AL-TP-1992-0001

AD-A247 426

TACTICAL STATE AND PERFORMANCE ASSESSMENT DURING AIR COMBAT MANEUVERING

William B. Raspotnik
Jeffrey L. Leeds

Logicon, Incorporated
Tactical and Training Systems Division
4010 Sorrento Valley Boulevard
P.O. Box 85158
San Diego, CA 92138-5158

Wayne L. Waag

HUMAN RESOURCES DIRECTORATE
AIRCREW TRAINING RESEARCH DIVISION
Williams Air Force Base, AZ 85240-6457

February 1992

Final Technical Paper for Period September 1986 - December 1991

Approved for public release; distribution is unlimited.

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5000

ARMSTRONG LABORATORY

NOTICES

WAYNE L. WAAG
Contract Monitor

DEE H. ANDREWS, Technical Director
Aircrew Training Research Division

W. M. KORNOVICH, JR., Lt Col, USAF
Chief, Aircrew Training Research Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
February 1992

3. REPORT TYPE AND DATES COVERED
Final – September 1986 – December 1991

4. TITLE AND SUBTITLE
Tactical State and Performance Assessment During Air Combat Maneuvering

5. FUNDING NUMBERS
C – F33615-86-C-0012
PE – 62205F
PR – 1123
TA – 35
WU – 06

6. AUTHOR(S)
William B. Raspotnik
Jeffrey L. Leeds
Wayne L. Waag

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Logicon, Incorporated
Tactical and Training Systems Division
4010 Sorrento Valley Boulevard
P.O. Box 85158
San Diego, CA 92138-5158

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)
Armstrong Laboratory
Human Resources Directorate
Aircrew Training Research Division
Williams Air Force Base, AZ 85240-6457

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AL-TP-1992-0001

11. SUPPLEMENTARY NOTES

Armstrong Laboratory Technical Monitor: Dr. Wayne L. Waag, (602) 988-6561, Ext. 113.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The purpose of this research was to investigate the feasibility of developing a diagnostic measure of tactical decision-making skill in the context of air combat maneuvering (ACM). Two experiments were conducted. Both employed a Tactical State Model (TSM) to characterize position and used a 7-point scale to rate tactical performance. Subjects were experienced F-15 and F-16 instructor pilots. In each experiment, they viewed two prerecorded ACM engagements and used the TSM to assign a tactical position or "state" to the designed aircraft. Subjects then rated the performance of the designated aircraft. Results indicated that a tactical state model can be used reasonably well by experts to characterize position during an air-to-air engagement. Relationships were also established for one of the engagements in Experiment Two between these state assignments and an objective measure of positional advantage. The results failed to establish sufficient reliability for the tactical performance rating scale developed by this work. Also, the results failed to establish a consistent relationship link between changes in position and quality of performance. It is concluded that development of a diagnostic measure of tactical decision making will require a different approach.

14. SUBJECT TERMS

Air combat engagement	Performance assessment	Tactical performance
Air combat maneuvering	Performance measurement	Tactical state
Decision making	Tactical decision making	Tactical state model

15. NUMBER OF PAGES
70

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

CONTENTS

SUMMARY 1

INTRODUCTION 2
Background 2
Rationale and Objectives for Present Investigation 3
ACM Model Selection 3
Tactical Performance Rating 5

EXPERIMENT ONE 5
Objectives 5
Method 6
Results 8
Discussion 11
Implications for Experiment Two 12

EXPERIMENT TWO 14
Objectives 14
Method 14
Results 15
Discussion 29

GENERAL DISCUSSION 34
Tactical State Characterization 34
Tactical Performance Ratings 35
Changes in Tactical State and TPR Assignments 36

CONCLUSIONS 37

REFERENCES 38

APPENDIX A 41

APPENDIX B 45

APPENDIX C 51

APPENDIX D 55

TABLES

Table No.

1. Interrater Correlations of TSM Assignments: Experiment One, Engagement One 9
2. Interrater Correlations of TSM Assignments: Experiment One, Engagement Two 9
3. Interrater Correlations of TPR Assignments: Experiment One, Engagement One 10
4. Interrater Correlations of TPR Assignments: Experiment One, Engagement Two 10
5. Interrater Correlations of TSM Assignments: Experiment Two, Engagement One 16
6. Interrater Correlations of TSM Assignments: Experiment Two, Engagement Two 16
7. Correlations between TSM Assignments and ROMP for Experiment Two, Both Engagements 17
8. Interrater Correlations of TPR: Experiment Two, Engagement One 20
9. Interrater Correlations of TPR: Experiment Two, Engagement Two 20
10. Correlations between TSM Changes and TPR Assignment: Experiment Two, Both Engagements 21
11. General Frequency Count 22
12. Ranked Category Frequencies Across States 24
13. Overall Frequency Count by Item for both Engagements 25
14. Category Frequencies by State for Engagement One 28
15. Category Frequencies by State for Engagement Two 28
D1. Frequency count by Item and Category: Experiment Two, Engagement One 53
D2. Frequency count by Item and Category: Experiment Two, Engagement Two 55

FIGURES

Figure No.

1. Maneuver Conversion Model 4
2. Tactical State Model 5
3. Example of ACM-PMS Display 7
4. STATE versus ROMP for each subject, Engagement One 18
5. STATE versus ROMP for each subject, Engagement Two 19
B1. Frequency of TSM Assignments: Experiment One, Engagement One . 45
B2. Frequency of TSM Assignments: Experiment One, Engagement Two . 45
B3. Frequency of TPR Assignments: Experiment One, Engagement One . 46
B4. Frequency of TPR Assignments: Experiment One, Engagement Two . 46
B5. Frequency of TSM Assignments: Experiment Two, Engagement One . 47
B6. Frequency of TSM Assignments: Experiment Two, Engagement Two . 47
B7. Frequency of TPR Assignments: Experiment Two, Engagement One . 48
B8. Frequency of TPR Assignments: Experiment Two, Engagement Two . 48

Preface

This research was performed by the Tactical and Training Systems Division of Logicon, Inc. for the Armstrong Laboratory Human Resources Directorate, Aircrew Training Research Division (AL/HRA), Williams AFB, Arizona. The authors wish to thank Dr Fritz Brecke for his assistance in developing the Tactical State Model.

This effort was conducted under Work Unit 1123-35-06, Validation and Refinement of Techniques for Air Combat Performance Assessment, and Contract No. F33615-86-C-0012. Contract monitor is Dr Wayne L. Waag, AL/HRAT.

TACTICAL STATE AND PERFORMANCE ASSESSMENT DURING AIR COMBAT MANEUVERING

SUMMARY

The purpose of this research was to investigate the feasibility of developing a diagnostic measure of tactical decision-making skill in the context of air combat maneuvering (ACM). To do this it was necessary to characterize the tactical position and rate the quality of maneuvering of one of the opponents in a 1 v 1 air combat engagement. The effort consisted of two experiments. Both employed a Tactical State Model (TSM) to characterize position and used a 7-point scale to rate tactical performance. Subjects were experienced F-15 and F-16 instructor pilots. In each experiment, they viewed two prerecorded ACM engagements and used the TSM to assign a tactical position or "state" to the designated aircraft. They also attempted to provide the underlying rationale for selecting the specific states. The subjects then rated the performance of the designated aircraft. Results of Experiment One led to changes in the TSM, the scale used to rate performance, and the methodology used to gather the underlying rationale for tactical state assignment. The analysis of Experiment Two data was also expanded to determine the relationship to an objective measure of positional advantage and to determine the relationship between changes in tactical state and maneuvering performance.

Results indicated that there was significant interrater agreement in assigning a tactical state from the model in both experiments, demonstrating that reliable judgments could be achieved for characterizing tactical position. However, the degree of interrater agreement among subjects using a performance rating scale was substantially less. The additional analysis in Experiment Two showed a positive relationship between state assignments and the objective measure of positional advantage for only one of the engagements. No consistent relationships between changes in tactical state and rated tactical performance were observed for individual subjects.

The goals of this research effort were only partially achieved. Results indicated that a tactical state model can be used reasonably well by experts to characterize position during an air-to-air engagement. However, no consistent relationships were established between these state assignments and an objective measure of positional advantage. Moreover, the results failed to establish sufficient reliability for the tactical performance rating scale developed by this work. And as might be expected, the results further failed to establish a consistent relationship link between changes in position and quality of performance. It is concluded that development of a diagnostic measure of tactical decision making will require a different approach.

INTRODUCTION

This research investigated the feasibility of developing a diagnostic measure of tactical decision-making skill within the context of air combat maneuvering (ACM). The importance of decision making was emphasized in a recent review of the ACM performance measurement literature (Brecke & Miller, 1991) in which the authors concluded that if

> . . . progress is to be made in ACM performance measurement, then the cognitive dimension (which we have dubbed "tactical decision making") has to be brought out of the closet and it has to be acknowledged as a key dimension.

Background

Interest in the measurement of ACM skills was spawned by lost firing opportunities for U.S. warplanes in Southeast Asia during the Vietnam conflict (Oberle, 1983). These missed opportunities were primarily attributed to pilots' lack of adequate weapon-envelope recognition during air combat engagements. In addition, the lack of standardized debriefing for routine ACM training has also been suggested to have played a major role (McGuinness, Forbes, & Rhoads, 1984). These factors were the precursor to efforts by researchers from a variety of scientific disciplines which attempted to develop objective measures of ACM performance using a variety of approaches.

Although the empirical research to date has produced numerous measures of ACM performance, all of these measures have one salient characteristic. For the most part, they have limited diagnostic power to directly support understanding and measurement of the cognitive decision-making aspects of ACM performance (Brecke & Miller, 1991). Among the existing objective measures, the All-Aspect Maneuvering Index (AAMI) (McGuinness, Forbes & Rhoads, 1984) and energy control measures, such as the Specific Excess Energy (P_s), may provide some diagnostic information, especially if they are used in conjunction with one another. However, they are only descriptive of aircraft state and as such yield virtually no information about the cognitive decision-making processes which led to a particular AAMI and P_s value. If training the cognitive aspects of ACM skill is to be supported in an objective and systematic fashion, then highly diagnostic measures of ACM decision-making skills are required.

To be of diagnostic value, measures of ACM decision-making skill must meet certain requirements. First, they must support the identification of errors in decision making. They must also sufficiently characterize the nature of the decision error and must support the generation of feedback which is useful in reducing the probability of a recurrence of that same error. Furthermore, because such measures would be used by instructor and student pilots alike, they should be expressed in language that is "natural" and/or "operational" to both groups of pilots. Finally, measures must be "human factored" or simplified so that human information processing capabilities of the user are not exceeded (Meister, 1971). In other words, the measure must be based on a model of ACM that reduces the apparent infinite variety of ACM decision-making situations to a finite number of situation classes or types which are manageable by operational personnel.

Rationale and Objectives for Present Investigation

Unfortunately, decision making cannot be directly observed but only inferred from its results. Within the context of ACM, "good" decisions would be characterized by achieving or maintaining a "better" tactical position than one's opponent. Therefore, in order to develop a diagnostic measure of tactical decision making, one must first be able to characterize the "present position" of a combatant in tactical terms and second, evaluate the maneuvering that took place between several "present positions" as "good" or "bad." Once a positive link is established between improved or degraded tactical position and "good" or "bad" maneuvering, then it follows that the decision to maneuver resulting in an improved tactical position was correct and one resulting in a degraded tactical position was incorrect.

To achieve the goal of understanding and evaluating decision making in a tactical environment, a number of questions about air-to-air combat engagements must be addressed. First, can such engagements be divided into "segments" based on time or some change in maneuvering or position (e.g. critical decision points)? Second, can a model of ACM be used to characterize the "tactical state" or relative position of an aircraft in these segments? Third, can movement within such a model from one state to another be characterized as "good" or "bad" and can a reliable "performance rating" be assigned to this movement? Fifth, does the assigned performance rating show any relation with movement within the model? In other words, if movement between segments of an engagement appears to be "good" according to the model, is such movement reflected in the performance rating?

If the answers to all of these questions are yes, it follows that tracking states throughout an engagement could identify the point within the model when movement either stops or changes direction. Analysis of this point should identify not only that a change took place, but it should also indicate which measurable parameters changed, causing the change in direction of movement. It could then be inferred that the underlying decision that directed the action to change these parameters was either correct or incorrect for the particular situation. The present experiments were designed to answer these questions and, in doing so, determine whether a diagnostic measure of ACM decision making could be developed using this approach.

ACM Model Selection

In attempting to address these research questions, the first step was to adopt a model of ACM that meets the requirements previously identified. After an exhaustive review of previous measurement approaches and existing tactical state models (Brecke & Miller, 1991), it was decided that the Maneuver Conversion Model (Oberle, 1974) most closely met the identified requirements. This model consists of seven ACM states which are expressed in operational terms (e.g., offensive, neutral, defensive). Figure 1 depicts the original model.

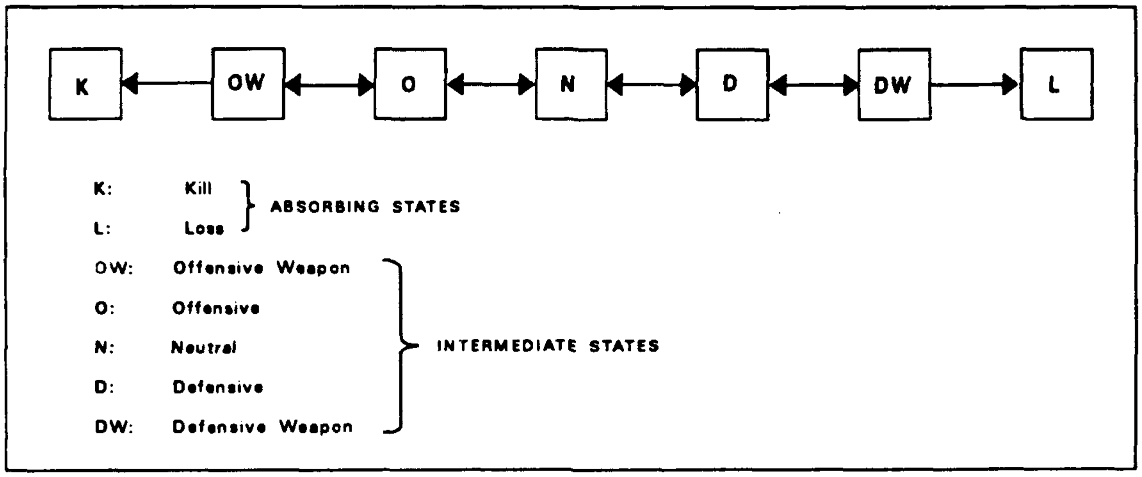
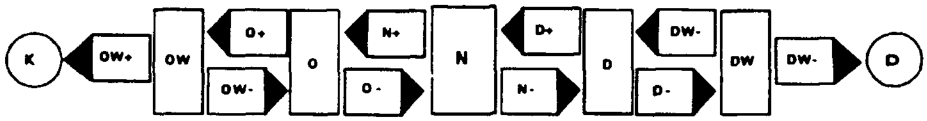


Figure 1. Maneuver Conversion Model

Although it was originally conceived as a conceptual framework for the mathematical treatment of operational testing, the model was also used by the U.S. Navy Fighter Weapons School (FWS) as a method for grading the proficiency and improvement of students as they progressed through the training syllabus (Oberle, 1974).

To test the suitability of using this model to characterize the relative position of combatants in an ACM engagement, the in-house subject matter expert (SME) attempted to apply it by characterizing tactical states in several randomly selected engagements. Unfortunately, these attempts to use the Maneuver Conversion Model proved unsuccessful. Frequently, it was impossible for the SME to unambiguously assign one particular state from the model. For example, one of the two opponents in a 1 v 1 engagement may be in a state which was neither clearly neutral nor clearly offensive. Often a participant may be moving away from a neutral state towards an offensive state, but is not yet clearly offensive. After reviewing numerous engagements, it became clear that the above situation was quite common (i.e., a participant was actually transitioning between two states and was not clearly in either of them).

Based on these initial findings, the original Maneuver Conversion Model (Oberle, 1974) was expanded to include transition states between each of the seven states. This expanded model shall be referred to as the Tactical State Model (TSM) and is shown in Figure 2. Note that two transition states were added between the original states (except for the Kill and Loss states) to permit discrimination between increasing offensiveness and increasing defensiveness. The resulting TSM consisting of 17 states was used in Experiment One.



K: Kill } ABSORBING STATES
L: Loss

OW: Offensive Weapon
O: Offensive
N: Neutral
D: Defensive
DW: Defensive Weapon
} INTERMEDIATE STATES

OW+: Transition State From Intermediate States OW To K
OW -: Transition State From Intermediate States OW To O
O+: Transition State From Intermediate States O To OW
O -: Transition State From Intermediate States O To N
N+: Transition State From Intermediate States N To O
N- : Transition State From Intermediate States N To D
D+: Transition State From Intermediate States D To N
D- : Transition State From Intermediate States D To DW
DW+: Transition State From Intermediate States DW To D
DW-: Transition State From Intermediate States DW To L
} TRANSITIONAL STATES

Figure 2. Tactical State Model

Tactical Performance Rating

The ability to characterize the tactical position of an aircraft engaged in 1 v 1 aerial combat, although essential, is only part of the equation leading to a diagnostic measure of tactical decision making. The second factor to be considered is the "correctness" or "quality" of the decision making that resulted in the tactical state. In order to assess maneuvering performance, a 7-point Tactical Performance Rating (TPR) scale was constructed with possible ratings ranging from +3 to -3. Instructions accompanying the 7-point scale stated that a score of +3 was to be assigned when the pilot made the correct move in the best possible fashion. Likewise, a score of -3 was to be assigned when the pilot made the worst possible move given the current situation. No other anchors were provided.

EXPERIMENT ONE

Objectives

The specific objectives of Experiment One were to: (a) determine whether experienced ACM pilots can reliably use the TSM to characterize the tactical position of an aircraft in an ACM engagement, (b) gather information on the underlying reasons or rationale for assigning a particular tactical state, and (c) verify that experienced pilots can reliably assign a TPR to maneuvering during an ACM engagement.

Method

Subjects

Participants were 4 F-15 and 2 F-16 (N = 6) instructor pilots (IPs) from Luke Air Force Base, AZ. Subjects had a mean of 1680 hours of total flight time, 958.3 hours of F-15/F-16 time, and 420.83 hours of instructor pilot time.

Equipment

The Simulator for Air-To-Air Combat (SAAC) was used to "fly" the two engagements which were recorded for use in Experiment One. The SAAC, which is located at Luke Air Force Base is used for training air-to-air combat techniques. The simulator has two visual domes which can be configured with either F-15 or F-16 cockpits and provides the pilot with an "out-the-window" field of view of 296 degrees in the horizontal dimension by 150 degrees in the vertical dimension.

The Air Combat Maneuvering Performance Measurement System (ACM-PMS) was used to present the recorded engagements to subjects for TSM and TPR assignments. The ACM-PMS is a state-of-the-art research tool which can provide a graphic reconstruction of an ACM engagement. The system consists of a Gould Model 32/67 mainframe computer and a Silicon Graphics Iris Model 2400 Turbo Graphics Display. The high resolution monitor is partitioned into three display windows. The top one third of the display is divided vertically into two equal size windows and the bottom two thirds make up the single third window. Each window is capable of providing several different displays. For this experiment, the work station was configured with out-of-cockpit view from #1 in the upper left, out-of-cockpit view from #2 in the upper right and the 3-D display in the large center window. The out-of-cockpit views provide some heads-up display (HUD)-like information as well as the pilot's view of the engagement. The 3-D display provides a three-dimensional view of the engagement.

User interface with the ACM-PMS is accomplished through a touch-screen to access various options and a mouse is used to manipulate the 3-D view. Both the viewing angle (azimuth and elevation) and the viewing perspective (distance from center) of the 3-D view can be changed. Figure 3 shows an example of the ACM-PMS display used in the present effort.

Materials

A graphic depiction of the TSM was provided for reference during assignment of tactical states. Subjects were also provided a copy of the 7-point TPR scale, as previously described, for use in rating tactical performance. A questionnaire was also developed (See Appendix A) which asked questions about each pilot's previous flight experience as well as questions about various aspects of the experiment. A standard cassette tape recorder was provided to record subject comments during the experiment.

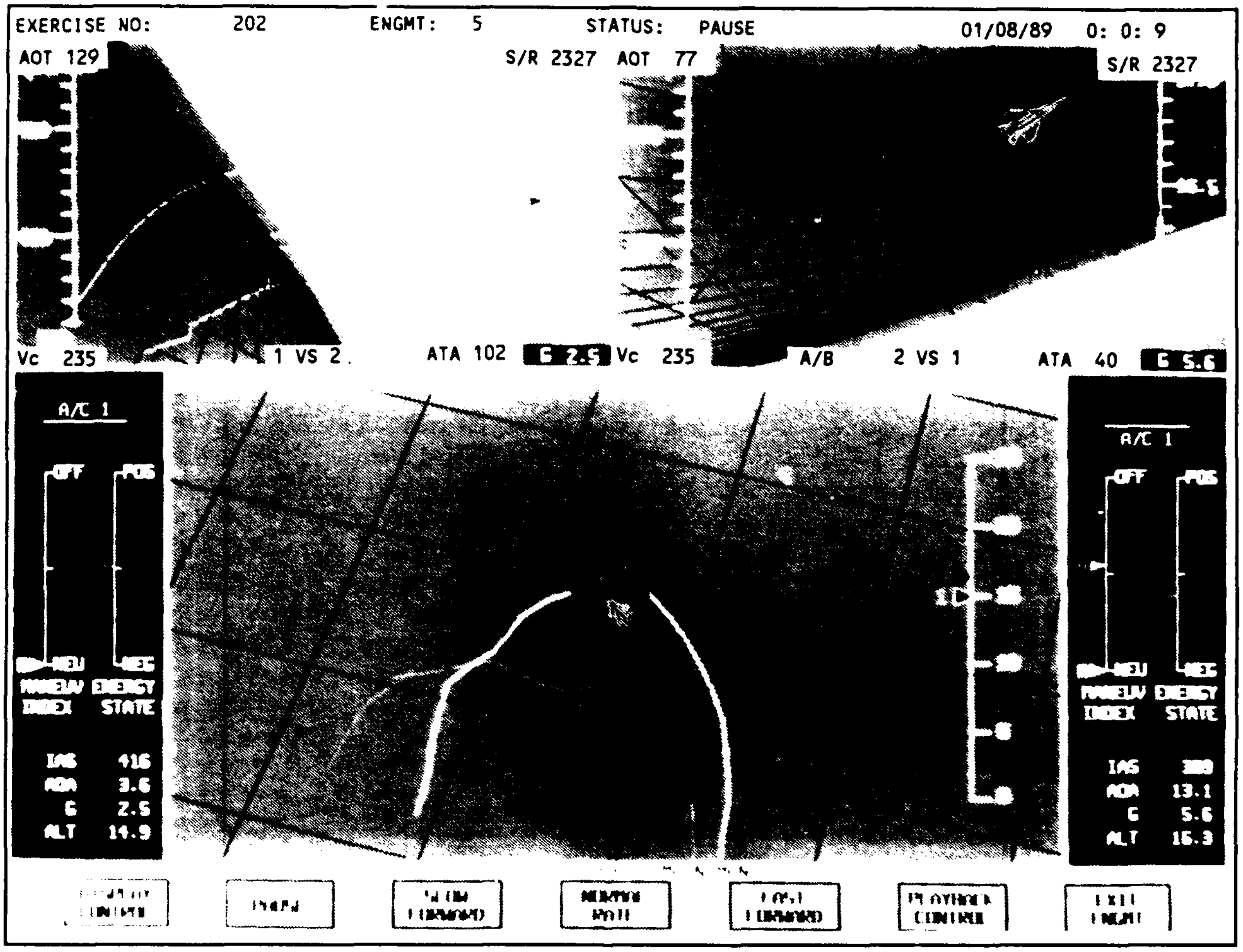


Figure 3. Example of ACM-PMS Display

Procedures

The two engagements used in this experiment were selected from a store of 1 v 1 engagements previously recorded on the ACM-PMS. They were selected with the criteria that each was at least 2 minutes in duration and ended with an ordnance kill. Both engagements started from a neutral position with the pilot's objective being to kill the opponent and survive the fight.

The subjects were told that the purpose of the project was to assist researchers in gaining some insight into tactical decision making. To accomplish this, they would be required to do two things. First, at every 5-second point throughout each of the two engagements, they would have to characterize the tactical state of one of the aircraft using the TSM and verbally provide the accompanying rationale for this characterization. Second, they would make another pass through the engagement to rate the tactical performance of that same aircraft using the 7-point TPR scale provided.

The experiment was conducted in the SAAC debrief room. The subject was seated at the ACM-PMS work station with a research staff member seated next to him to operate the ACM-PMS. The subject began the experiment by listening to taped instructions describing the experimental tasks. Subjects were then given an explanation of the ACM-PMS display to familiarize them with the aircraft information available on the display to assist them in performing the task. They were also instructed on use of the mouse, which enabled them to manipulate the 3-D view of the engagement. Finally, subjects were allowed to practice each task to ensure they thoroughly understood the instructions. When the researcher was satisfied that the subject understood the task, the experiment began.

Engagements were arranged so that each could be presented sequentially from beginning to end. Once the experiment began, the experimenter proceeded through the engagement, stopping every 5 seconds. At each five-second point, the subject spoke into the tape recorder denoting the time, the TSM choice and the rationale for the assignment. When assigning tactical state, if the state assigned to successive 5-second points were not adjacent states on the model, the subject performed a "zoom." This was accomplished by resetting the engagement to the previous 5-second point and having the subject assign states at 1 second intervals between the nonadjacent 5-second points. This procedure was done to determine whether the TSM met the Markovian stochastic assumption of smooth transitions from one state to another.

Due to the extensive time that was necessary to collect all of the data, the experimental tasks were split up into a 2-day testing period. On the first day, subjects went through Engagement One providing both state and TPR assignments. On the second day of testing, subjects performed the same task on Engagement Two and completed the questionnaire. Engagement One lasted 2 minutes 5 seconds while Engagement Two lasted 2 minutes 25 seconds, thereby providing 25 and 29 data collection points for the two engagements, respectively.

Results

Initial perusal of the data revealed no apparent differences between F-15 and F-16 pilots. Consequently, all data were combined and considered as a whole set. Frequencies of each state assignment over time for both engagements are presented in Appendix B. In order to assess the degree of interrater agreement, a number of statistical analyses were performed. First, Pearson product-moment correlations among the raters were computed. In order to estimate the overall reliability of the ratings, intraclass correlations were also computed (Guilford, 1954). Such an analysis produced two reliability estimates. The first, $\bar{r}_{11}$, provides an estimate of the mean reliability of a single rater. The second, r_{kk}, estimates the reliability of all raters as a group. It can be shown that r_{kk} is equivalent to Cronbach's coefficient alpha which is generally used to provide an overall estimate of the internal consistency of a set of measures.

For state assignments using the TSM, the results showed a fairly high level of agreement among subjects (see Figures B1 and B2 for the frequency of state assignments over time). Such agreement is further evidenced by significant correlations among raters as shown in Tables 1 and 2. In fact, all correlations were found to be statistically significant. The mean reliabilities for a single rater were found to be .59 and .74 for Engagements One and Two, respectively. When considered as a group, the reliabilities increased to .92 and .96 respectively.

Table 1. Interrater Correlations of TSM Assignments: Experiment One, Engagement One.

	Subject ID Number				
	2	3	4	5	6
1	.61**	.58**	.69**	.59**	.46*
2		.73**	.80**	.66**	.55**
3			.84**	.77**	.58**
4	$\bar{r}_{11}$ = .59			.88**	.69**
5	r_{kk} = .92				.58**

*P<.05
**P<.01

Table 2. Interrater Correlations of TSM Assignments: Experiment One, Engagement Two.

	Subject ID Number				
	2	3	4	5	6
1	.85	.85	.79	.75	.71
2		.87	.83	.82	.81
3			.89	.82	.80
4	$\bar{r}_{11}$ = .74			.75	.76
5	r_{kk} = .96				.67

p<.01 For all correlations

Overall, there appeared to be a lot more variability for TPR assignments as evidenced in Figures B3 and B4. Tables 3 and 4 present the interrater correlations. As shown, only 8 of the 15 correlations for Engagement One were statistically significant while only 4 of the correlations were significant for Engagement Two. This general lack of agreement is further confirmed by the mean reliabilities for a single rater of .34 and .22 for the two engagements. Even when considered as a group, the reliabilities were found to be only .80 and .70, respectively.

Table 3 Interrater Correlations of TPR Assignments: Experiment One, Engagement One.

	Subject ID Number				
	2	3	4	5	6
1	.35	.71**	.78**	.58**	.19
2		.47*	.22	.49*	.36
3			.60**	.52**	.21
4	$\bar{r}_{11}$ = .34			.31	.09
5	r_{kk} = .80				.52**

*p<.05
**p<.01

Table 4. Interrater Correlations of TPR Assignments: Experiment One, Engagement Two.

	Subject ID Number				
	2	3	4	5	6
1	.24	.47*	.21	.64**	-.04
2		.28	.37*	.18	.35
3			.32	.58**	.13
4	$\bar{r}_{11}$ = .22			.23	-.03
5	r_{kk} = .70				.15

*p<.05
**p<.01

Questionnaire data provided additional insight into the utility and validity of the TSM. All of the subjects indicated that the TSM was useful to them, and that they were able to use it with little acclimation time. Subjects also indicated that the TSM was able to characterize the tactical state of an aircraft most of the time. There were several exceptions, which are described in detail in the discussion.

Data gathered from recorded comments regarding the rationale for assigning tactical states proved to be of limited use. Originally, these comments were intended to provide the foundation for developing objective, parameter-based definitions of each state in the TSM. However, analysis of the rationales provided by the subjects presented several problems which are also considered in detail in the Discussion section.

Discussion

The obtained levels of interrater agreement suggested that the TSM was potentially a reliable tool for characterizing the tactical state of a participant in an ACM engagement. In addition all subjects stated that the model made good intuitive sense to them and that the framework appeared sound. Several subjects compared it to the Basic Fighter Maneuvers (BFM) or "BFM Ladder" that they had used before in Fighter Lead-In BFM training.

Subjects also agreed that the TSM was comprehensive enough to describe nearly all of the tactical states they had viewed. Two notable exceptions, however, were discovered. The first occurred in head-on situations, where some confusion developed as to whether an aircraft was Offensive Weapons Plus (OW+) or Defensive Weapons Minus (DW-). This confusion, due to the availability of reliable all-aspect short range missiles, resulted in lower levels of agreement among subjects in assigning states to the head-on situation. The second gray area, pointed out by the subjects and supported by the data, occurred in a high-speed overshoot situation. In such instances, the rated aircraft went almost instantaneously from OW+ to DW-. Agreement level again was lower because subjects differed in acknowledging when the overshoot actually occurred.

Use of the "zoom" technique to analyze segments which appeared to skip over cells in the model proved quite successful. In every instance where a "zoom" was required, movement was found to make smooth transitions from one state to another therefore upholding the Markovian assumption (Dynkin, 1982).

The observed reliabilities for TPRs were substantially lower than those for state assignments. In fact, the mean reliabilities for a single rater, which is the most important of the two indices computed, were extremely low indicating very little agreement. One difficulty with the initial TPR scale may have been that it contained no intrinsic frame of reference for subjects to use in assigning ratings. While subjects indicated on the questionnaire that the TSM made good intuitive sense to them in a general way, and that it reminded them of other metrics that they had used before, they reported that the scale used to rate tactical performance did not possess any familiar framework. This lack of benchmarking may have contributed to reduced agreement among subjects. It also appeared that subjects occasionally had difficulty in using the scale and were reluctant to give ratings at various times in the engagement. For example, they might repeatedly ask for a portion of the engagement to be replayed, or verbally indicate that the scale just did not seem to "fit." In such cases, the accuracy of the resulting rating would appear to be highly questionable.

Analysis of the transcripted rationale provided by the subjects for assigning tactical state revealed several problems which prevented the development of objective definitions of the tactical states. The first and foremost problem identified was that the rationale voiced by the subjects for assigning a state was primarily prescriptive rather than descriptive in nature. Subjects tended to talk about what the aircraft was or was not doing rather than describing where the aircraft was in relation to the opponent. When actual physical parameters such as Gs, angles and positions were mentioned, they were often not discussed by all or even a majority of subjects in conjunction with the same 5 second point. In addition, when there was no consensus for the state assigned to a specific 5 second point, any specific parameters mentioned could not be associated with a single state. Even when all or most subjects agreed on the state of a 5 second point, the comments on parameters were so varied as to provide no commonality. For example, in one instance, a subject said that a "hard" turn contributed to the state assignment. Another subject mentioned "defensive" turn associated with that point and yet another said it was a "good" turn. None of these descriptions was objective enough to be used as part of a definition for the given state. Although central themes did run through many of the transcripts, significant interpretation of the data would have been required to achieve quantitative assessment of these rationales.

Implications for Experiment Two

The results of Experiment One pointed out a number of areas in need of improvement if the goals of this research were to be realized. The first of these was an expansion of the neutral state of the TSM. During Experiment One, it was observed that many subjects were having some difficulty assigning a tactical state when the combatants were head-on to each other. In this situation with all-aspect weapons, each participant could bring weapons to bear on the other, which created confusion in assigning the offensive weapons or defensive weapons state. Although the addition of the transition states made a significant improvement in the model, there still appeared to be a problem characterizing the head-on situation.

This ambiguity in the model was eliminated by treating neutral as a complex state. The "new" neutral state would be made up of two parts, "Neutral/Neither at Risk" and "Neutral/Both at Risk". With the expansion of the neutral classification, it was felt that subjects should be able to indicate the exact nature of a neutral position. These new classifications were expected to help eliminate the low agreement level encountered in the head-on situation and give subjects more descriptive terminology to work with when describing the neutral states. Cursory attempts by the SME to use the model with the divided neutral state were encouraging.

Another area that required improvement was the 7-point scale used to rate tactical performance. While the initial attempt to rate tactical performance proved to be less than satisfactory, there was an overall trend for subjects to classify engagements in the same direction. In other words, most subjects would consistently rate a move as either on the positive or negative side of the scale. However, there appeared to be little agreement in terms of the magnitude of the rating. Although the complete solution to this problem was not immediately clear, it seemed reasonable that the scale could be improved by providing a better "frame of reference" in terms of explicit anchor points. In Experiment One subjects were only instructed that a rating of +3 indicated that the pilot had performed the proper maneuver in the best possible way, with all other ratings falling somewhere under that. To improve the scale, explicit word descriptions were developed for each increment.

An additional area concerned the actual engagements used for data collection during Experiment One. These two engagements could be characterized as relatively "straightforward" in that the rated combatant obtained the advantage early in the engagement and maintained that advantage until a kill was achieved. Consequently, only parts of the TSM were actually being used. For this reason, it was decided that two new engagements that would more fully exercise the model would be used in the follow-on experiment. The first engagement was specifically flown for use in the experiment and was actually "choreographed" by the in-house SME. The objective of this engagement was to use the entire TSM. The rated aircraft started neutral, gained the advantage, lost the advantage, and was ultimately killed with ordnance. The second engagement was selected from the existing data base of engagements. It was selected because the two aircraft exchanged the advantage several times. This repeated crossing through the neutral state was expected to challenge the raters and test the ability of the model to characterize this more ambiguous situation.

The last implication from the first experiment concerned the rationale or reasons used to assign a particular tactical state. After transcribing all the tapes, it was apparent that although pilots explain things in similar terms, they do not use exactly the same words. As a result it was impossible to develop any kind of empirically based consensus as to the reasoning behind assigning a particular state. However, after qualitative analysis of the transcripts, certain themes and variables were identified. These results led to a comprehensive listing of the reasons given to assign states. This list was then made into a checklist which subjects used in the follow-on experiment. It was reasoned that by providing all subjects the same list of underlying "reasons" to choose from, the resulting data would be more standardized and certainly more amenable to quantitative analysis.

EXPERIMENT TWO

Objectives

The objectives of Experiment Two were essentially the same as those in Experiment One, however, two additional objectives were added. Because the first experiment resulted in high levels of agreement among pilots in making TSM assignments, it seemed reasonable to determine the relationship between such state ratings and a computationally-derived objective measure of positional advantage. The score chosen for this analysis was the Relative Offensive Maneuvering Potential (ROMP) which is derived from the combination of AAMI scores of the two opponents (McGuiness, Forbes, & Rhoads, 1984). The ROMP is essentially the difference of the AAMI scores of the two opponents. This index is scaled from -100 to 100 and is an indicator of the relative positional advantage at any time during the engagement. Additionally, an attempt was made to determine if a relationship existed between changes in state position and the quality of maneuvering performance.

The specific objectives of Experiment Two were to: (a) validate the use of the expanded TSM in characterizing the tactical position of an aircraft in an ACM engagement, (b) determine the relationship between the state ratings using the TSM and the ROMP score, (c) verify that the quality of movement within the model can be reliably assessed using the revised TPR scale, (d) gather information on the rationale for assigning a tactical state, and (e) determine the relationship between the changes in TSM assignments and TPRs.

Method

Subjects

Participants were 4 F-15 and 4 F-16 Replacement Training Unit (RTU) IPs stationed at Luke Air Force Base. Subjects had a mean of 2056 hours of total flight time, 1471 hours in fighter aircraft, and 1085 hours in F-15/F-16 aircraft. Subjects had a mean age of 34.

Equipment

The same equipment used in Experiment One was used in Experiment Two.

Materials

Two data collection forms were developed for this experiment. The first, used for tactical state assignment, had a graphic depiction of the TSM printed across the top. The second had the revised TPR scale across the top. Both forms included the same checklist of items, although the checklist portion was not used in the second part of the experiment when tactical performance was rated. Examples of these data collection forms can be found in Appendix B. The same questionnaire used in Experiment One was also administered.

Procedures

The procedures used in Experiment Two were the same as those used in Experiment One with the following exceptions. Subjects' comments were not tape recorded. Rather, the checklist, as described above, was used at each 5 point. Subjects placed an "X" in the appropriate cell of the TSM and then placed a check mark on those items that had influenced their selection. The TPR task was performed in the same manner except that the checklist with the rating scale printed at the top was used. Subjects did not mark influencing items when rating performance. Although the experiment proceeded at a faster pace using the checklist, two days were still required with the subject completing one of the two engagements on each day. Durations for Engagements One and Two were 1 minute 45 seconds and 2 minutes 20 seconds, respectively, thus providing 21 and 28 data collection points for the two engagements.

Results

TSM Interrater Agreement

The analysis procedures for determining the level of interrater agreement used in Experiment One were also employed in Experiment Two. Again, the frequencies of state assignments over time were computed and are presented in Appendix B. Tables 5 and 6 present the correlation matrices of the TSM assignments for Engagements One and Two, respectively. For Engagement One, all correlations were quite high and statistically significant. In fact, the mean reliability coefficient for a single rater was found to be .95. For all raters, the reliability coefficient was .99. For Engagement Two, all intercorrelations were statistically significant, but the magnitudes were substantially lower. In fact, r_{11} was found to be only .68. Overall, these results indicate substantial agreement in the assignment of state ratings with the highest levels of agreement for Engagement One.

Table 5. Interrater Correlations of TSM Assignments: Experiment Two, Engagement One.

	Subject ID Number						
	2	3	4	5	6	7	8
1	.96	.96	.97	.94	.96	.96	.92
2		.95	.96	.97	.96	.95	.92
3			.96	.95	.97	.97	.96
4				.96	.96	.97	.93
5					.96	.97	.95
6						.98	.95
7							.96

$\bar{r}_{11} = .95$

$r_{kk} = .99$

p<.01 For all correlations

Table 6. Interrater Correlations of TSM Assignments: Experiment Two, Engagement Two.

	Subject ID Number						
	2	3	4	5	6	7	8
1	.86**	.82**	.46*	.71**	.86**	.62**	.79**
2		.78**	.44*	.54**	.70**	.57**	.67**
3			.54**	.66**	.78**	.71**	.81**
4				.39*	.59**	.64**	.50**
5					.79**	.64**	.80**
6						.71**	.79**
7							.81**

$\bar{r}_{11} = .68$

$r_{kk} = .94$

* p<.05
** p<.01

Relationship of TSM Assignments with ROMP

The ROMP score was used to determine the relationship between the tactical states assigned by subjects and an objective measure of positional advantage. ROMP scores were calculated for each 5-second point and were then correlated with each subject's tactical state rating. Graphs of each subject's TSM assignment and the ROMP are presented in Figures 4 and 5 for Engagements One and Two, respectively. The resulting correlations for each subject as well as the average TSM assignments across all subjects are presented to the right in Table 7. The results of this analysis indicated that there was a strong relationship between these scores only for Engagement One. In fact, there was almost perfect agreement as the correlations for all subjects as well as their average approached one. However, no significant correlations were obtained for Engagement Two, although the trends were in the expected direction and a number did approach statistical significance.

Table 7. Correlations between TSM Assignments and ROMP for Experiment Two, Both Engagements.

Subject ID Number	Engagement Number 1	2
1	.95**	.34
2	.95**	.28
3	.96**	.23
4	.96**	.35
5	.96**	.21
6	.96**	.33
7	.96**	.33
8	.91**	.30
Avg.	.97**	.35

** $p<.01$

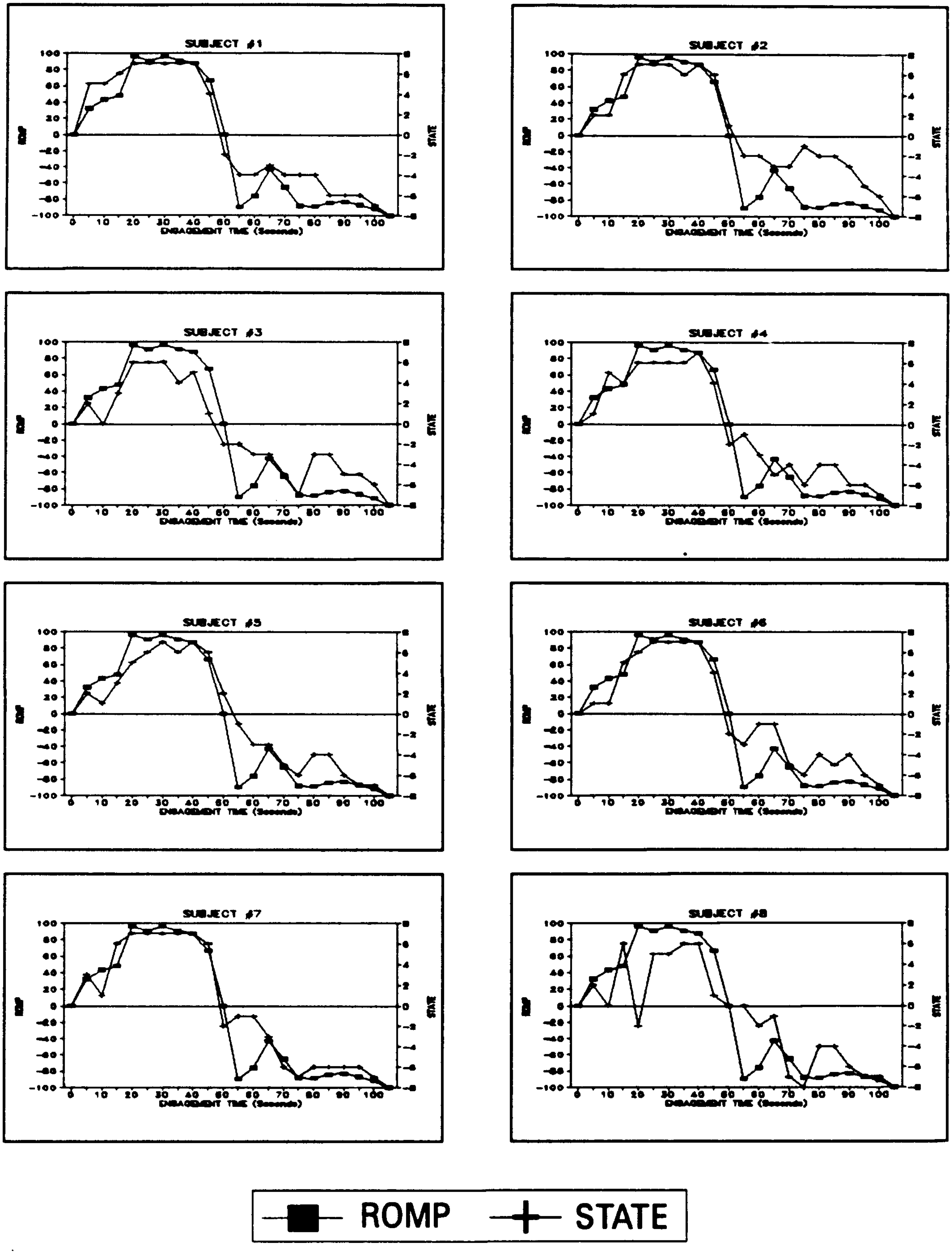


Figure 4. STATE versus ROMP for each subject, Engagement One.

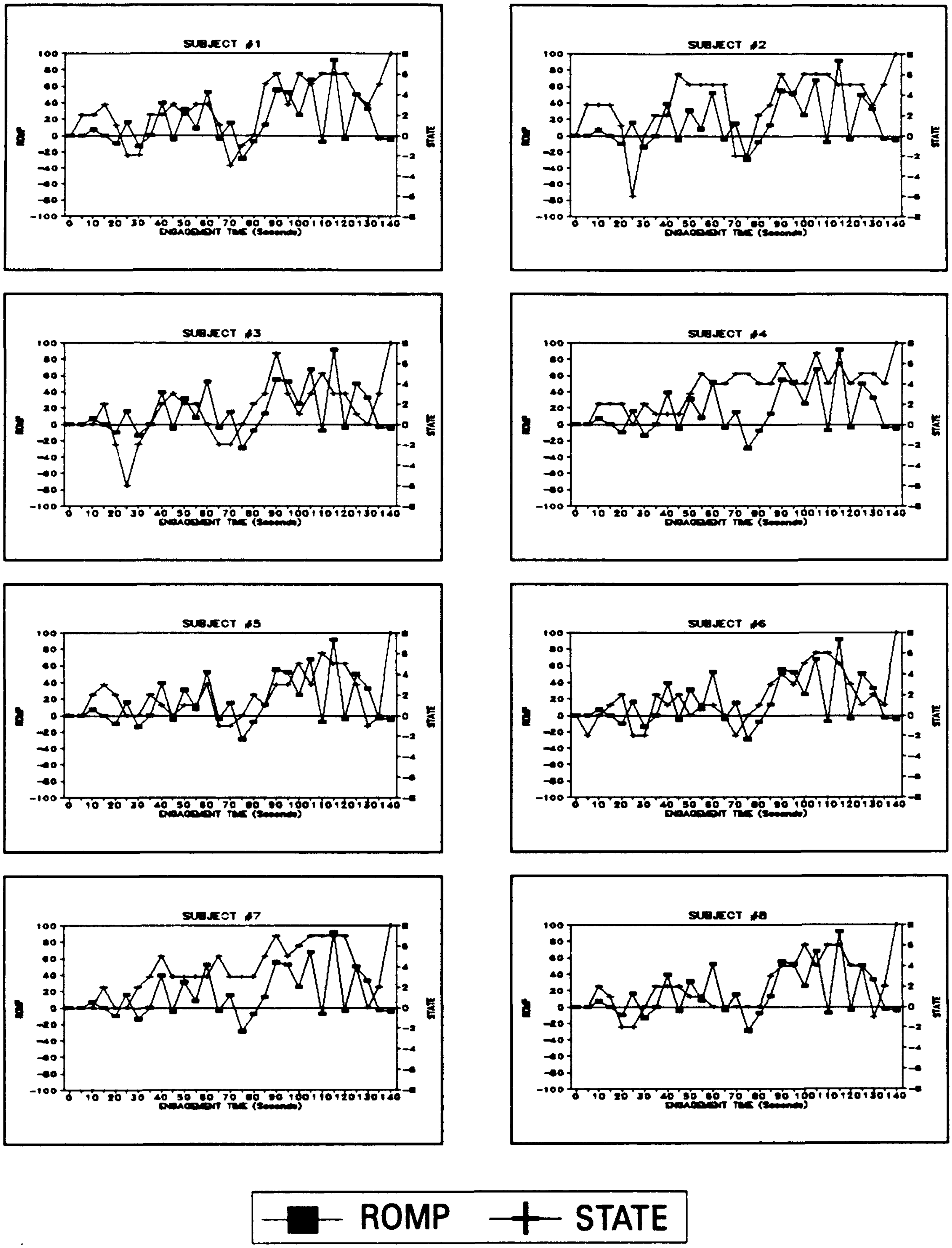


Figure 5. STATE versus ROMP for each subject, Engagement Two.

TPR Interrater Agreement

Similar to Experiment One, the overall interrater agreement for TPRs was substantially less than the observed agreement for tactical state assignment as shown in Figures B7 and B8. Tables 8 and 9 present the interrater correlations for both engagements. For Engagement One, most of the interrater correlations were statistically significant with a mean reliability for a single rater of .69. For the entire group, the reliability coefficient was .95. However, for Engagement Two, the degree of agreement was substantially reduced. As shown in Table 9, less than half of the intercorrelations achieved statistical significance. Moreover, the mean reliability for a single rater dropped to .31, a level similar to that found in Experiment One.

Table 8. Interrater Correlations of TPR: Experiment Two, Engagement One.

	Subject ID Number						
	2	3	4	5	6	7	8
1	.79**	.70**	.66**	.63**	.93**	.77**	.56**
2		.71**	.78**	.81**	.82**	.85**	.45*
3			.73**	.83**	.74**	.85**	.41
4				.71**	.74**	.70**	.42
5					.60**	.89**	.51*
6	$\bar{r}_{11}$ = .69					.78**	.45*
7	r_{kk} = .95						.64**

* p<.05
** p<.01

Table 9. Interrater Correlations of TPR: Experiment Two, Engagement Two.

	Subject ID Number						
	2	3	4	5	6	7	8
1	-.07	.21	.47**	.31	.25	.41*	.48*
2		.01	.09	.14	.41*	.15	.35
3			.04	.27	.22	.42*	.27
4				.28	.40*	.20	.21
5					.63**	.42*	.47*
6	$\bar{r}_{11}$ = .31					.57**	.51**
7	r_{kk} = .79						.32

* p<.05
** p<.01

Relationship between State Change and Rated Performance

Analyses were performed to determine if <u>changes</u> in tactical state were related to TPR ratings. It was initially hypothesized that improvements in tactical position would be positively related to ratings of maneuvering performance and conversely that a deterioration of tactical position would be negatively related to ratings of maneuvering performance. For each interval, a change score was computed by taking the difference of the current state and the previous state. These change scores were then correlated with TPR assignments. The results of the analysis for both engagements are presented in Table 10. As shown, no consistent relationships were found. For Engagement One, three of the eight correlations were significant, while five were significant for Engagement Two. However, only one subject produced significant correlations for both engagements. Despite these negative findings for individual subjects, the averages across subjects produced significant correlations for both engagements.

Table 10. Correlations between TSM Changes and TPR Assignment: Experiment Two, Both Engagements.

Subject ID Number	Engagement Number 1	Engagement Number 2
1	-.31	-.52**
2	-.38	.08
3	-.40	-.50*
4	-.48*	-.02
5	-.35	-.43*
6	-.48*	-.41*
7	-.51*	-.32
8	-.01	-.64**
Avg.	-.56**	-.64**

* p<.05
** p<.01

Rationale for State Assignment

Data indicating the pilots' reason or rationale for assigning a particular tactical state were obtained from the checklist portion of the TSM data collection form (see Appendix C). The first step in the analysis was to code for each subject the state assignment and items selected for each time interval within each engagement. These data were then summarized across subjects and time intervals for each engagement. The resulting frequency counts for each item across the 18 tactical states are presented in Appendix D. From this basic data set, a number of analyses were performed.

The first analysis determined the overall frequency for selection of each category and item summed across all states. These results are presented in Table 11. As shown, the table provides frequency counts for each category and item within each category for both engagements. From this table, it is clear that wide variation occurred in terms of the frequencies for both items and categories. For example, the category "Range" received the most counts within each engagement and other categories, such as "Circle," were used very infrequently. Likewise, certain items such as "3-9 position advantage" were selected quite often while others such as "Outside parameters" and "In-range Heat" were chosen very infrequently. From these data, it is clear that differences occurred not only in the frequencies for categories and items, but also differences across engagements. For example, the "Airspeed" category was given as a reason for assigning a particular state almost twice as often for Engagement Two. Even more striking are differences in the category "G" which was given as a reason for state assignment 45 times for Engagement One but only 7 times for Engagement Two.

TABLE 11. General Frequency Count

ENGAGEMENT ONE				ENGAGEMENT TWO			
CATEGORY	ITEM NAME	TOTALS ITEM	TOTALS CAT.	CATEGORY	ITEM NAME	TOTALS ITEM	TOTALS CAT.
AIRSPEED				AIRSPEED			
1	AIRSPEED < 200	4		1	AIRSPEED < 200	3	
2	AIRSPEED > 450	0		2	AIRSPEED > 450	0	
3	A/S EXCESSIVE	1		3	A/S EXCESSIVE	0	
4	AIRSPEED ADV	21		4	AIRSPEED ADV	20	
5	AIRSPEED DISADV	6		5	AIRSPEED DISADV	30	
6	AIRSPEED SAME	3		6	AIRSPEED SAME	15	
7	ACCELERATING	1	36	7	ACCELERATING	3	71
RANGE				RANGE			
8	RANGE < 1000 Ft	9		8	RANGE < 1000 Ft	8	
9	RANGE 1000 - 3500 Ft	18		9	RANGE 1000 - 3500 Ft	20	
10	RANGE > 3500Ft	9		10	RANGE > 3500Ft	0	
11	IN PARAMETERS	11		11	IN PARAMETERS	13	
12	OUTSIDE PARAMETERS	1		12	OUTSIDE PARAMETERS	4	
13	NEAR PARAMETERS	31		13	NEAR PARAMETERS	22	
14	IN-RANGE GUN	6		14	IN-RANGE GUN	14	
15	IN-RANGE HEAT	3	88	15	IN-RANGE HEAT	1	82
BOGIE POSITION				BOGIE POSITION			
16	BOGIE IN WINDSCREEN	7		16	BOGIE IN WINDSCREEN	16	
17	BOGIE IN HUD FOV	12		17	BOGIE IN HUD FOV	9	
18	BOGIE NEAR CANOPY BOW	4		18	BOGIE NEAR CANOPY BOW	14	
19	BOGIE LEFT	0		19	BOGIE LEFT	2	
20	BOGIE RIGHT	1		20	BOGIE RIGHT	1	
21	LOST SIGHT OF BOGIE	2	26	21	LOST SIGHT OF BOGIE	0	42
CIRCLE				CIRCLE			
22	ONE CIRCLE FIGHT	1		22	ONE CIRCLE FIGHT	1	
23	TWO CIRCLE FIGHT	0	1	23	TWO CIRCLE FIGHT	1	2

Table 11. Continued

ENGAGEMENT ONE				ENGAGEMENT TWO			
CATEGORY	ITEM NAME	TOTALS ITEM	TOTALS CAT.	CATEGORY	ITEM NAME	TOTALS ITEM	TOTALS CAT.
ALTITUDE				ALTITUDE			
24	ALTITUDE ADVANTAGE	2		24	ALTITUDE ADVANTAGE	17	
25	ALTITUDE DISADVANTAGE	3		25	ALTITUDE DISADVANTAGE	5	
26	ALTITUDE DIFFERENT	0		26	ALTITUDE DIFFERENT	1	
27	ALTITUDE SAME	1	6	27	ALTITUDE SAME	5	28
SWITCHES				SWITCHES			
28	SWITCHES IN GUNS	10		28	SWITCHES IN GUNS	15	
29	SWITCHES IN HEAT	12		29	SWITCHES IN HEAT	3	
30	SWITCHES WRONG	4		30	SWITCHES WRONG	5	
31	SWITCHES LATE	1	27	31	SWITCHES LATE	0	23
RADAR				RADAR			
32	RADAR LOCK	8		32	RADAR LOCK	1	
33	RADAR NO LOCK	0		33	RADAR NO LOCK	0	
34	RADAR CLOSE TO LOCK	0	8	34	RADAR CLOSE TO LOCK	0	1
NOSE POSITION				NOSE POSITION			
35	NOSE ON	25		35	NOSE ON	37	
36	NOSE OFF	2		36	NOSE OFF	6	
37	NOSE IN LEAD	2		37	NOSE IN LEAD	6	
38	NOSE IN LAG	5		38	NOSE IN LAG	7	
39	NOSE WITHIN 60°	2		39	NOSE WITHIN 60°	7	
40	HEAD-ON	1		40	HEAD-ON	0	
41	NOSE LEVEL	1	38	41	NOSE LEVEL	0	63
ENERGY				ENERGY			
42	GAINING ENERGY	8		42	GAINING ENERGY	5	
43	MAINTAINING ENERGY	24		43	MAINTAINING ENERGY	14	
44	LOSING ENERGY	6	38	44	LOSING ENERGY	19	38
Gs				Gs			
45	HIGH G's	8		45	HIGH G's	3	
46	LOW G's	11		46	LOW G's	0	
47	SAME G's	3		47	SAME G's	2	
48	NOT USING G's	21		48	NOT USING G's	0	
49	TOO SOON G's	0		49	TOO SOON G's	1	
50	GOD'S G's	2	45	50	GOD'S G's	1	7
SHOTS				SHOTS			
51	GUN SHOTS	8		51	GUN SHOTS	5	
52	HEAT SHOTS	9		52	HEAT SHOTS	0	
53	HI ANGLE GUN SHOTS	1		53	HI ANGLE GUN SHOTS	2	
54	MISSED SHOTS	1		54	MISSED SHOTS	0	
55	LATE SHOTS	3	21	55	LATE SHOTS	2	9
3-9 POSITION				3-9 POSITION			
56	POSITION ADVANTAGE	34		56	POSITION ADVANTAGE	56	
57	LOST ADVANTAGE	18	52	57	LOST ADVANTAGE	10	66
LIFT VECTOR				LIFT VECTOR			
58	CONTROL-GOOD	19		58	CONTROL-GOOD	48	
59	CONTROL-POOR	29		59	CONTROL-POOR	8	
60	ABOVE HORIZON	4		60	ABOVE HORIZON	4	
61	BELOW HORIZON	1		61	BELOW HORIZON	8	
62	ON OPPONENT	1		62	ON OPPONENT	8	
63	CLOSE	1	55	63	CLOSE	0	76

Table 11. Concluded

ENGAGEMENT ONE				ENGAGEMENT TWO			
CATEGORY	ITEM NAME	TOTALS ITEM	CAT.	CATEGORY	ITEM NAME	TOTALS ITEM	CAT.
AOA				AOA			
64	HIGH AOA	3		64	HIGH AOA	2	
65	LOW AOA	4		65	LOW AOA	0	
66	GOOD AOA	0		66	GOOD AOA	0	
67	SAME AOA	0	7	67	SAME AOA	0	2
CLOSING VELOCITY				CLOSING VELOCITY			
68	HIGH	5		68	HIGH	5	
69	LOW	2	7	69	LOW	0	5
ASPECT				ASPECT			
70	LESS THAN 60°	27		70	LESS THAN 60°	7	
71	60° to 90°	5		71	60° to 90°	8	
72	GREATER THAN 90°	2	34	72	GREATER THAN 90°	7	22
OTHER				OTHER			
73	OTHER	59	59	73	OTHER	31	31

To determine the relative importance of the categories and items, these data were ranked by frequency of selection for each engagement. Table 12 presents data ranked by category and Table 13 presents the data ranked by specific items. For both engagements, the categories of "Range" with reference primarily to weapons parameters and "Lift Vector Control" were cited most often as indicative of state position. "3-9 Position" was ranked third for Engagement One and fourth for Engagement Two. In general, there appeared to be fair agreement with certain exceptions such as "Airspeed" and "G" as mentioned earlier. There were 72 unique items available on the checklist. For Engagement One, 62 items were actually selected with 49 items selected two or more times. For Engagement Two, 54 items were selected with 46 selected two or more times. It was also found that a group of 12 items accounted for 50% of the selections made in Engagement One and 52% of the selections made in Engagement Two. Again, these items come primarily from the top rated categories of "Range," "Lift Vector Control," and "3-9 Position."

TABLE 12. Ranked Category Frequencies Across States

ENGAGEMENT ONE				ENGAGEMENT TWO			
CATEGORY	FREQ	%	CUM %	CATEGORY	FREQ	%	CUM %
RANGE	88	18%	18%	RANGE	82	15%	15%
LIFT VECTOR	55	11%	29%	LIFT VECTOR	76	14%	29%
3-9 POSITION	52	11%	40%	AIRSPEED	71	13%	43%
Gs	45	9%	49%	3-9 POSITION	66	12%	55%
ENERGY	38	8%	57%	NOSE POSITION	63	12%	67%

Table 12. Concluded

ENGAGEMENT ONE				ENGAGEMENT TWO			
CATEGORY	FREQ	%	CUM %	CATEGORY	FREQ	%	CUM %
NOSE POSITION	38	8%	65%	BOGIE POSITION	42	8%	74%
AIRSPEED	36	7%	72%	ENERGY	38	7%	82%
ASPECT	34	7%	79%	ALTITUDE	28	5%	87%
SWITCHES	27	6%	84%	SWITCHES	23	4%	91%
BOGIE POSITION	26	5%	90%	ASPECT	22	4%	95%
SHOTS	21	4%	94%	SHOTS	9	2%	97%
RADAR	8	2%	96%	Gs	7	1%	98%
CLOSING VELOCITY	7	1%	97%	CLOSING VELOCITY	5	1%	99%
AOA	7	1%	99%	CIRCLE	2	0%	99%
ALTITUDE	6	1%	100%	AOA	2	0%	100%
CIRCLE	1	0%	100%	RADAR	1	0%	100%
TOTAL	489			TOTAL	537		

Table 13. Overall Frequency Count by Item for both Engagements

ENGAGEMENT ONE				ENGAGEMENT TWO			
ITEM	FREQ	%	CUM %	ITEM	FREQ	%	CUM %
3-9 POSITION ADVANTAGE	34*	7.0%	7.0%	3-9 POSITION ADVANTAGE	56*	10.4%	10.4%
NEAR PARAMETERS	31*	6.3%	13.3%	LIFT VECTOR CONTROL-GOOD	48*	8.9%	19.4%
LIFT VECTOR CONTROL-POOR	29*	5.9%	19.2%	NOSE ON	37*	6.9%	26.3%
ASPECT LESS THAN 60°	27	5.5%	24.7%	AIRSPEED DISADVANTAGE	30	5.6%	31.8%
NOSE ON	25*	5.1%	29.9%	NEAR PARAMETERS	22*	4.1%	35.9%
MAINTAINING ENERGY	24*	4.9%	34.8%	RANGE 1000 - 3500 Ft	20*	3.7%	39.7%
NOT USING G's	21	4.3%	39.1%	AIRSPEED ADVANTAGE	20*	3.7%	43.4%
AIRSPEED ADVANTAGE	21*	4.3%	43.4%	LOSING ENERGY	19	3.5%	46.9%
LIFT VECTOR CONTROL-GOOD	19*	3.9%	47.2%	ALTITUDE ADVANTAGE	17	3.2%	50.1%
3-9 LOST ADVANTAGE	18*	3.7%	50.9%	BOGIE IN WINDSCREEN	16*	3.0%	53.1%
RANGE 1000 - 3500 Ft	18*	3.7%	54.6%	AIRSPEED SAME	15	2.8%	55.9%
SWITCHES IN HEAT	12	2.5%	57.1%	SWITCHES IN GUNS	15*	2.8%	58.7%
BOGIE IN HUD FOV	12*	2.5%	59.5%	BOGIE NEAR CANOPY BOW	14	2.6%	61.3%

Table 13. Continued

ENGAGEMENT ONE				ENGAGEMENT TWO			
ITEM	FREQ	%	CUM %	ITEM	FREQ	%	CUM %
LOW G's	11	2.2%	61.8%	IN-RANGE GUN	14	2.6%	63.9%
IN PARAMETERS	11	2.2%	64.0%	MAINTAINING ENERGY	14*	2.6%	66.5%
SWITCHES IN GUNS	10*	2.0%	66.1%	IN PARAMETERS	13*	2.4%	68.9%
RANGE < 1000 Ft	9	1.8%	67.9%	3-9 LOST ADVANTAGE	10*	1.9%	70.8%
RANGE > 3500Ft	9	1.8%	69.7%	BOGIE IN HUD FOV	9	1.7%	72.4%
HEAT SHOTS	9	1.8%	71.6%	LIFT VECTOR CONTROL-POOR	8*	1.5%	73.9%
RADAR LOCK	8	1.6%	73.2%	ASPECT 60° to 90°	8	1.5%	75.4%
HIGH G's	8	1.6%	74.8%	LFT VECTOR BELOW HORIZON	8	1.5%	76.9%
GAINING ENERGY	8	1.6%	76.5%	RANGE < 1000 Ft	8	1.5%	78.4%
GUN SHOTS	8	1.6%	78.1%	LIFT VECTOR ON OPPONENT	8	1.5%	79.9%
BOGIE IN WINDSCREEN	7*	1.4%	79.6%	NOSE WITHIN 60°	7	1.3%	81.2%
AIRSPEED DISADV	6	1.2%	80.8%	ASPECT LESS THAN 60°	7	1.3%	82.5%
IN-RANGE GUN	6	1.2%	82.0%	NOSE IN LAG	7	1.3%	83.8%
LOSING ENERGY	6	1.2%	83.2%	ASPECT GREATER THAN 90°	7	1.3%	85.1%
HIGH CLOSING VELOCITY	5	1.0%	84.3%	NOSE IN LEAD	6	1.1%	86.2%
NOSE IN LAG	5	1.0%	85.3%	NOSE OFF	6	1.1%	87.3%
ASPECT 60° to 90°	5	1.0%	86.3%	GUN SHOTS	5	0.9%	88.3%
LFT VECTOR ABOVE HORIZON	4	0.8%	87.1%	ALTITUDE DISADVANTAGE	5	0.9%	89.2%
AIRSPEED < 200	4	0.8%	87.9%	ALTITUDE SAME	5	0.9%	90.1%
SWITCHES WRONG	4	0.8%	88.8%	SWITCHES WRONG	5	0.9%	91.1%
BOGIE NEAR CANOPY BOW	4	0.8%	89.6%	HIGH CLOSING VELOCITY	5	0.9%	92.0%
LOW AOA	4	0.8%	90.4%	GAINING ENERGY	5	0.9%	92.9%
LATE SHOTS	3	0.6%	91.0%	OUTSIDE PARAMETERS	4	0.7%	93.7%
HIGH AOA	3	0.6%	91.6%	LFT VECTOR ABOVE HORIZON	4	0.7%	94.4%
SAME G's	3	0.6%	92.2%	AIRSPEED < 200	3	0.6%	95.0%
ALTITUDE DISADVANTAGE	3	0.6%	92.8%	HIGH G's	3	0.6%	95.5%
AIRSPEED SAME	3	0.6%	93.5%	SWITCHES IN HEAT	3	0.6%	96.1%
IN-RANGE HEAT	3	0.6%	94.1%	ACCELERATING	3	0.6%	96.6%
NOSE IN LEAD	2	0.4%	94.5%	HIGH AOA	2	0.4%	97.0%
LOW CLOSING VELOCITY	2	0.4%	94.9%	SAME G's	2	0.4%	97.4%
GOD'S G's	2	0.4%	95.3%	HI ANGLE GUN SHOTS	2	0.4%	97.8%
LOST SIGHT OF BOGIE	2	0.4%	95.7%	LATE SHOTS	2	0.4%	98.1%

Table 13. Concluded

ENGAGEMENT ONE				ENGAGEMENT TWO			
ITEM	FREQ	%	CUM %	ITEM	FREQ	%	CUM %
NOSE WITHIN 60°	2	0.4%	96.1%	BOGIE LEFT	2	0.4%	98.5%
ASPECT GREATER THAN 90°	2	0.4%	96.5%	TWO CIRCLE FIGHT	1	0.2%	98.7%
NOSE OFF	2	0.4%	96.9%	ALTITUDE DIFFERENT	1	0.2%	98.9%
ALTITUDE ADVANTAGE	2	0.4%	97.3%	IN-RANGE HEAT	1	0.2%	99.1%
ONE CIRCLE FIGHT	1	0.2%	97.5%	RADAR LOCK	1	0.2%	99.3%
LIFT VECTOR ON OPPONENT	1	0.2%	97.8%	GOD'S G's	1	0.2%	99.4%
SWITCHES LATE	1	0.2%	98.0%	TOO SOON G's	1	0.2%	99.6%
LIFT VECTOR CLOSE	1	0.2%	98.2%	ONE CIRCLE FIGHT	1	0.2%	99.8%
OUTSIDE PARAMETERS	1	0.2%	98.4%	BOGIE RIGHT	1	0.2%	100.0%
MISSED SHOTS	1	0.2%	98.6%				
BOGIE RIGHT	1	0.2%	98.8%	TOTAL	537	100%	
A/S EXCESSIVE	1	0.2%	99.0%				
LFT VECTOR BELOW HORIZON	1	0.2%	99.2%				
NOSE LEVEL	1	0.2%	99.4%				
ACCELERATING	1	0.2%	99.6%				
HEAD-ON	1	0.2%	99.8%				
ALTITUDE SAME	1	0.2%	100.0%				
TOTAL	489	100%					

The second analysis attempted to determine whether specific categories and items were associated with specific state assignments. Frequency counts for each category across each state are presented in Tables 14 and 15 for Engagements One and Two, respectively. Again, the individual item data across each state can be found in Appendix D. The results for Engagement Two were fairly straightforward. As shown in Table 15, the state assignments for Engagement One were primarily neutral and offensive. The category "Range" was found to be cited most often for Offensive and Offensive Weapons states and only rarely for the Neutral states. "Lift Vector Control," on the other hand, was cited most often for the Neutral and Offensive states and much less often for the Offensive Weapons states. The "3-9 Position" was selected most often for the Offensive states and less often for the Offensive Weapons and Neutral states.

Table 14. Category Frequencies by State for Engagement One

CATEGORY	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	TOT	%	CUM %
	STATES																				
RANGE		15	17	4	3	3	1	2	3				17	6	2	9	5	1	88	18%	18.0%
LIFT VECTOR		1	1	3	1		1	3	2		2	4	4	8	2	13	6	4	55	11%	29.2%
3-9 POSITION		8	9	3	2	1	2	4	1		4	7	3	4		2	2		52	11%	39.9%
Gs		1	5	1		1	3	3			3	1	3	6	1	8	6	3	45	9%	49.1%
ENERGY		10	2	3	2		1				3	2	2	5		6	1	1	38	8%	56.9%
NOSE POSITION		7	11	3	3		3	1	1			1		3	1	2	1	1	38	8%	64.6%
AIRSPEED		2	4		2	1	1		4		6	2	2	6	1	3	1	1	36	7%	72.0%
ASPECT		2	12	2	1	1	3		2		1	2	1	3	1	2	1		34	7%	78.9%
SWITCHES		12	8	2	2									1			2		27	6%	84.5%
BOGIE POSITION		6	9		2	2	1	1	1			1	3						26	5%	89.8%
SHOTS		8	4	1	1									2		2	1	2	21	4%	94.1%
RADAR		5	1		1	1													8	2%	95.7%
Vc			4					2						1					7	1%	97.1%
AOA			3			1	1		1								1		7	1%	98.6%
ALTITUDE						1	1		1		2	1							6	1%	99.8%
CIRCLE			1																1	0%	100.0%
																			489	100%	

Table 15. Category Frequencies by State for Engagement Two

CATEGORY	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	TOT	%	CUM %
	STATES																				
RANGE	5	3	27	17	6	15	4	1		2						2			82	15%	15.3%
LIFT VECTOR	1	2	4	6	5	11	16	9	12	5	2	2				1			76	14%	29.4%
AIRSPEED	1	2	3	6	10	10	8	11	6	4	1	9							71	13%	42.6%
3-9 POSITION	1	3	7	10	6	16	7	7		2	4	2	1						66	12%	54.9%
NOSE POSITION	3	2	9	8	8	16	8	9											63	12%	66.7%
BOGIE POSITION		1	10	9	2	7	7	3	1	2									42	8%	74.5%
ENERGY		1		8	0	5	10	6	2	2	1	3							38	7%	81.6%
ALTITUDE		1	1	2	2	4	5	4	3	3	1	2							28	5%	86.8%
SWITCHES	2	2	8	2	5	3				1									23	4%	91.1%
ASPECT	2		2		6	1	3	5	1	1		1							22	4%	95.2%
SHOTS	3	1	2		3														9	2%	96.8%
Gs				1			3		2			1							7	1%	98.1%
Vc		1	1		1					1	1								5	1%	99.1%
CIRCLE							1		1										2	0%	99.4%
AOA												2							2	0%	99.8%
RADAR										1									1	0%	100.0%
																			537	100%	

Unfortunately, the results for Engagement One, which was by far the simpler of the two engagements, were not as clear. These results even suggest that there may have been some confusion for some subjects in view of their responses. For example, consider the category "Range" which is primarily concerned with weapons parameters. As such, it is reasonable to expect that items comprising this category be marked primarily during

Offensive and Offensive Weapons states. However, such was not the case for the Engagement Two data. As shown in Table 14, there is a somewhat even split between the Offensive and Defensive states. An examination of the individual item distributions (Appendix D) revealed, for example, that the item "Near Parameters" was given as a reason for the Defensive (D) state 16 times. The question arises, how can the rated combatant be "near parameters" and still be defensive? On the other hand, if the rated combatant's <u>opponent</u> were "near parameters," then indeed the correct state of the rated aircraft would be Defensive. Perusal of the other items revealed a similar pattern. For example, the item "switches in guns" was primarily associated with Offensive Weapons states, except in two cases in which it is given as the reason for a DW- state assignment. Clearly, these two responses make little sense unless one assumes that it is the combatant's <u>opponent</u> who has his "switches in guns" and therefore the rated aircraft is highly defensive. In view of these problems, no further attempts were made to relate specific categories and items to states within the model.

Questionnaire Findings

Results from the questionnaire revealed that all subjects judged the TSM to be "appropriate and useful" ($\bar{X} = 2.1$) and that when using the TSM they were able to "characterize all of the tactical states in all of the time sequences" ($\bar{X} = 1.8$). Responses to the open-ended questions reflected similar favorable reactions to the TSM.

Subjects were less enthusiastic about use of the TPR scale. Subjects mostly agreed that the scale did "provide the opportunity to rate how well a pilot accomplished his task" ($\bar{X} = 2.0$) and they also acknowledged that the "ACM-PMS provided adequate information to assign goodness-of-choice in maneuvers" ($\bar{X} = 2.1$). However, when asked if the "numerical rating scale would be a useful tool for training purposes," subjects were decidedly neutral ($\bar{X} = 2.9$).

Since this was the first research use of the ACM-PMS as a debrief station, subjects were also asked some general questions about their satisfaction with the system. In response to these queries, subjects agreed that the ACM-PMS was a "good debriefing tool" ($\bar{X} = 1.6$), and that "all the necessary information was available to assign tactical state" ($\bar{X} = 1.5$). The responses to the open-ended questions indicated favorable reactions towards the system as well.

Discussion

Tactical State Assignment

The interrater correlations for tactical state assignments demonstrated excellent agreement among subjects for Engagement One, but not for Engagement Two. In fact, these differences were rather pronounced. It seems likely that these differences are largely due to differences in the two engagements. As mentioned earlier, the primary difference in the two engagements lies in the number of times the aircraft cross the neutral states in the model. The first engagement uses almost the entire spectrum of the TSM in a linear fashion. In other words, the rated aircraft starts off neutral, goes offensive and then defensive and is eventually killed. In contrast, the second engagement crosses through neutral a number of times before the rated aircraft becomes definitely offensive and makes the kill. It was anticipated that Engagement Two would be more difficult for raters to characterize and the lower interrater correlations confirm this difficulty. It also reinforces the need for objective definitions of the tactical states if accurate performance measurement is to be achieved.

Although the correlations for the most part were quite strong for state assignment, there were certain time frames in the engagements with reduced levels of agreement among raters. One possible explanation is that some subjects may have entered into a mind set of being offensive, neutral or defensive. As a result, when the rated aircraft began to transition out of one stage of the model (e.g., from offensive states to neutral ones), a lag for changes in the state rating occurred. These subjects appeared to require some major event, such as an overshoot, to finally shift to the other side of the model. It appeared that this occurred throughout the offensive, defensive and neutral sections of the TSM. It is possible that with additional training and experience using the model, this problem could be avoided and the level of agreement improved on relatively complex engagements such as Engagement Two.

An important research question was whether or not the TSM represented an appropriate framework for the characterization of ACM tactical states. The data gathered, albeit strictly opinion, was very positive. In response to this inquiry on the questionnaire, all subjects stated that the model made good intuitive sense to them and that the framework was sound. In fact, subjects indicated that the model reminded them of other training metrics, such as the BFM Ladder, that they had used while in Fighter Lead-In training. When asked if they believed that the TSM could adequately describe all of the situations that they had viewed in the two engagements, subjects responded that the model was in fact comprehensive enough to characterize all situations.

Relationship of TSM Assignments to ROMP

The results of the analysis correlating the state ratings to actual ROMP scores showed a high level of agreement only for Engagement One. In this case, a strong relationship was found between subjects' state ratings and ROMP scores, with over 90% of the variance in subjects' ratings explained by the ROMP scores. This finding would indicate that there was strong association between the All-Aspect Maneuvering Index algorithm, which produces the ROMP curve, and the state ratings assigned by subjects. Unfortunately, Engagement Two produced no significant relationships, although the trends were in the expected direction. It seems likely that the characteristics of the engagement previously discussed may have influenced these results at least to some degree. Clearly, Engagement Two was much more difficult to assess as evidenced by the reduced reliabilities for both TSM and TPR assignments. Since reliability represents an upper limit for the validity of a measure, such reductions in correlations would be expected, although not to the levels found in this investigation. Obviously, further work is required to establish the relationships between TSM assignments and the ROMP.

The results of these analyses are considered important although they do not bear directly on the stated goals of this research effort. The positive relationship between the subjective TSM assignments and the computationally-based objective measure, at least for one of the engagements, served as a demonstration of the concurrent validity of the ROMP measure as an indicator of positional advantage during ACM. These results suggest that the ROMP may be a valid representation of the relative advantage or disadvantage of an aircraft engaged in 1 v 1 aerial combat. In fact, it may be that the ROMP is a more accurate indicator of the relative advantage of two combatants than could be ascertained by any subject matter expert, since it relies strictly on empirical calculations. These results also suggest that the ROMP may be useful in predicting other criterion measures of ACM performance such as engagement outcome. In fact, the AAMI algorithm which produces the ROMP has been shown to be a strong indicator of engagement outcome (Waag, Leeds, and Raspotnik, 1991). The results of the present work suggest that the ROMP be evaluated against other criteria.

If further work were to validate the ROMP as a measure of positional advantage, it is possible that it may have a great deal of potential as a training aid. By itself, the ROMP is only an index of positional advantage and consequently cannot be used in a diagnostic sense to judge students' performance in terms of mistakes, errors, good moves, etc. However, it may provide the instructor a means of gauging the results of various maneuvers that the student attempted in terms of its effects upon positional advantage. In this way, the instructor would have a means of providing objective feedback and clearly demonstrating the effectiveness or ineffectiveness of the student's maneuvering.

Tactical Performance Ratings

Overall, the interrater correlations for the TPRs were not as strong as those for the tactical state ratings. Like the TSM results, the reliabilities obtained for Engagement Two were much lower than those for Engagement One. This was probably due to the more ambiguous nature of Engagement Two as discussed earlier. This seems to indicate that Engagement One was much easier to characterize than Engagement Two, regardless of the type of judgement that was being made.

There are several possible reasons which may account for the lower level of agreement for the TPR when compared with TSM assignments. The first problem could be with the scale itself. One difficulty with the TPR scale was that it lacked a definitive frame of reference or set of anchors for the rater to use. As mentioned earlier in this paper, subjects indicated that the TSM made good intuitive sense to them in a general way, reminding them of other metrics (BFM Ladder) that they had used. On the other hand, the TPR scale possesses no such past frame of reference, and this could have resulted in reduced agreement among subjects. Even though the IPs use a 6-point scale for grading students during the course of their normal instruction, none of them mentioned this as being a reference point for rating tactical performance.

Change of State and Performance

The results indicated that there was no clear relationship between change of state and rated performance. Although some of the correlations achieved statistical significance, there was little consistency. Of some interest, however, was the finding that the averages across raters did produce significant results for both engagements. This would lend some support to the initial hypothesis that achieving a better tactical position, i.e., positive state changes, is positively related to the "quality" of maneuvering. Unfortunately, these results were obtained only for the "group" data and not for the individual subjects with the exception of one subject as previously discussed. It seems likely that the obtained reliabilities of TSM and TPR scales contributed to such findings.

Rationale for State Assignments

Analysis of the underlying rationale or reasons given by subjects for making TSM assignments proved to be inconclusive, although a number of interesting findings did emerge. For example, considering the entire data set, the categories selected most often were "Range," "3-9 Position," and "Lift Vector Control." These findings are certainly reasonable from the standpoint that the goal of the combat pilot is to maneuver his aircraft into a positional advantage so that weapons may be employed. One result of particular interest concerned maneuvering performance. Although other "performance" items could be selected (airspeed, G, closing velocity, angle of attack, altitude), the element of maneuvering chosen most often was "lift vector control." In addition to this category being selected most often, two specific items in the category, "lift vector control-good" and "lift vector control-poor" appear to be synonymous with offensive and defensive positions. During Engagement One, "lift vector control-poor" was selected 29 times, 26 of these (89%) were associated with defensive states (see Appendix D). During Engagement Two, on the other hand, "lift vector control-good" was selected 48 times, 38 of these (79%) associated with offensive states. Often during ACM instruction, an instructor will say "roll to put your lift vector on him and pull." The results of this analysis appear to confirm that instructors do indeed believe that lift vector control is a key factor in ACM performance.

For individual items, "3-9 advantage" was selected most often in both engagements. It was ranked first in importance in both engagements and was also selected most often in describing offensive states. Again, this was no surprise, since the basic ingredient in determining tactical state is the relative position of the players and 3-9 position is perhaps the simplest and most often cited means of describing relative position.

Some differences also occurred as a function of the engagements used in the experiments. As mentioned, "G" was cited 45 times for Engagement

One but only 7 times for Engagement Two. This category appears to be one that is primarily associated with defensive states. In Engagement One, the rated aircraft spends about one half of the engagement in defensive states. In fact, of the 45 times Gs were selected, 31 of those were associated with defensive states. During Engagement Two, the rated aircraft was never clearly defensive and there is a corresponding decrease in the number of times that Gs were selected. Likewise, there were also "Airspeed" differences with roughly twice as many selections of "Airspeed" items for Engagement One.

A second part of the analysis attempted to determine whether specific categories and their respective items were related to state assignment. In order to understand the data, all items were categorized as either "position," "performance," or "weapons employment." Unfortunately, the results failed to clearly link these categories to state assignment. While some trends were evident for Engagement Two, there were numerous inconsistencies for Engagement One. As previously mentioned, it seems likely that there was some confusion by some subjects in terms of how the checklist was to be used. For this reason, the results of these analyses are inconclusive. The current data can only point to certain "trends." It would appear, for example, that "Position" items were used most often to describe offensive states, while "Performance" items were used most to describe neutral/defensive states. And as expected, "Weapons Employment" items were used most often to describe either Offensive Weapons or Defensive Weapons states.

A speculative hypothesis for the association of "Position" items with offensive states and "Performance" items with neutral/defensive states is that the rater tends to identify with the rated aircraft when it is offensive, which is "good," and subsequently the rater selects those factors that describe the aircraft's position. On the other hand, when the rated aircraft is defensive, which is "bad," the rater tends to become diagnostic and use performance factors to describe how the aircraft became defensive or what it is doing wrong rather than describing where it is. This hypothesis grew out of the observation of the subjects while performing the rating task. When the rated aircraft was doing good (i.e., offensive), subjects often maintained an attitude of "we" are doing this and "we" are doing that. However, when the rated aircraft was doing poorly (i.e., defensive), the subjects were more likely to say that "he" was doing this and "he" was doing that.

Although the overall results may be considered "reasonable" and consistent with our general understanding of ACM, they fail to provide a great deal of additional insight. In fact, it can be argued that the data do little more than confirm the obvious. Certainly, in terms of the original goal of this analysis, namely, identifying quantitative parameters that are related to tactical state assignment, the results provide little useful information. The rationale for developing the checklist was to provide subjects with a common set of terms and to also facilitate the data analysis. While this was accomplished to some degree, it is apparent that such "simplification" also produced somewhat "simplistic" results. It is clear that any attempt to gain insight and further understanding into what may be considered as one of the most complex of human behaviors will require a different approach.

GENERAL DISCUSSION

We turn now to a general discussion of the results of the two experiments in terms of how they relate to the original objectives. At this point, it may be instructive to briefly restate the underlying rationale for the approach taken in this investigation. Three major questions were addressed. The first question was whether experts can reliably assign tactical states at specific moments or points in an ACM engagement. Second, could they also reliably assign a rating to the maneuvering which took place between these points? And third, does movement in one direction or the other within the TSM correlate with "good" or "bad" TPRs? If the results indicated positive answers to these questions, it was reasoned that the assignment of a tactical state to end points of segments and a TPR to the maneuvering in between those points would be a powerful combination. A positive correlation between movement within the model and TPRs should lead directly to the quantification of "good" or "bad" performance. Further examination of the points or segments where movement within the model either stops of changes direction should provide the specific information necessary to accomplish diagnosis and remediation of poor ACM performance. Discussion now turns to the results in terms of these three issues.

Tactical State Characterization

The first question considered was whether tactical state could be reliably characterized. In other words, can pilots reliably assign states using the TSM? The data gathered in these two experiments indicate the answer is "yes", but only for fairly straightforward, simple engagements. As the difficulty of the engagement increased, the reliability of state assignments was found to diminish. In fact, the single rater reliability for Engagement Two in the second experiment was only .70, a level generally considered inadequate for practical use (Nunnaly, 1967). On the other hand, reliability estimates when raters were considered as a group were quite high, i.e., greater than .90 in both experiments. However, from a practical standpoint, it is highly unlikely that a group of six or eight IPs would ever be available to evaluate single engagements during training. It is unknown whether additional refinement of the model or perhaps additional training would lead to improvements in its reliability for individual raters.

It is suggested that a more fruitful approach to characterizing tactical states would be to concentrate on empirically-derived measures. The positive correlations between the ROMP and TSM assignments for Engagement One certainly support such a recommendation. Although no significant correlations were found for Engagement Two, a perusal of the actual time series data presented in Figure 5 suggests that a relationship nonetheless does exist. It is clear that one of the problems for Engagement Two was a restriction of range. In other words, most of the engagement stayed fairly neutral and consequently did not exercise the full range of either the ROMP or TSM. For this reason, the observed correlations would be reduced. Moreover, since the single rater reliability for this engagement was also low, the correlation would further be restricted in terms of its upper bound.

Furthermore, there appear to be some anomalies. For example, the ROMP score for the last sample point in the engagement was near zero, despite the fact that a kill occurred. In actuality, a high-aspect gun kill occurred from a position in which both combatants were offensive. Consequently, the ROMP score would be considered "correct" up to the point of time that the kill occurred. Likewise, the "correct" TSM rating should have been "neutral--both at risk" until the kill occurred. It should also be pointed out that the shifting back and forth about the zero or neutral point is also evident in the TSM ratings, although the magnitudes differ and there appear to be certain "lags" in some of the raters. Again, due to the difficulty of the engagement, such variations would not be unexpected. For these reasons, it seems reasonable to conclude that some empirical measure of positional advantage would be preferred to the use of the TSM scale. Moreover, there is no reason that such an empirically-based measure could not be "displayed" in TSM terminology thereby maintaining face validity with the operational user.

A further concern at the outset of this investigation was to obtain information on the pilot's reason or rationale for making specific state assignments. It was hoped that such queries would enable the identification of specific flight parameters that would characterize each state. In Experiment One pilots simply provided a verbal rationale for selecting a particular state which was recorded. This procedure was abandoned due to the complexities of data reduction, and subsequently in Experiment Two, a checklist was used. While the data from the second experiment was easier to quantify, the results were not particularly insightful. As stated before, such results did little more than to confirm the obvious. Items relating to weapons parameters were cited most often in the Offensive Weapons and Defensive Weapons states. Items reflecting position were most cited in the Offensive/Defensive states while items reflecting maneuvering performance were cited most often in the Neutral states. The one useful finding to come from these data was the importance of lift vector control as a candidate measure of maneuvering performance. In the continuing effort to develop measures of performance that not only relate to overall performance but have potential for diagnosing elements of poor performance, this finding may prove to be one of the most important results for the entire investigation.

Tactical Performance Ratings

The second major question in this investigation was whether IPs could reliably assign a TPR to judge the "goodness" or "quality" of maneuvering performance. These results were perhaps the most disappointing. Obtained single rater reliability estimates were .34 and .22 for Experiment One and .69 and .31 for Experiment Two. Even the highest of these, which was for Engagement One of Experiment Two is considered to be too low to be of practical value. Although group reliabilities were higher, there remains the practical limitations of multiple raters. While it seems likely that the addition of anchor points for Experiment Two did improve the reliability somewhat, these improvements were only marginal.

In light of the existing literature on the topic, these results are not atypical. In fact, the interrater agreement achieved in this investigation appears to be similar to that normally achieved for this type of assessment exercise (Borman, 1978). This would support the hypothesis suggested by Mount and Thompson (1987) that there may be strict limits on the possible level of interrater agreement for performance assessment. This hypothesis is supported by considerable experimental evidence (Cooper, 1981; Feldman, 1981; and Landy & Farr, 1980).

Second, it is possible that individual differences among the raters may have accounted for such low agreement. For example, Mount and Thompson (1987) speculate that it is not so much the performance metrics themselves, but rather individual differences which create the low interrater agreement on performance. If this is the case, then any further revision to the TPR Scale would be questionable. In fact, Feldman (1981) has called for a moratorium on further development of performance scales because of the low levels of return. Borman (1978) states that even in a perfect rating environment, agreement will be far less than perfect. Borman (1978) further adds that the development of perfect error-free performance rating scales may be an unattainable goal.

A third possible reason for the lack of agreement may have been insufficient training using the scale. Specialized training is a strategy that has been tried in many research efforts for ameliorating poor interrater agreement levels (Hedge & Kavanagh, 1988). Although some authors have proposed specific techniques that improved rater training in experimental conditions (Bernardin & Pence, 1980), the literature has shown inconsistent results. For example, Borman (1979) was able to reduce halo effects in college students with specific instruction. However, while this worked for one job performance scenario, it did not for another. Bernardin & Pence (1980) were also able to eliminate the halo effect with specific training, but in doing so, the overall accuracy of raters was reduced.

Changes in Tactical State and TPR Assignments

The final question to be addressed was whether changes in tactical state were related to maneuvering performance. Unfortunately, this question could not be equivocally answered due to the low reliability of the TPR assignments. Nevertheless, the analyses were performed and as might be expected, no consistent relationships were found. It is of interest that only the average state changes and TPR assignments produced significant correlations for both engagements. These findings provide some support that gaining offensiveness is positively related to maneuvering performance, while becoming increasingly defensive is negatively related to the quality of performance. It seems likely that the failure to obtain such relationships on an individual subject basis was due largely to the poor reliability of the TPR scale. This combined with the decreased reliability of any difference or change score seem to be the most likely reasons for failure to obtain the hypothesized relationships.

CONCLUSIONS

The primary goal of this research investigation, namely the development of a technique for assessing tactical decision making during ACM, was not accomplished. The three questions posed at the outset of the investigation were successfully addressed only to a very limited degree. First, the tactical state model developed in this investigation was only partially successful in being able to reliably characterize position during an air-to-air engagement. Second, the quality of maneuvering performance could not be reliably measured using the TPR scale developed in this investigation. In all instances, the observed reliabilities were quite low. And third, the research failed to establish consistent relations between changes in position and quality of performance on an individual subject basis, although the group data produced correlations in the expected direction. Based on these findings, it is concluded that development of a diagnostic measure of tactical decision making will require a different approach.

Despite these negative findings with regard to the original intent of the investigation, the results do point to several fruitful areas for further work. Specifically, the observed correlation between subjective characterization of states by experts and an objective measure of performance suggests that future work should focus on empirically-derived measures of positional advantage. In fact, the current investigation provides some degree of concurrent validity for the measure that was evaluated. Moreover, the finding that lift vector control was one of the most frequently cited reasons for assigning a tactical state suggests that it should be considered as a candidate measure in future measurement development efforts.

REFERENCES

Bernardin, H. J., & Pence, E. C. (1980). Effects of rater training: Creating new response sets and decreasing accuracy. Journal of Applied Psychology, 65, 60-66.

Borman, W. C. (1978). Exploring the upper limits of reliability and validity in job performance ratings. Journal of Applied Psychology, 63, 135-144.

Borman, W. C. (1979). Format and training effects on rating accuracy and rater errors. Journal of Applied Psychology, 64, 410-421.

Brecke, F.H. & Miller, D.C. (1991). Aircrew performance measurement in the air combat maneuvering domain: A critical review of the literature. (AL-TR-1991-0042). Williams AFB AZ: Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

Ciavarelli, A. P., Williams, A. M., & Pettigrew, K. W. (1981). Performance assessment and appraisal system (PAAS) (NAVTRAEQUIPCEN 78-C-0136-1). Orlando FL: Naval Training Equipment Center.

Cooper, W. H. (1981). Ubiquitous halo. Psychological Bulletin, 90, 218-244.

Dynkin, E. B. (1982). Markov processes and related problems of analysis. New York, NY: Cambridge University Press.

Feldman, J. (1981). Beyond attribution theory: Cognitive processes in performance appraisal. Journal of Applied Psychology, 66, 127-148.

Guilford, J.P. (1954). Psychometric Methods. New York: McGraw-Hill Book Co.

Hedge, J. W., & Kavanagh, M. J. (1988). Improving the accuracy of performance evaluations: Comparison of three methods of performance appraiser training. Journal of Applied Psychology, 73, 68-73.

Landy, F. J., & Farr, J. L. (1980). Performance rating. Psychological Bulletin, 87, 72-107.

McGuinness, J., Forbes, J. M., & Rhoads, J. E. (1984). Air combat maneuvering performance measurement system design (AFHRL-TP-83-56, B081060L). Williams Air Force Base, AZ: Air Force Human Resources Laboratory.

Meister, D. (1971). Human factors: Theory and practice. New York, NY: John Wiley and Sons, Inc.

Mount, M. K., & Thompson, D. E. (1987). Cognitive categorization and quality of performance ratings. Journal of Applied Psychology, 72, 240-246.

Nunnaly, J.C. (1967). Psychometric Theory. New York: McGraw-Hill.

Oberle, R. A. (1974). An air combat maneuver conversion model (Report No. CRC 274). Washington, DC: Office of Naval Research.

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

Waag, W.L., Leeds, J.L. & Raspotnik, W.B. (1991). The prediction of engagement outcome during air combat maneuvering. In Proceedings of the Sixth International Symposium on Aviation Psychology. Columbus OH: The Ohio State University.

APPENDIX A

QUESTIONNAIRE

<u>QUESTIONNAIRE</u>

Date: ____/____/____ Time: _________ Subj. # _______________________

Name and Rank ___________________________ Squadron ______________________

Total Flying Time ________________ Total F-15/F-16 Time __________________

Items 1-6 reference the State Matrix Model. Please mark your response on the continuum line. 1 indicates that you strongly agree and 5 indicates that you strongly disagree with the statement.

1. The State Matrix Model was useful to characterize tactical states.

   1 (strongly agree) — 2 — 3 — 4 — 5 (strongly disagree)

2. Overall, the State Matrix Model was realistic.

   1 (strongly agree) — 2 — 3 — 4 — 5 (strongly disagree)

3. Overall, the State Matrix Model was easy to use when characterizing tactical states in an ACM engagement.

   1 (strongly agree) — 2 — 3 — 4 — 5 (strongly disagree)

4. The State Matrix Model was sufficient enough to characterize <u>all</u> of the tactical states which you viewed.

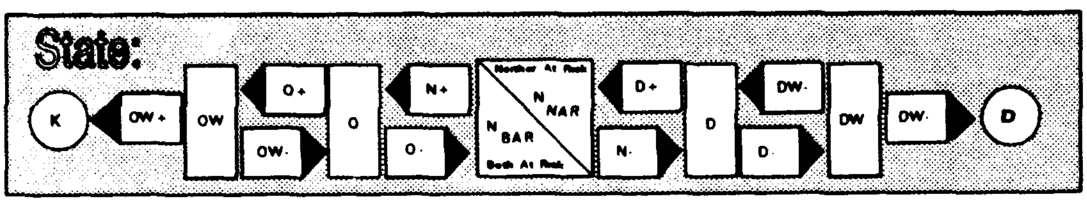
   1 (strongly agree) — 2 — 3 — 4 — 5 (strongly disagree)

5. If you did not agree with the statement in Item 4, please comment specifically on those tactical states you feel the State Matrix Model could not characterize.

   ____

   ____

6. If you could revise the State Matrix Model, what would you move, change, add or subtract? Mark the model below in any way that illustrates your point. (Skip if you wouldn't make changes)



Items 7-9 contain items pertaining to the numerical scale you used to characterize ACM maneuvering.

7. Overall, the numerical rating format provided the opportunity to rate how well a pilot accomplished their task, in terms of effecting a maneuver to produce a tactical state change?

1	2	3	4	5
strongly agree				strongly disagree

8. The ACM PMS provided enough information necessary to review this "goodness-of-choice" in maneuvers to effect state changes.

1	2	3	4	5
strongly agree				strongly disagree

9. The numerical rating scale would be a useful tool for training purposes.

1	2	3	4	5
strongly agree				strongly disagree

Items 10-21 provide questions concerning the ACM PMS hardware itself.

10. The ACM PMS has great potential as a briefing tool.

1	2	3	4	5
strongly agree				strongly disagree

11. The ACM PMS has great potential as a debriefing tool.

1	2	3	4	5
strongly agree				strongly disagree

12. What would you change, add or subtract from the ACM PMS to make it a better briefing/debriefing tool?

____

____

13. The ACM PMS provided all necessary information to completely review an ACM engagement.

1	2	3	4	5
strongly agree				strongly disagree

14. What displayed information did you use most to evaluate the pilots ACM ability?

a. Graphic depiction of both aircraft
b. Numerical outputs
c. Out-of-cockpit views
d. Other ____________________

15. Overall, what did you like about the display?

____

____

16. Overall, what didn't you like about the display?

____

____

17. This display would be an effective way to present multi-ship (e.g., 4 v 4) engagements.

1	2	3	4	5
strongly agree				strongly disagree

18. If you disagreed with Item 17, add specific suggestions which could help us adapt this display for multi-ship use.

____

____

19. Please add any additional comments or suggestions that you may have about this experiment or the equipment that you used. Use back of sheet if more space is required.

____

____

APPENDIX B

TACTICAL STATE AND PERFORMANCE DATA

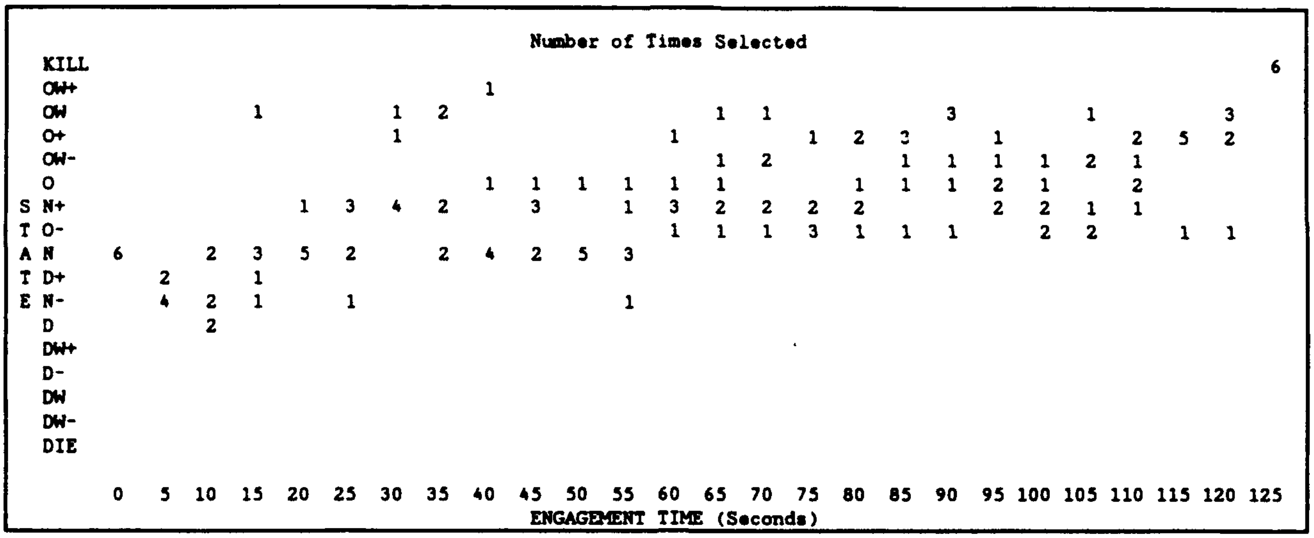


Figure B1. Frequency of TSM Assignments: Experiment One, Engagement One.

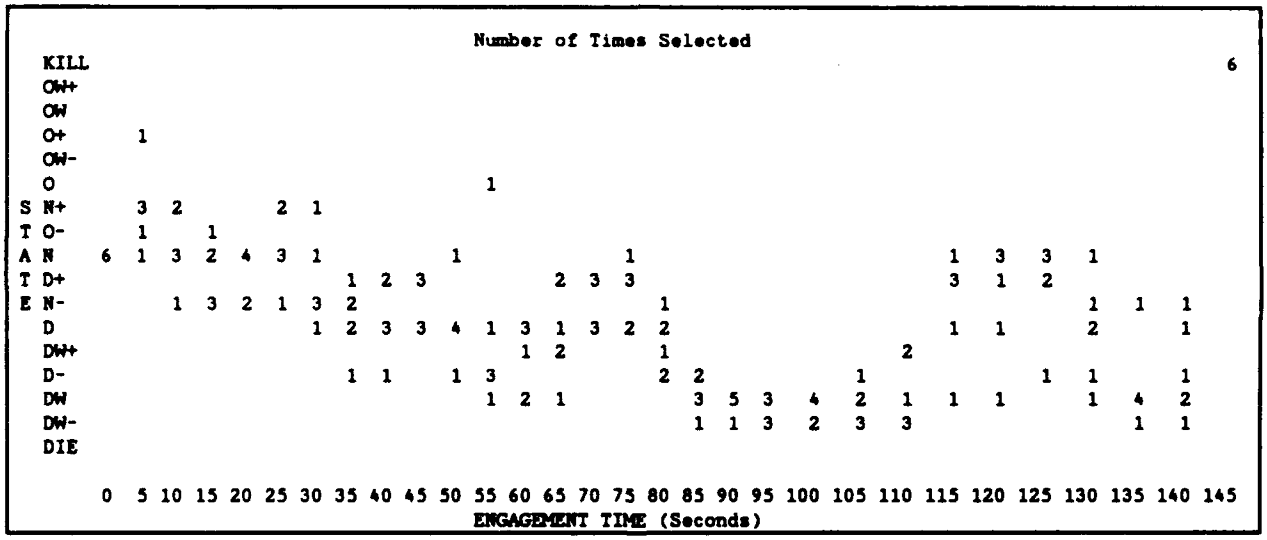


Figure B2. Frequency of TSM Assignments: Experiment One, Engagement Two.

Number of Times Selected

RATING	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100	105	110	115	120	125
3			3	3	3	2	2	1	2	3	4	4	4			3	4	2	1		3	1	1	4	2	4
2				1		3	3	2	1	1		2	2	1	2	3	2	4	1	4	2	4	2	1	3	1
1					3	1	1	2	2		2			3	2				2	1		1	3	1	1	
0	6			2					1	2				2	1				1		1					1
-1		3	2												1				1	1						
-2		1						1																		
-3		2	1																							

ENGAGEMENT TIME (Seconds)

Figure B3. Frequency of TPR Assignments: Experiment One, Engagement One.

Number of Times Selected

RATING	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100	105	110	115	120	125	130	135	140	145
3		4	2	3	1	2			1	2			2		1	2	1	1	1	1			1	3	2	3	1	4	5	3
2		2		1	3				2		1			4	2	1	1				3		2	1	1	1	4	1	1	1
1			1		1	1	1			2	1	2	2		2				1		2	1	1		1					
0	6		1	1			2		1	1	2		1	1				2				2	1	2		1		1		1
-1				1	1	2	2	1		1	2	1			1	2	2	1	2	2		1	1							
-2			2			1	1	1	1			1	1	1		1	1	1	1	1		1			2	1				
-3								4	1			2					1	1	1	2	1	1					1			1

ENGAGEMENT TIME (Seconds)

Figure B4. Frequency of TPR Assignments: Experiment One, Engagement Two

Number of Times Selected

STATE	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100	105
KILL																						
OW+					3	4	5	3	6													
OW				4	3	3	2	4	1	3												
O+		1	2	1	1	1	1		1													
OW-				1				1		3												
O		1		2																		
N+		4	1								1											
O-		2	3							2	1											
N	8		2								1	1										
D+												3	2	2		1						
N-					1						5	2	2				1	1				
D												1	3	5	1		1	1	1			
DW+												1	1		2	1	5	3	1			
D-														1	3			1	1	2		
DW															1	3	1	2	5	4	2	
DW-															1	2				2	6	
DIE																1						8

ENGAGEMENT TIME (Seconds)

Figure B5. Frequency of TSM Assignments: Experiment Two, Engagement One

Number of Times Selected

STATE	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100	105	110	115	120	125	130	135	140
KILL																													8
OW+																			2			2	1	1	1				
OW										1									3		4	2	5	3	1				
O+									1		1	2	1	2	1	1		2		1	2	1	1	3	2	2	1	2	
OW-													1	1			1	1	2	3	1	1	1		2	3		1	
O		1	1	3				1		3	2	2	3		1	1	1	4	1	4		2		1	2	1	2	1	
N+		1	4	3	3		2	5	4	2	2	1					3										1	2	
O-				2	2			1	3	1	2	3	1	1			1	1			1					2		1	
N	8	5	3		1	3	3	1		1	1		2	2	1	4	2										2	1	
D+														1	1	1											2		
N-		1			2	3	3							1	3	1													
D															1														
DW+																													
D-																													
DW						2																							
DW-																													
DIE																													

ENGAGEMENT TIME (Seconds)

Figure B6. Frequency of TSM Assignments: Experiment Two, Engagement Two

Number of Times Selected

RATING	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100	105
R 3			2	4				1														
A 2		5	3	3	3	2	2	1				1										
T 1		3	2	1	4	2	2	2	1			3	2		1							
I 0	8					1	2	1	1						1	2	2	1	1			
N -1						1	1	1	4	1		3	3		1	3	1		1	1		
G -2			1		1	1	1	2	2	6	3		3	2	3	1	2	2	1	1	1	1
-3						1				1	5	1		6	2	2	3	5	5	6	7	7

Figure B7. Frequency of TPR Assignments: Experiment Two, Engagement One

Number of Times Selected

RATING	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70
R 3											1				
A 2		3	6	8	1		2	6	3	2	2	5	1		1
T 1		5	2		2	2	5	2	3	3	2	2	1	3	4
I 0	8				3	2			1	1	1			3	2
N -1					1	3			1	2	2	1	3	1	1
G -2					1		1						3	1	
-3						1									

RATING	75	80	85	90	95	100	105	110	115	120	125	130	135	140
R 3			4					1					2	3
A 2	3	6	3	3	2	4	2	2	2	2	1	2	4	3
T 1	3	1		1	2	3	2		1	2	1	3	2	2
I 0	1	1		1	3	1		1	2	4		2		
N -1	1		1	2	1			3	2					
G -2							3	1	1		5	1		
-3				1			1				1			

Figure B8. Frequency of TPR Assignments: Experiment Two, Engagement Two

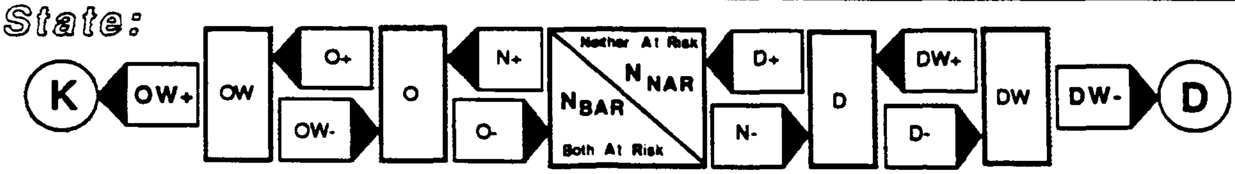
APPENDIX C

DATA COLLECTION FORMS

State Assignment and Rationale

Experiment: Engagement: Airplane: Subject: Time:

State:



Reasons:

Bogie Position
- In Windscreen ______
- In HUD FOV ______
- Near Canopy Bow ______
- Left ______
- Right ______
- Lost Sight ______

Nose Position
- Nose On ______
- Nose Off ______
- In Lead ______
- In Lag ______
- Within 60° ______
- Head-On ______
- Level ______

3-9 Position
- Advantage ______
- Lost Advantage ______

Aspect
- Less than 60° ______
- 60° to 90° ______
- Greater than 90° ______

Circle
- One Circle Fight ______
- Two Circle Fight ______

Energy
- Gaining ______
- Maintaining ______
- Losing ______

Lift Vector Control
- Good ______
- Poor ______
- Above Horizon ______
- Below Horizon ______
- On Opponent ______
- Close ______

Other

Airspeed
- Low < 200 ______
- Hi > 450 ______
- Excessive ______
- Speed Advantage ______
- Speed Disadvantage ______
- Speed Same ______
- Accelerating ______

Altitude
- Advantage ______
- Disadvantage ______
- Different ______
- Same ______

Gs
- High ______
- Low ______
- Same ______
- Not Using ______
- Too Soon ______
- God's G ______

AOA
- High ______
- Low ______
- Good ______
- Same ______

Range
- <1000 Ft ______
- 1000 to 3500 Ft ______
- >3500 Ft ______
- In Parameters ______
- Out of Parameters ______
- Near Parameters ______
- Gun ______
- Heat ______

Switches
- In Guns ______
- In Heat ______
- Wrong ______
- Late ______

Radar
- Lock ______
- No Lock ______
- Close to Lock ______

Shots
- Gun ______
- Heat ______
- Hi Angle Gun ______
- Miss ______
- Late ______

Closing Velocity
- High ______
- Low ______

Tactical Performance Rating

Experiment:	Engagement:	Airplane:	Subject:	Time:

Quality:

+1 ACCEPTABLE maneuver or ACCEPTABLE execution	0 Little or no action taken, maintaining status quo	−1 POSSIBLY WRONG maneuver or WEAK execution
+2 GOOD maneuver or GOOD execution		−2 WRONG choice of maneuver or BAD execution
+3 PERFECT maneuver or PERFECT execution		−3 ABSOLUTELY WORST choice of maneuver or WORST execution

Reasons:

Range
- <1000 Ft ______
- 1000 to 3500 Ft ______
- >3500 Ft ______
- In Parameters ______
- Out of Parameters ______
- Near Parameters ______
- Gun ______
- Heat ______

Switches
- In Guns ______
- In Heat ______
- Wrong ______
- Late ______

Radar
- Lock ______
- No Lock ______
- Close to Lock ______

Shots
- Gun ______
- Heat ______
- Hi Angle Gun ______
- Miss ______
- Late ______

Closing Velocity
- High ______
- Low ______

Bogie Position
- In Windscreen ______
- In HUD FOV ______
- Near Canopy Bow ______
- Left ______
- Right ______
- Lost Sight ______

Nose Position
- Nose On ______
- Nose Off ______
- In Lead ______
- In Lag ______
- Within 60° ______
- Head-On ______
- Level ______

3-9 Position
- Advantage ______
- Lost Advantage ______

Aspect
- Less than 60° ______
- 60° to 90° ______
- Greater than 90° ______

Circle
- One Circle Fight ______
- Two Circle Fight ______

Energy
- Gaining ______
- Maintaining ______
- Losing ______

Lift Vector Control
- Good ______
- Poor ______
- Above Horizon ______
- Below Horizon ______
- On Opponent ______
- Close ______

Other

Airspeed
- Low < 200 ______
- Hi > 450 ______
- Excessive ______
- Speed Advantage ______
- Speed Disadvantage ______
- Speed Same ______
- Accelerating ______

Altitude
- Advantage ______
- Disadvantage ______
- Different ______
- Same ______

Gs
- High ______
- Low ______
- Same ______
- Not Using ______
- Too Soon ______
- God's G ______

AOA
- High ______
- Low ______
- Good ______
- Same ______

APPENDIX D

ITEM AND CATEGORY FREQUENCY COUNT DATA

Table D1. Frequency count by Item and Category: Experiment Two, Engagement One.

CATEGORY	ITEM NAME	STATE 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	TOTALS
AIRSPEED	1 < 200														2		2			4
	2 > 200																			0
	3 EXCESSIVE			1																1
	4 ADVANTAGE		2	2		2				2		6	1	1	3	1			1	21
	5 DISADVANTAGE						1	1		1			1		1			1		6
	6 SAME			1						1				1						3
	7 ACCELERATING																1			1
	CATEGORY TOTALS	0	2	4	0	2	1	1	0	4	0	6	2	2	6	1	3	1	1	36
RANGE	8 < 1000 Ft		3	2		2			2											9
	9 1000 - 3500 Ft			7						3					2		3	3		18
	10 > 3500Ft		1	1	1	1	1	1						1	1		1			9
	11 IN PARAMETERS		4	3			1										1	2		11
	12 OUT OF PARAMETERS						1													1
	13 NEAR PARAMETERS		2	2	2									16	3	2	4			31
	14 IN-RANGE GUN		2	2	1														1	6
	15 IN-RANGE HEAT		3																	3
	CATEGORY TOTALS	0	15	17	4	3	3	1	2	3	0	0	0	17	6	2	9	5	1	88
BOGIE POSITION	16 IN WINDSCREEN		1	1		1	2	1		1										7
	17 IN HUD FOV		5	7																12
	18 NEAR CANOPY BOW			1		1			1					1						4
	19 LEFT																			
	20 RIGHT												1							1
	21 LOST SIGHT OF													2						2
	CATEGORY TOTALS	0	6	9	0	2	2	1	1	1	0	0	1	3	0	0	0	0	0	26
CIRCLE	22 ONE CIRCLE FIGHT			1																1
	23 TWO CIRCLE FIGHT																			
	CATEGORY TOTALS	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
ALTITUDE	24 ADVANTAGE											2								2
	25 DISADVANTAGE						1	1					1							3
	26 DIFFERENT																			
	27 SAME									1										1
	CATEGORY TOTALS	0	0	0	0	0	1	1	0	1	0	2	1	0	0	0	0	0	0	6
SWITCHES	28 IN GUNS		6	1		1												2		10
	29 IN HEAT		5	5	2															12
	30 WRONG		1	1		1									1					4
	31 LATE			1																1
	CATEGORY TOTALS	0	12	8	2	2	0	0	0	0	0	0	0	0	1	0	0	2	0	27
RADAR	32 LOCK ON		5	1		1	1													8
	33 NO LOCK																			0
	34 CLOSE TO LOCK																			0
	CATEGORY TOTALS	0	5	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	8
NOSE POSITION	35 NOSE ON		7	8	2			3							2		2	1		25
	36 NOSE OFF			1		1														2
	37 NOSE IN LEAD			1											1					2
	38 NOSE IN LAG			1	1	2				1										5
	39 NOSE WITHIN 60								1							1				2
	40 HEAD-ON																		1	1
	41 NOSE LEVEL											1								1
	CATEGORY TOTALS	0	7	11	3	3	0	3	1	1	0	0	1	0	3	1	2	1	1	38

Table D1. (Concluded)

CATEGORY	ITEM NAME	STATE 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	TOTALS
ENERGY	42 GAINING		1	1	1	1						1			2		1			8
	43 MAINTAINING		9	1	2	1		1				1	1	1	2		5			24
	44 LOSING											1	1	1	1			1	1	6
	CATEGORY TOTALS	0	10	2	3	2	0	1	0	0	0	3	2	2	5	0	6	1	1	45
Gs	45 HIGH			1	1		1	3	1			1								8
	46 LOW			1								1		1	2	1	3	1	1	11
	47 SAME			3																3
	48 NOT USING		1									1	1	2	4		5	5	2	21
	49 TOO SOON																			0
	50 GOD'S G's								1											2
	CATEGORY TOTALS	0	1	5	1	0	1	3	3	0	0	3	1	3	6	1	8	6	3	45
SHOTS	51 GUN		5	1		1													1	8
	52 HEAT		1	1	1										2		2	1	1	9
	53 HI ANGLE GUN																			
	54 MISSED			1																1
	55 LATE		2	1																3
	CATEGORY TOTALS	0	8	4	1	1	0	0	0	0	0	0	0	0	2	0	2	1	2	21
3-9 POSITION	56 ADVANTAGE		8	9	3	2	1	1	3	1				1	2		2	1		34
	57 LOST ADVANTAGE							1	1			4	7	2	2			1		18
	CATEGORY TOTALS	0	8	9	3	2	1	2	4	1	0	4	7	3	4	0	2	2	0	52
LIFT VECTOR	58 CONTROL-GOOD		1		3			1	3			1	1	1	5		3			19
	59 CONTROL-POOR			1		1				1			2	3	2	2	8	5	4	29
	60 ABOVE HORIZON											1	1		1		1			4
	61 BELOW HORIZON									1										1
	62 ON OPPONENT																1			1
	63 CLOSE																	1		1
	CATEGORY TOTALS	0	1	1	3	1	0	1	3	2	0	2	4	4	8	2	13	6	4	55
AOA	64 HIGH						1	1		1										3
	65 LOW			3														1		4
	66 GOOD																			
	67 SAME																			
	CATEGORY TOTALS	0	0	3	0	0	1	1	0	1	0	0	0	0	0	0	0	1	0	7
Vc	68 HIGH			2					2						1					5
	69 LOW			2																2
	CATEGORY TOTALS	0	0	4	0	0	0	0	2	0	0	0	0	0	1	0	0	0	0	7
ASPECT	70 LESS THAN 60		2	11	2	1	1	3		1		1		1		1	2	1		27
	71 60o to 90			1						1			1		2					5
	72 GREATER THAN 90												1		1					2
	CATEGORY TOTALS	0	2	12	2	1	1	3	0	2	0	1	2	1	3	1	2	1	0	34
OTHER	73 OTHER	0	3	6	1	1	1	0	3	1	0	2	6	7	7	5	5	6	5	59

Table D2. Frequency count by Item and Category: Experiment Two, Engagement Two.

CATEGORY	ITEM NAME	STATE 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	TOTALS
AIRSPEED	1 < 200				1				1				1							3
	2 > 200																			
	3 EXCESSIVE																			
	4 ADVANTAGE	1	1	2	2	5	3	1	1	1	2		1							20
	5 DISADVANTAGE				2	2	3	4	5	4	2	1	7							30
	6 SAME			1		3	4	3	4											15
	7 ACCELERATING		1		1					1										3
	CATEGORY TOTALS	1	2	3	6	10	10	8	11	6	4	1	9	0	0	0	0	0	0	71
RANGE	8 < 1000 Ft	3		1	4	2					1						1			12
	9 1000 - 3500 Ft		2	7		1	5		1											16
	10 > 3500Ft																			
	11 IN PARAMETERS	1	1	8	1		2													13
	12 OUT OF PARAMETERS				1		1	1			1									4
	13 NEAR PARAMETERS	1		3	7	3	5	3												22
	14 IN-RANGE GUN			8	4		1										1			14
	15 IN-RANGE HEAT						1													1
	CATEGORY TOTALS	5	3	27	17	6	15	4	1	0	2	0	0	0	0	0	2	0	0	82
BOGIE POSITION	16 IN WINDSCREEN			6	3		3	3			1									16
	17 IN HUD FOV		1	4	1		2				1									9
	18 NEAR CANOPY BOW				5	1	1	4	3											14
	19 LEFT					1	1													2
	20 RIGHT									1										1
	21 LOST SIGHT OF																			
	CATEGORY TOTALS	0	1	10	9	2	7	7	3	1	2	0	0	0	0	0	0	0	0	42
CIRCLE	22 ONE CIRCLE FIGHT							1												1
	23 TWO CIRCLE FIGHT									1										1
	CATEGORY TOTALS	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	2
ALTITUDE	24 ADVANTAGE		1		2	2	3	3	1	2	2	1								17
	25 DISADVANTAGE							1	2		1		1							5
	26 DIFFERENT						1													1
	27 SAME			1				1	1	1			1							5
	CATEGORY TOTALS	0	1	1	2	2	4	5	4	3	3	1	2	0	0	0	0	0	0	28
SWITCHES	28 IN GUNS	2	2	7		3														14
	29 IN HEAT				1	1	1													3
	30 WRONG			1	1	1	1				1									5
	31 LATE						1													1
	CATEGORY TOTALS	2	2	8	2	5	3	0	0	0	1	0	0	0	0	0	0	0	0	23
RADAR	32 LOCK ON										1									1
	33 NO LOCK																			
	34 CLOSE TO LOCK																			
	CATEGORY TOTALS	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
NOSE POSITION	35 NOSE ON	2	2	9	4	2	7	6	5											37
	36 NOSE OFF				1	3	1	1												6
	37 NOSE IN LEAD				2	2			2											6
	38 NOSE IN LAG	1				1	5													7
	39 NOSE WITHIN 60				1		3	1	2											7
	40 HEAD-ON																			
	41 NOSE LEVEL																			
	CATEGORY TOTALS	3	2	9	8	8	16	8	9	0	0	0	0	0	0	0	0	0	0	63

Table D2. (Concluded)

CATEGORY	ITEM NAME	STATE 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	TOTALS
ENERGY	42 GAINING				2		2		1											5
	43 MAINTAINING		1		3		1	6	2	1										14
	44 LOSING				3		2	4	3	1	2	1	3							19
	CATEGORY TOTALS	0	1	0	8	0	5	10	6	2	2	1	3	0	0	0	0	0	0	38
Gs	45 HIGH				1			2												3
	46 LOW																			
	47 SAME							1		1										2
	48 NOT USING																			
	49 TOO SOON												1							1
	50 GOD'S G's									1										1
	CATEGORY TOTALS	0	0	0	1	0	0	3	0	2	0	0	1	0	0	0	0	0	0	7
SHOTS	51 GUN	2	1	1		1														5
	52 HEAT																			
	53 HI ANGLE GUN			1		1														2
	54 MISSED																			
	55 LATE	1				1														2
	CATEGORY TOTALS	3	1	2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	9
3-9 POSITION	56 ADVANTAGE	1	3	7	9	6	16	7	5		1	1								56
	57 LOST ADVANTAGE				1				2		1	3	2	1						10
	CATEGORY TOTALS	1	3	7	10	6	16	7	7	0	2	4	2	1	0	0	0	0	0	66
LIFT VECTOR	58 CONTROL-GOOD	1	1	1	4	5	10	11	5	7	1	1	1							48
	59 CONTROL-POOR		1	1				1	1		2	1					1			8
	60 ABOVE HORIZON				2					1			1							4
	61 BELOW HORIZON							3	2	3										8
	62 ON OPPONENT			2			1	1	1	1	2									8
	63 CLOSE																			
	CATEGORY TOTALS	1	2	4	6	5	11	16	9	12	5	2	2	0	0	0	1	0	0	76
AOA	64 HIGH												2							2
	65 LOW																			
	66 GOOD																			
	67 SAME																			
	CATEGORY TOTALS	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	2
Vc	68 HIGH		1	1		1					1	1								5
	69 LOW																			
	CATEGORY TOTALS	0	1	1	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	5
ASPECT	70 LESS THAN 60	2		1		3			1											7
	71 60o to 90			1		1	1	1	2	1	1									8
	72 GREATER THAN 90					2		2	2				1							7
	CATEGORY TOTALS	2	0	2	0	6	1	3	5	1	1	0	1	0	0	0	0	0	0	22
OTHER	73 OTHER	1	3	2	3	2	4	6	2	0	3	0	3	1	0	0	1	0	0	31