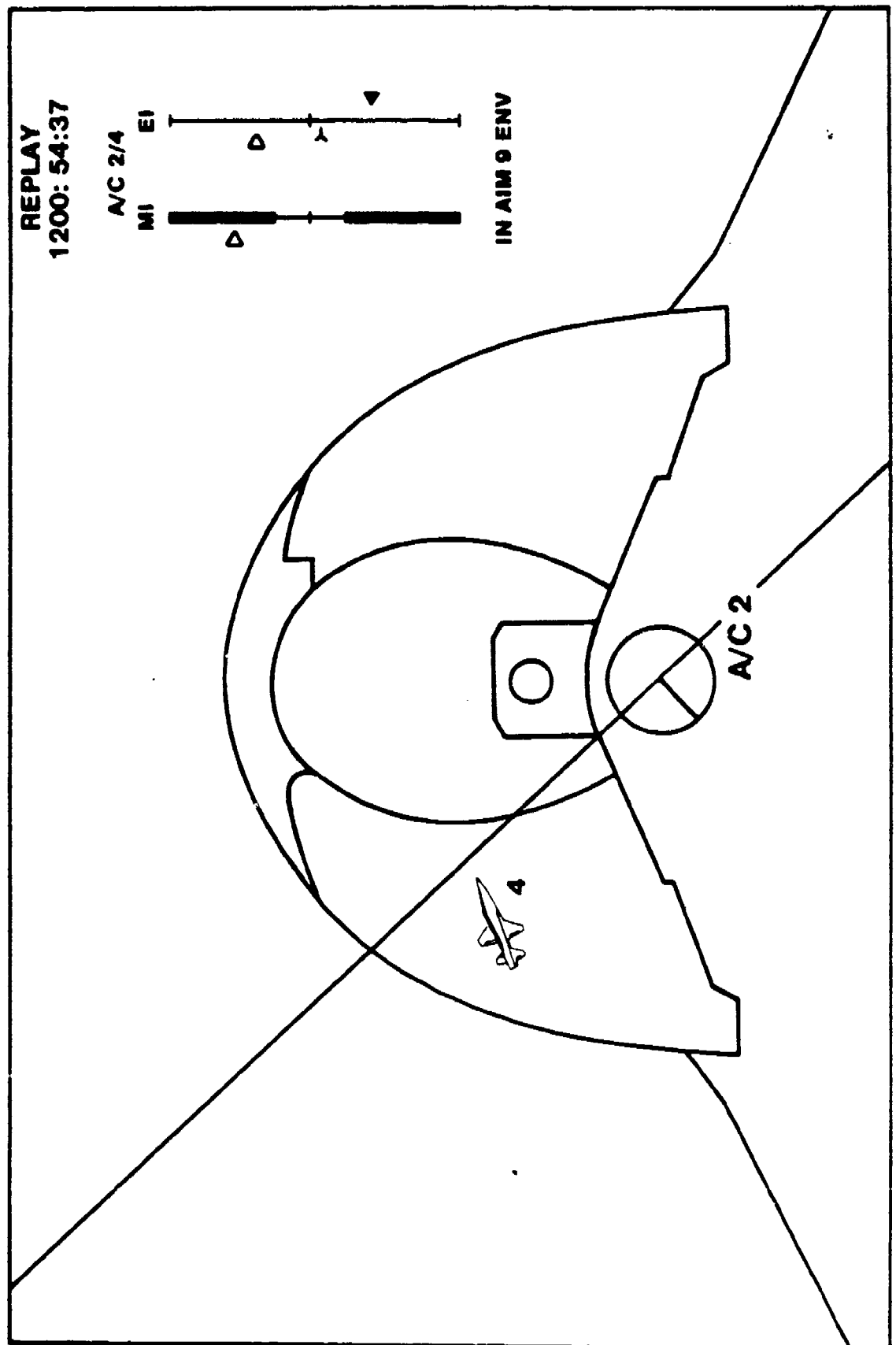
Figure 7.0-7 Display Enhancements to Pilot Cockpit View Display

depicted by the filled arrowhead symbol (◀) located to the right of the EI bar. Together, the EI and MI bars could enable the user to assess aircrew trade-offs of energy for position advantage, and vice versa.

A most important feature provided by the enhancements shown in Figure 7.0-7 is the alphanumeric information presented beneath the MI and EI indicator bars. This information specifies when the adversary is inside a weapon envelope of the fighter and which weapon type is the optimum one to use. Envelope status information is an extremely useful diagnostic. For example, occurrences of a pilot who consistently misses shot opportunities might indicate a need for additional specialized instruction to teach weapon envelope recognition skills.

In actual operation, the user would be able to view the MI, EI, envelope status or other parameters as desired options. Once selected, these enhancements would be displayed as windowed overlays on the basic pilot cockpit view display.

The introduction of high-speed raster graphics systems to TACTS/ACMI would provide a wide range of capabilities for developing innovative display enhancements. An example made possible by this latest technology is the maneuver diagram shown in Figure 7.0-8. This diagram depicts aircraft flight paths and maneuvers flown. A history of the flight path is shown as a ribbon which follows the aircraft. Energy management information could be easily integrated with the maneuver information by coding it into the ribbon of the maneuver diagram. For example, if the aircraft is in a positive energy state, or gaining energy, the ribbon could be coded as a specific color and texture. When an aircraft is in a negative energy state, or losing energy, the ribbon would show a contrasting color and texture. A third color, located between the other two colors on the color spectrum, and a distinct texture could signify when the aircraft is neither gaining nor losing energy. Details of the color scheme will not be elaborated upon since it is a subject for future study. It should be mentioned, however, that redundant color and texture coding is desirable for viewing by color blind operators and with non-color terminals and printers.

* Displays designed to assess training have not been given much emphasis in the past. However, with increased military training requirements and resource constraints, it is becoming increasingly important to have an assessment capability built into training systems. There is virtually no objective means to determine whether learning has occurred without an assessment capability. At the individual level, diagnostic assessment displays could provide the aircrew with a rich source of feedback to determine his level of task mastery. Moreover, the displays could help pinpoint strengths and weaknesses so appropriate training resources could be applied to expedite the learning process.

A key ingredient to a diagnostic assessment capability is a data base management system (DBMS). The DBMS stores historical data collected during the training cycle. Because storage is an obvious limitation to any system, only the most essential data points or measures can be stored. The quantity of data can be reduced by storing only those measures which relate directly to tasks identified from mission and task analysis. Diagnostic assessment displays must then be based upon these measures and be formatted in a simplified manner which is meaningful to aircrews.

Due to time constraints for reviewing exercises during TACTS/ACMI debriefs, it is possible that important events or trends may be overlooked or forgotten. This potential problem may be particularly heightened with large-scale TACTS/ACMI systems. Fortunately, the historical nature of the data makes diagnostic assessment displays well suited for off-line viewing (e.g. squadron location).
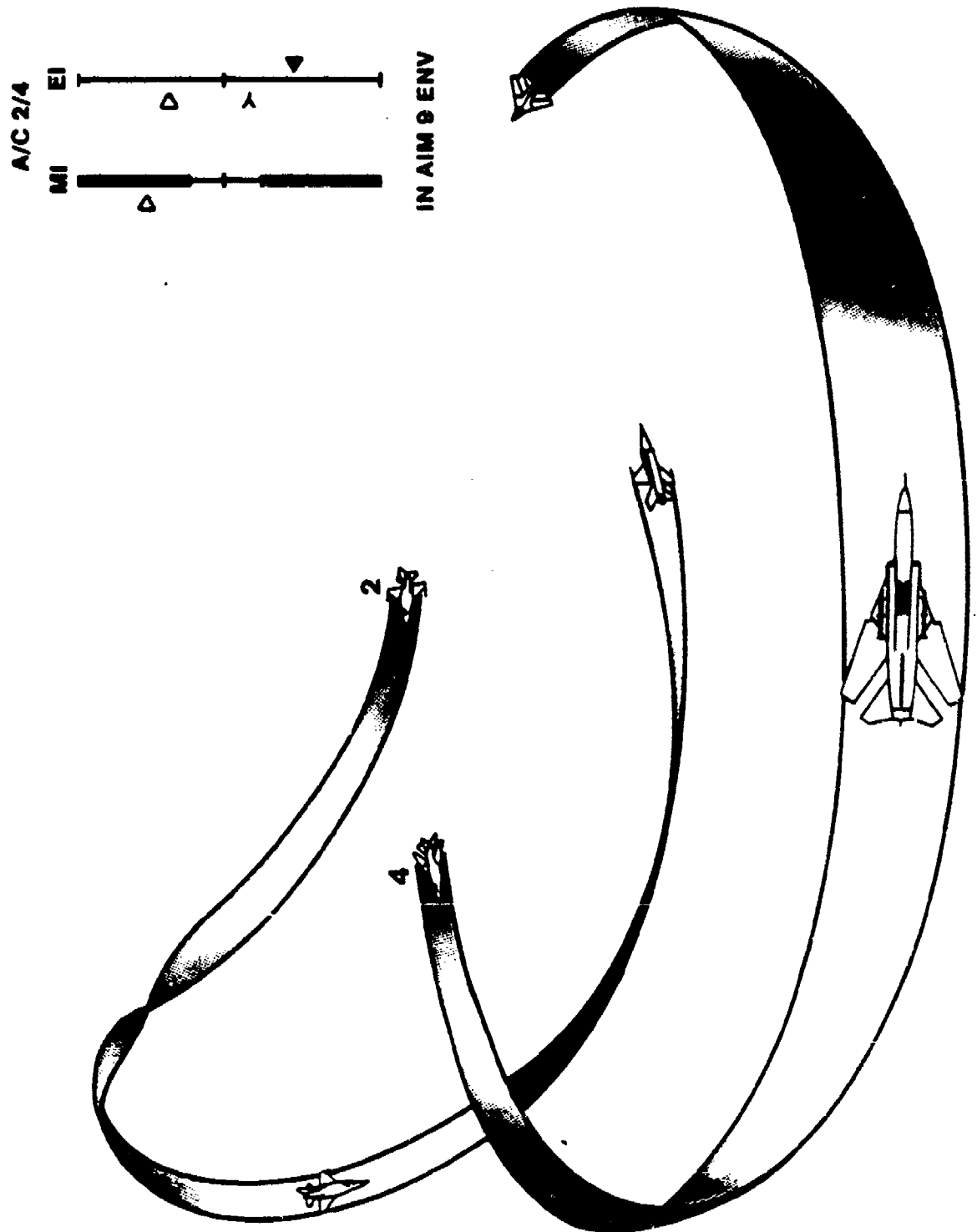
Figure 7.0-8 Maneuver Diagram

An example of a diagnostic assessment display is the MI profile which is illustrated in Figure 7.0-9. The MI profile is basically a plot of the MI over the course of the engagement. MI values are shown along the y-axis, and engagement time is displayed across the x-axis. Superimposed on the graph is weapon envelope information. In the example presented in Figure 7.0-9, times are marked where the fighter maneuvered to get the adversary within his AIM-7F and AIM-9I missile envelopes. Simulated missiles fired during the TACTS/ACMI engagement are indicated near the bottom of the MI profile.

During the engagement shown in Figure 7.0-9, the fighter started in a neutral position, proceeded to a slightly offensive position and then drifted to a defensive position. As the fighter maneuvered into an offensive position, the aircrew fired an AIM-7F missile which was scored as a no-kill. It is important to note that the fighter shot prior to entering the missile envelope. The fighter then began to lose his position advantage to the adversary who subsequently fired a simulated missile and missed. The engagement ended after the fighter obtained a position advantage a second time and successfully fired a simulated AIM-9L missile inside the weapon envelope.

The MI profile provides a historical perspective of the fight and would be a good source of diagnostic feedback to aircrews, particularly with respect to envelope recognition performance. Summarizing MI profiles and plotting them over successive training days would show the aircrew's learning curve. An example of this type of assessment output is illustrated in Figure 7.0-10. The learning curve could show when aircrew performance levels drop off due to extended layoffs. This information could be valuable to determine when to allocate training resources to maintain aircrews at peak readiness levels. Interestingly enough, the learning curve could also be valuable to determine which training methods or devices are actually working.

A final example of an output for use during the assessment phase of training is illustrated in Figure 7.0-11. This output reveals the aircrew performance for several key tasks identified in air-to-air combat. For each task, the aircrew average is shown. The individual's performance is then compared to established training standards, not to other aircrews. The percentage of a training standard accomplished reveals overall training strengths and weaknesses which might require additional training. Although emphasis has been placed in this report on air-to-air missions, a summary out ut, as illustrated in Figure 7.0-11, would be applicable to other missions s h as air-to-surface and electronic combat.
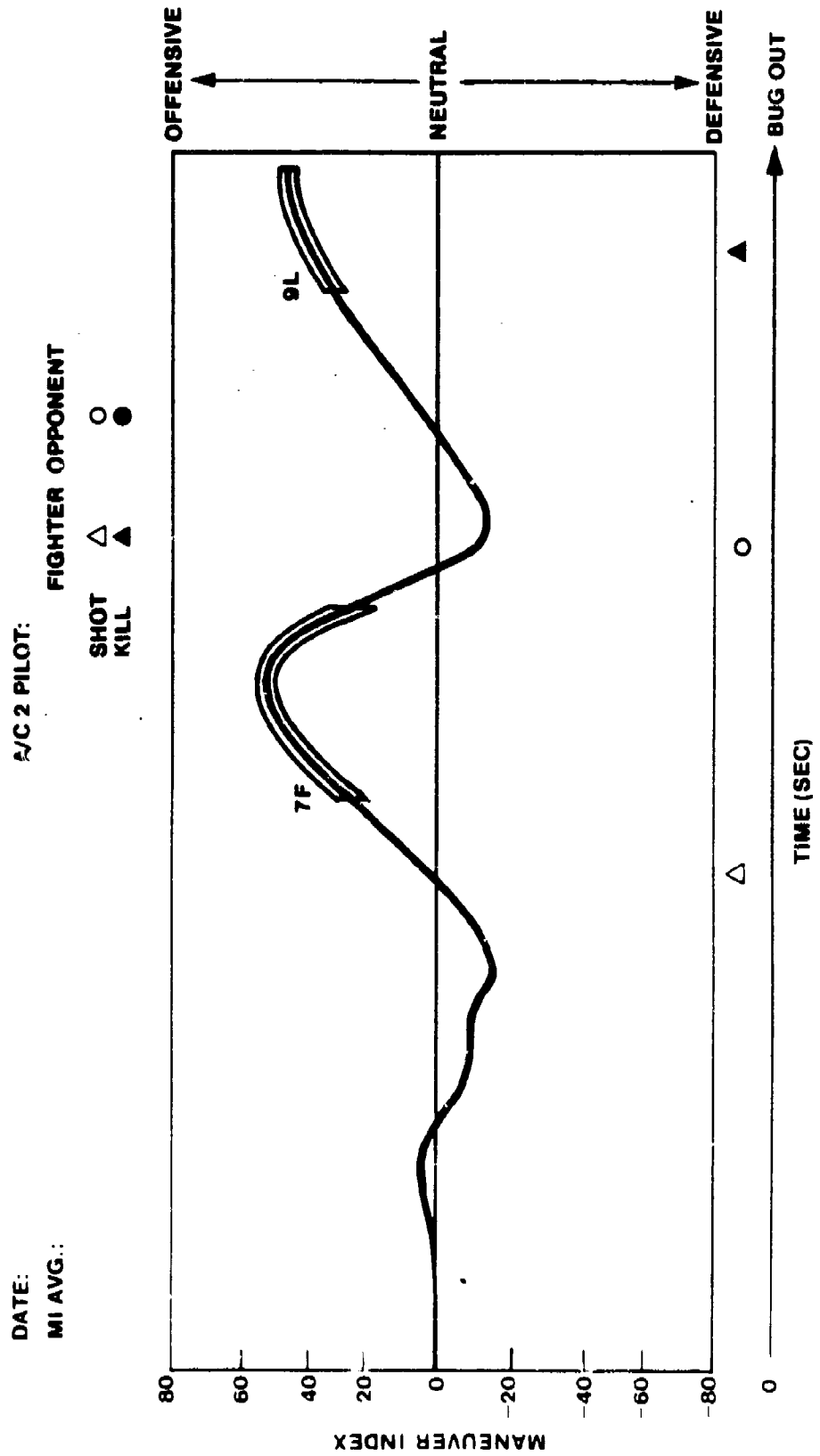
Figure 7.0-9 Diagnostic Assessment: Maneuver Index Profile
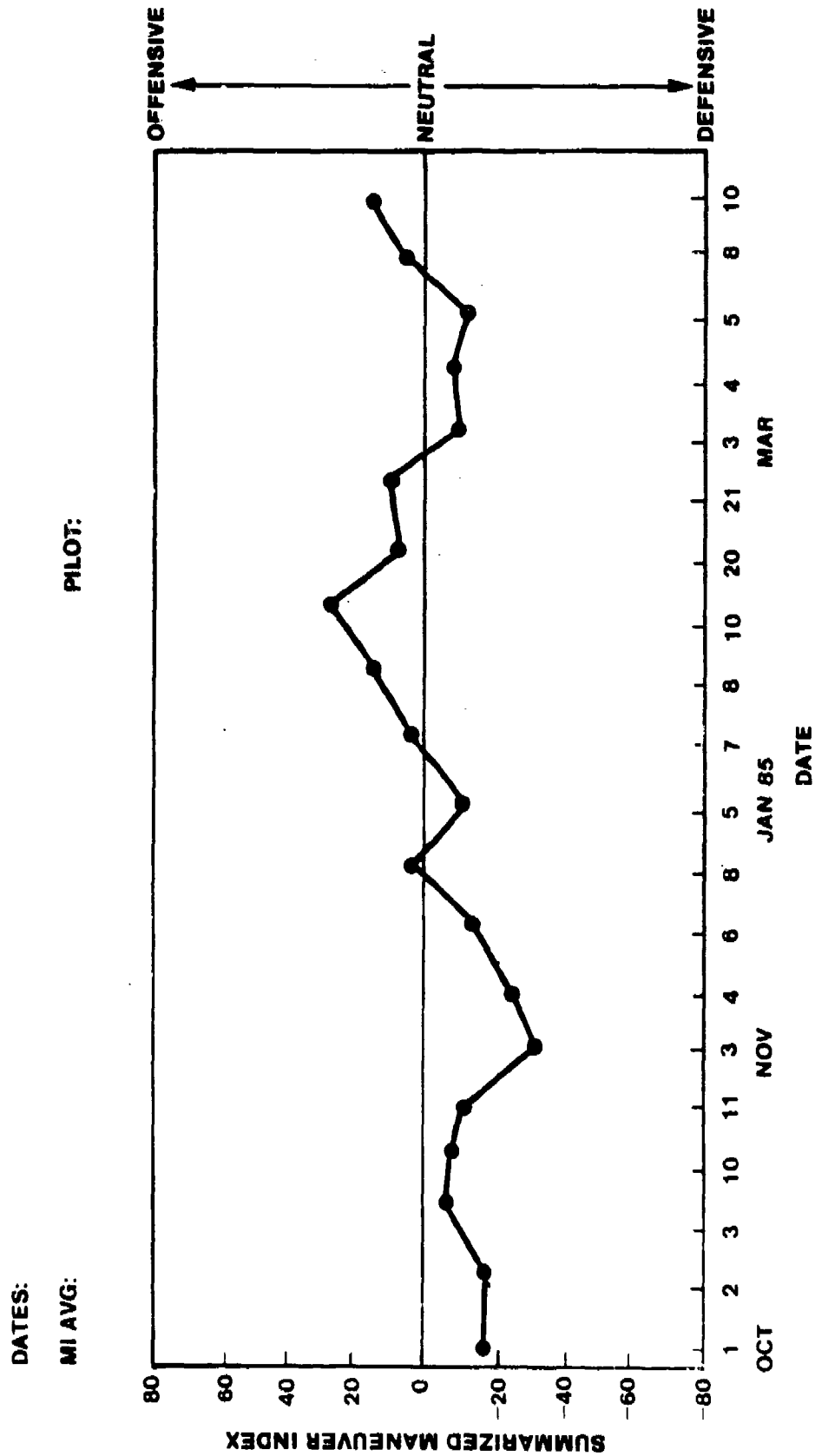
Figure 7.0-10 Diagnostic Assessment: Maneuver Index Profile Learning Curve

Date:

Eng:

Pilot:

Overall %
Stand Accom:

AIR-TO-AIR TASKS	PILOT AVERAGE	TRAINING STANDARD FOR EXPERT/LEVEL	% STANDARD ACCOMPLISHED
SEARCH PROCEDURES			
Radar Contact (%/rng)*			
Tally Ho (%/rng)			
Visual ID (%/rng)			
TACTICS AND MANEUVERS			
1st Shot (%)			
Maneuver Index			
Energy Index			
Missed Opportunities			
WEAPON PERFORMANCE			
Shot Accuracy			
Guns			
AIM 7			
AIM 9			

*Percentage refers to % of engagements that task was successfully completed.

Figure 7.0-11 Diagnostic Assessment: Pilot Performance Compared to Air-to-Air Training Standards

8.0 SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

This report presents technical progress and results from a study contract performed by Cubic Corporation for the Naval Air Systems Command. A primary objective for undertaking the research was to develop up-to-date algorithms for use in assessing aircrew performance during the maneuvering portion of the air-to-air combat engagement conducted within visual range. Another objective was to integrate available energy management metrics with output from the position advantage algorithms.

The report presents a rationale and framework for the development of performance measures. Requirements are outlined for assessing the validity and reliability of performance measures. Applications for the use of performance measures in an aircrew training environment are also described.

Prior to developing performance algorithms, a review and analysis was performed of measurement models that have been developed for air-to-air combat over the past 10 years. Strengths and weaknesses of each model are described. A major weakness of the position advantage models is that they do not adequately accommodate aircraft equipped with all-aspect weapons.

A major breakthrough in the study was the development of a maneuver index. The MI is designed to depict maneuvering performance of a fighter aircraft during vectored air-to-air engagement as flown on TACTS/ACMI. The MI comprises three basic components. An angular geometry component depicts the position advantage of the fighter relative to an adversary and is expressed in terms of direction angle parameters (AOT and ATA). The AGC computation method was adapted from Simpson's (1976) performance index algorithm. The second component of the MI is the weapon range component. This component is essentially a distance measure expressed as deviation or error from the optimum range of a weapon envelope. The third component of the MI is the scale factor which is included to set the origin and end points of the MI scale.

The report presents a historical perspective of energy metrics that have been used by design engineers over the years to compare aircraft performance. While point comparisons of aircraft performance such as maximum air speed, maximum altitude and thrust-to-weight ratio were prevalent up through the 1960's, newer metrics are becoming significant due to the improved maneuvering performance of aircraft. The work by Skow and Hamilton (1984) is described which introduces metrics such as turn agility, energy onset rate and barrel roll rate. Many of these metrics are new and have not been fully tested. It is not known if the new metrics will be available or applicable for use as aircrew training feedback. At this point it is recommended that simplified metrics such as specific energy or specific excess power be used in conjunction with the MI.

The report describes the Energy Maneuverability Display which has been operating at TACTS/ACMI sites the past few years. The EMD is used primarily by Navy instructors who use the display to evaluate tactics when aircraft are flown at the edge of their performance envelope. Unfortunately, the EMD appears to be used very little by aircrews during their daily training exercises. A possible reason for its lack of use

may be the display format which makes a difficult concept even more difficult to understand. Suggestions for improving the EMD format are provided.

Although the EMD could potentially be modified and improved, changes would be very difficult to implement. The EMD has been maintained and updated by operators at different field sites with a different version evolving at each site. This problem is compounded by a lack of complete documentation. The EMD has been removed from some TACTS/ACMI sites because it is expensive in terms of computer memory resources. To revise display formats and update the EMD, it may be necessary to reprogram the software, especially for new-generation systems (e.g. NAS Fallon) which have a different computer architecture.

Various display formats are recommended in the report to improve training feedback for aircrews. Unique displays are designed for viewing during the practice and assessment phases of TACTS/ACMI training. During the practice phase, it is recommended that the pilot cockpit view display be enhanced by the introduction of vertical indicators which can be windowed. The indicators would show aircraft position advantage information based upon the MI and a simplified energy metric such as specific energy or its derivative. A display made possible by high-speed raster graphics technology is the maneuver diagram. This type of graph shows actual aircraft flight paths with energy information color and texture coded on a ribbon which trails each aircraft. Graphs recommended for assessment are based upon cumulative training data. They highlight aircrew learning curves and performance profiles. Due to the historical nature of the data, assessment displays could be viewed at off-line locations, if desired.

It is recommended that a phased research effort be conducted to develop and test the MI, assess the reliability and validity of the metric and to evaluate user acceptance and training utility. Initially, the MI algorithm should be coded to run on a TACTS/ACMI compatible computer. MI outputs could then be tested against engagements stored on TACTS/ACMI tapes. Ideally, performance during the maneuvering portion of the engagement as reflected by the MI should predict discrete engagement outcomes. Next, the graphic display formats described above should be verified by air combat subject matter experts and then developed to provide a proof of concept demonstration. The final phase of the research effort for the air-to-air mission should focus on assessing the validity and reliability of the MI in a training setting. Due to requirements for large sample size, experimental control and repeatability of conditions, validation of the training system should be conducted on a flight simulator. TACTS/ACMI should then be used as a vehicle to test transfer of training from the simulator to "real world" conditions. Following validation, prototype development and testing, the completed performance assessment system should be implemented on existing TACTS/ACMI ranges and air combat training simulators. The authors believe that its implementation at that point would prove to be a considerable asset to aircrews undergoing training, and would provide a useful training effectiveness test bed.

9.0 REFERENCES

Allen, M.J., and Yen, W.M. (1979). Introduction to measurement theory. Monterey, CA: Brooks - Cole Publishing Co.

American Psychological Association, American Educational Research Association, and National Council on Measurement in Education. (1974). Standards for educational and psychological tests. Washington, D.C.: American Psychological Association.

Baker, R.T., Mednick, B.R., and Brock, W. (1980). An application of causal modeling techniques to prospective longitudinal data bases. Unpublished manuscript. Los Angeles: University of Southern California, Center for Interactional Research in Human Development and Learning.

Benson, J., and Clark, F. (1982). A guide for instrument development and validation. American Journal for Occupational Therapy, 36, 12, 789-800.

Bittner, Jr., A.C., Carter, R.C., Kennedy, R.S., Harbeson, M.M., and Krause, M. (1984). Performance evaluation tests for environmental research (PETER): the good, bad and ugly. In Proceedings of the Human Factors Society 28th Annual Meeting. Santa Monica, CA: Human Factors Society.

Blaiwes, A.S, Puig, J.A., and Regan, J.J. (1973). Transfer of training and the measurement of training effectiveness. Human Factors, 15, 6, 523-533.

Bloom, B. (Ed.) (1972). Taxonomy of educational objectives: The classification of educational goals, Handbook I: Cognitive domain. New York: David McKay.

Burden, R.L., Faires, J.D., and Reynolds, A.C. (1981). Numerical analysis (2nd ed.). Boston: PWS Publishers.

Campbell, D.T., and Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

Ciavarelli, A.P. (1982). Methodology to assess in-flight performance for air-to-air combat training. In Proceedings of the 4th Interservice/Industry Training Equipment Conference: Volume I. Washington, D.C.: National Security Industrial Association.

Ciavarelli, A.P., Pettigrew, K.W., and Brictson, C.A. (1981a). Development of a computer-based air combat maneuvering range debrief system (NAVTRAEQUIPCEN 78-C-0136-1). Orlando: Naval Training Equipment Center.

Ciavarelli, A.P., Williams, A.M., and Pettigrew, K.W. (1981b). Performance assessment and appraisal system (PAAS) (NAVTRAEQUIPCEN 80-D-0011-1). Orlando: Naval Training Equipment Center.

Cronbach, L.J. (1971). Test validation. In R.L. Thorndike (Ed.), Educational measurement (2nd ed.). Washington, D.C.: American Council on Education.

Cronbach, L.J., and Meehl, P.E. (1955). Construct validity in psychological tests. Psychological Bulletin, 92, 281-302.

REFERENCES (Continued)

Ebel, R. (1979). Essentials of educational measurement (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Fitzpatrick, A.R. (1983). The meaning of content validity. Applied Psychological Measurement, 7, 1, 3-13.

Fleishman, E.A. (1967). Performance assessment based on an empirically derived task taxonomy. Human Factors, 6, 349-366.

Gagne, R.M., and Briggs, L.J. (1979). Principles of instructional design (2nd ed.). New York: Holt, Rinehart and Winston.

Glaser, R., and Klaus, D.J. (1971). Criterion referenced measurement. In M.D. Merrill (Ed.), Instructional design readings. Englewood Cliffs, NJ: Prentice-Hall.

Gronlund, N.E. (1976). Measurement and evaluation in teaching (4th ed.). New York: Macmillan Publishing Co., Inc.

Gunston, B., and Spick, M. (1983). Modern air combat. New York: Crescent Books.

Harrow, A. (1972). Taxonomy of psychomotor domain. New York: David McKay.

Hocevar-Page, S., and Hocevar, D. (1982). Use of confirmatory factor analysis in evaluating factorial validity. Paper presented at the conference of the Academy of Management.

Jones, M.B., Kennedy, R.S, and Bittner, Jr., A.C. (1981). A video game for performance testing. American Journal of Psychology, 94, 1, 143-152.

Krathwohl, D., Bloom, B., and Bertram, M. (1972). Taxonomy of educational objectives: The classification of educational goals, Handbook II: Affective domain. New York: David McKay.

Kress, G., and Breidenbach, S.T. (1983). A performance assessment and appraisal system for air combat training. Orlando: Naval Training Equipment Center.

Livingston, S.A. (1971a). A classical test-theory of criterion-referenced measures. Paper presented at the annual meeting of the American Educational Research Association.

Livingston, S.A. (1971b). The reliability of criterion-referenced measures. Paper presented at the annual meeting of the American Educational Research Association.

Lovett, H.T. (1977). Criterion-referenced reliability estimated by ANOVA. Educational and Psychological Measurement, 37, 21-29.

Martin, T., Luter, M., Caddy, M., and Mueller, W. (1984, September). A review of air combat maneuvering performance presentations (Update 1) (NADC-84043-60). Washington, D.C.: Department of the Navy, Naval Air Systems Command.

REFERENCES (Continued)

McGuinness, J., Forbes, J.M., and Rhoads, J.E. (1983, September). Design of an air combat maneuvering performance measurement system for the simulator for air-to-air combat and the air combat maneuvering instrumentation range. Alexandria, VA: Person - System Integration, Ltd.

Mehrens, W., and Lehman, L. (1973). Measurement and evaluation in education and psychology (2nd ed.). New York: Holt, Rinehart and Winston.

Merrill, M.D. (1972). Psychomotor taxonomies, classifications and instructional theory. In R.N. Singer (Ed.), The psychomotor domain: movement behavior. Philadelphia: Lee and Fabinger.

Messick, S. (1975). The standard problem: Meaning and values in measurement and evaluation. American Psychologist, 30, 955-966.

Nunnally, J.C. (1975). Psychometric theory - 25 years ago and now. Educational Researcher, 4, 10, 7-20.

Oberle, R.A. (1974, November). An air combat maneuver conversion model (Report CRC 274). Washington, D.C.: Office of Naval Research.

Oberle, R.A. (1983, May). Air combat evaluation: The reduced dimension measures (RES 83-6-2). Orlando: Naval Training Equipment Center.

Pedhazur, E.J. (1982). Multiple regression in behavioral research: Explanation and prediction (2nd ed.). New York: Holt, Rinehart and Winston.

Popham, W.J. (1978). Criterion-referenced measurement. Englewood Cliffs, NJ: Prentice-Hall.

Pruitt, V.R. (1979, April). Energy management training aid for the Navy's air combat maneuvering range (ACMR) (Contract N00123-78-C-1371). St. Louis: McDonnell Aircraft Co.

Pruitt, V.R., Moroney, W.F., and Lau, C. (1980, August). Energy maneuverability display for the air combat maneuvering range/tactical training system (ACMR/TACTS). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Robinson, J.C., Drinnon, D.D., Eubanks, J.L. and Eddowes, E.E. (1981, April). Evaluation of pilot air combat maneuvering performance changes during TAC aces training. Nellis AFB, NV: Tactical Air Command, Tactical Fighter Weapons Center.

Roscoe, S.N. (Ed.) (1980). Aviation psychology. Ames: The Iowa State University Press.

Simpson, W.R. (1976, July). Development of a time-variant figure of merit for use in analysis of air combat maneuvering engagements (TM 76-1 SA). Patuxent River, MD: Naval Air Test Center.

REFERENCES (Continued)

Simpson, W.R., and Oberle, R.A. (1977, June). The numerical analysis of air combat engagements dominated by maneuvering performance (TM 77-2 SA). Patuxent River, MD: Naval Air Test Center.

Skow, A.M., and Hamilton, W.L. (1984, September). Advanced fighter agility metrics (TR 84-05). Torrance, CA: Eidetics International.

Swezey, R.W. (1978). Aspects of criterion - referenced measurement in performance evaluation. Human Factors, 20, 2, 169-178.

Tenopyr, M.L. (1975). Content - construct confusion. Personnel Psychology, 30, 677-680.

Vreuls, D., and Wooldridge, L. (1977). Aircrew performance measurement. In J. Pope and D. Meister (Eds.), Proceedings of symposium on productivity enhancement and personnel assessment in Navy systems. San Diego: Navy Personnel Research and Development Center.

Weiss, D.J., and Davison, M.L. (1981). Test theory and methods. Annual Review of Psychology, 32, 629-658.

Wooldridge, L., Kelly, M.J., Obermayer, R.W., Vreuls, D., Nelson, W.H., and Norman, D.A. (1982, October). Air combat maneuvering performance measurement state space analysis (NAVTRAEQUIPCEN III-342/AFHRL-TR-82-15). Brooks AFB, TX: HQ Air Force Human Resources Laboratory.

APPENDIX A

TACTS/ACMI DESCRIPTION

APPENDIX A
TACTS/ACMI DESCRIPTION

The Navy's TACTS enables aircrews to monitor various air combat exercises in real time. Through its replay capability, TACTS permits debrief and evaluation of pilot tactics, maneuvers and weapon delivery accuracy. (The U.S. Air Force second-generation version of this system is referred to as Air Combat Maneuvering Instrumentation (ACMI).) Four major subsystems comprise the TACTS/ACMI System:

o   **Airborne Instrumentation Subsystem (AIS)** - A pod attached to the aircraft which measures flight dynamics information, senses weapon firing signals, and transmits data to the ground through the Tracking Instrumentation Subsystem (TIS).

o   **Tracking Instrumentation Subsystem (TIS)** - A series of unmanned remote tracking stations communicating with a master tracking station to monitor AIS-equipped aircraft in a specified airspace.

o   **Control and Computation Subsystem (CCS)** - Converts data received from the TIS into suitable form for display.

o   **Display and Debriefing Subsystem (DDS)** - Serves as a control center and display station.

Figures A-1 and A-2 illustrate the major TACTS/ACMI subsystems and their inter-relationships. Some of the more important training features of TACTS/ACMI follow.

o   Real-time tracking including position, velocity, acceleration, attitude, and angular rate measurement of aircraft engaged in air combat training

o   Tape playback of flight history data, complete with pictorial display of the air-to-air engagement and voice transmissions

o   Both digital and graphic hard-copy printouts of flight data, aircraft state vector, cockpit view of engaged aircraft, and mission summary data

o   Computer-generated results of weapon firing.

In addition to air-to-air combat training capabilities of TACTS, engineering upgrades have been made to provide ground attack training missions. Some of these capabilities include No-Drop Bomb Scoring, electronic warfare, and No-Drop Mine Laying. Thus, TACTS has evolved into a multi-mission range instrumentation system capable of providing training across the entire spectrum of air warfare activities.

AIS — Airborne Instrumentation Subsystem
- Pod Physically Similar to AIM-9, Carried on Aircraft
- Transmission Link Air-to-Ground, Ground-to-Air

TIS — Tracking Instrumentation Subsystem
- Multilateration Equipment
- One Master Station
- Seven Interrogator Stations

CCS — Control and Computation Subsystem
- Computers
  - Aircraft State Vector
  - Weapon Trajectory Simulation
  - Aircraft Status

DDS — Display & Debriefing Subsystem
- Display Consoles
- Graphics and Alphanumerics Displays
- Live and Replay Capability
- Projection Large Screen Displays

Figure A-1 Tactical Aircrew Combat Training System (TACTS)

AIRBORNE INSTRUMENTATION SUBSYSTEM (POD)

DISPLAY & DEBRIEFING SUBSYSTEM

CONTROL & COMPUTATION SUBSYSTEM

TRACKING INSTRUMENTATION SUBSYSTEM MASTER

TRACKING INSTRUMENTATION SUBSYSTEM INTERROGATOR STATION

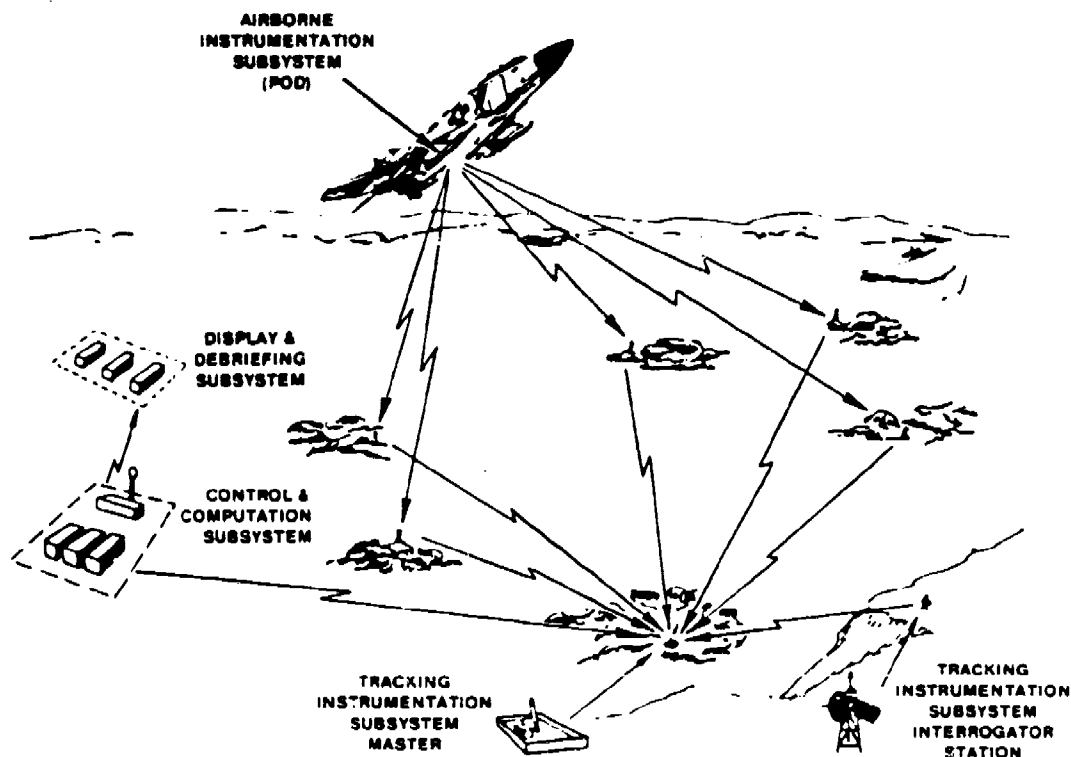**Figure A-2 General Configuration of TACTS Subsystems**

APPENDIX B
GLOSSARY OF AIR COMBAT TERMS

APPENDIX B
GLOSSARY OF AIR COMBAT TERMS

AAMI	All - Aspect Maneuvering Index
ACM	Air Combat Maneuvering
ACMI	Air Combat Maneuvering Instrumentation
ACM State	A descriptor of the ACM situation as offensive, defensive, and neutral positions
AB	After Burner
AIM	Air Intercept Missile
ALT	Altitude
AOA	Angle of Attack
AOT	Angle Off Tail; angle between longitudinal axis of target and line of sight from target to fighter in wing plane of target aircraft, measured in degrees
ATA	Angle between longitudinal axis of fighter and target aircraft, measured in degrees (pointing angle)
Bogey	A term applied to an ACM opponent (suspected unfriendly or adversary)
Bugout	Aircraft leaving arena of engagement or attempting to terminate fight (ie, escape from bogey)
Contact	A call made by an aircrew member (pilot or RIO) upon obtaining radar contact with a target
Corner Velocity	Velocity which corresponds to the maximum instantaneous turn rate
DDS	Display and Debriefing System
Defensive	An engagement state in which a particular aircraft is in a threatened position according to specific mathematical rules
EMD	Energy Maneuverability Display
Energy	Kinetic and potential energy state of aircraft engaged in air-to-air combat; can be defined in terms of IAS and cornering trade-off (kinetic), fuel state and altitude (potential)
Energy Management	Relates to the efficient use of potential and kinetic energy, including stored energy from fuel, to attain specific mission objectives

APPENDIX B (Continued)
GLOSSARY OF AIR COMBAT TERMS

Energy Maneuverability
Capability of aircraft under dynamic flight conditions to perform a change or combination of changes in direction, altitude, and airspeed, expressed in terms of energy and energy rate

Energy Onset Rate
$\Delta P_s/\Delta\tau$; the increment in specific excess power between the maximum power, minimum drag configuration and the minimum power, maximum drag configuration divided by the time to go from minimum to maximum power and from maximum to minimum drag

Engaged Fighter
A fighter whose primary responsibility is to kill or control bogey. It should be in an offensive position

Envelope
Weapon boundary limits within which a missile or guns should be fired. An envelope is defined in terms of distance (range) and angles off tail (degrees) between shooter aircraft and target

E_s
Specific Energy; sum of potential and kinetic energies divided by aircraft gross weight

Fox 1
Call made by aircrew member (usually pilot) indicating that a Sparrow (AIM 7) missile has been fired

Fox 2
Call made by pilot indicating that a Sidewinder (AIM 9) has been fired

G
A unit of force acting on a body being accelerated; unit is equal to the gravitational force applied to the object at the earth's surface (eg 3G's = 3 times the object's weight)

g
Normal Acceleration; measured in units of $g = 322$ ft/sec^2

GCI
Ground Control Intercept

IAS
Indicated Airspeed; airspeed for aircraft measured in knots

Lock-on
Electronically locking the radar system on a particular target

Mach Number
Ratio of the speed of an object to the speed of sound in the surrounding atmosphere

MC Model
Model for estimaticn of position advantage using a semi-Markov process

Neutral
An engagement state in which a particular aircraft is in neither an advantaged or disadvantaged state, according to specific mathematical rules

n_w
Wind axis load factor; n_w = sum of force normal to wind axis divided by gross weight

APPENDIX B (Continued)
GLOSSARY OF AIR COMBAT TERMS

Offensive   An engagement state in which a particular aircraft is threatening an opponent (see above defensive, neutral)

1v1   An engagement involving one friendly versus (v) one bogey aircraft

PI   Performance Index; a metric for position advantage based on the product of continuous functions of AOT, ATA, Range and closing velocity

Placard Limit   Maximum velocity of an aircraft in level flight

PMI   Performance Measurement Index

P_s Specific Excess Power; change in the sum of kinetic plus potential energy divided by time (first derivative of E_s)

Range or R   Distance in feet or nautical miles (nm) between fighter and aircraft

RIO   Radar Intercept Officer

RTO   Range Training Officer

SEAM   Sidewinder Extended Acquisition Mode

Section   Two aircraft that fight as coordinated unit in an air-to-air engagement

TACTS   Tactical Aircrew Combat Training System

Tally Ho   A call made by an aircrew member upon obtaining visual contact with a target

Turn Agility   Turn rate divided by the time required for a complete bank angle change of 90 degrees at that turn rate

Turn Rate   Ability of aircraft to turn expressed in degrees of arc per second

2v1   An air-to-air engagement involving two friendly aircraft versus (v) one bogey aircraft

2v2   An air-to-air engagement involving two friendly aircraft versus (v) two bogey aircraft

UHF   Ultra High Frequency communication channel for radio transmission between aircraft

Vc   Closing Velocity, positive or negative, between fighter and target aircraft, measured in knots

VTAS   Visual Target Acquisition System (helmet mounted gunsight used to slave and point weapon seeker)

APPENDIX B (Continued)
GLOSSARY OF AIR COMBAT TERMS

Wingman   Second aircraft in flight section; also referred to as "wingy"; see
                 definition for "Section"