Proceedings of the 2005 Winter Simulation Conference
M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, eds.

THE PAIRWISE ESCAPE-G METRIC: A MEASURE FOR AIR COMBAT MANEUVERING PERFORMANCE

Antoinette M. Portrey

Lockheed Martin Systems Management
Air Force Research Laboratory
6030 South Kent Street
Mesa, AZ 85212, U.S.A.

Brian Schreiber

S&D Statistical Consulting Services
Air Force Research Laboratory
6030 South Kent Street
Mesa, AZ 85212, U.S.A.

Winston Bennett, Jr.

Air Force Research Laboratory
Warfighter Readiness Research Division
6030 South Kent Street
Mesa, AZ 85212, U.S.A.

ABSTRACT

The Air Force Research Laboratory, Warfighter Readiness Research Division, is continuously researching tools to measure performance of knowledge and skills from an individual level to the Command and Control (C^2) level, within both high fidelity distributed simulation environments and live training environments. Using the Performance Effectiveness Tracking System (PETS), we ran preliminary testing of a metric called Pairwise Escape-G that uses a concept called the Theoretical Instantaneous Probability of Weapon Intercept (TIPWI). TIPWI takes into account the current geometry of one aircraft against another for each given weapon (i.e., the physics-based envelope parameters) and is the weapon's probability of threat intercept at any instant during an engagement. This paper will describe the initial application of the Escape G metric within the Distributed Mission Operations Testbed (four high-fidelity F-16 simulators, one Airborne Warning and Controller System console, and Instructor Operator Station), preliminary outcomes, and suggested applications for this metric.

1 INTRODUCTION

"The primary mission of fighters is air superiority; that is, ensuring use by friendly aircraft of the airspace and denying use of that airspace to the enemy," (Shaw 1985). Since aviation became a major military component during WWI, fighter pilots have been trained to seek and achieve a positional / energy advantage over their adversary in order to destroy them. These air combat tactics involve dynamic, four-dimensional (x, y, z, and time), maneuvering at, or surpassing, supersonic speeds. There are a limited number of maneuvers available to the pilot in a given situation, dictated more by the weapon, relative positions and energy states of himself and his opponent than by potential technical advantages available in today's aircraft (Gunston and Spick 1983).

The relationship between aircraft in combat is dynamic as each pilot maneuvers to counteract the maneuvers of the other. Thus, any valid performance measurement system must take into account the maneuvering of both proponent and opponent aircraft in order to obtain an accurate indication of pilot performance (Wooldridge et al. 1982). As a four-dimensional challenge of both space and time, maneuvering a high-performance aircraft into an optimum position / energy profile is an extremely difficult skill to teach, learn, and assess. In the past, pilot performance was evaluated through dual flights with instructor pilots using performance record sheets, or instrument procedures through trainers such as the Link Trainer.

The measurement of this performance requires sensitive, continuous measures that accurately assess intercept geometry, the weapon engagement zone (WEZ), and energy management between fighter aircraft. Intercept geometry is the range and angular limits (e.g., aspect angle) between aircraft that are used to determine weapon launch boundaries. The WEZ is a hypothetical area surrounding an aircraft called the weapons envelope in which an adversary is vulnerable to a shot. The actual minimum and maximum WEZ ranges, relative to the enemy position are based on many factors, such as type of weapon, aircraft speed, relative altitudes, and geometry. Energy management involves a combination of kinetic and potential energy of the aircraft. As an aircraft maneuvers in combat, the aerodynamic design, speed, and thrust of the aircraft dictates the amount of potential and kinetic energy available to the pilot. Pruitt (1979) has shown that pilots who can exploit an aircraft's energy maneuverability more effectively achieve greater success in air combat. Shaw (1985) states that the purpose of the energy fight is to gain an energy advantage over an opponent without yielding a decisive position advantage; thus allowing the pilot to convert this energy into a lethal position advantage.

Developing performance measures for air combat tasks is challenging because these tasks are complex, dynamic, and transpire rapidly, thus the amount of information that is available is extensive (Dixon 1990). Historically, simple outcome measures such as kill ratios and missile-hit ratios were used to measure pilot and team performance. However, these outcome measures only provided end results and did little to reveal how well a pilot or team is performing during a sortie, or to reveal associations between process measurements and important sortie events. Since the early 1960's, research on performance measurement has been an important focus for the U.S. military. Brecke and Miller (1991) identified three approaches of research that have been continuously investigated over the past four decades.

An approach by Brictson et al. (1978, cited in Dixon 1990) focused on Discrete Event Measures (alternately known as Non-Maneuvering Aircrew Assessment). This approach used data collected at specific discrete points in a simulated air combat sortie to determine a measure of overall performance. This system was limited to a small number of discrete events occurring prior to a close-in maneuvering phase of the sortie (Brecke and Miller 1991). Alternately, the Energy Maneuverability research by Pruitt (1973, cited in Dixon 1990) was oriented towards presenting pilots with graphic displays to measure a pilot's ability to make use of the kinetic and potential energy loss/ gain characteristics of his/ her aircraft. Pilots traded altitude for speed or kinetic energy. Likewise, the pilot could convert aircraft speed back into altitude or potential energy.

The Positional Advantage research initiated by Oberle (1974) and followed by McGuiness, Forbes, and Rhoads (1985) produced the All Aspect Maneuvering Index (AAMI) – a measure that estimates how well pilots maneuver within the WEZ; outputting kill/no kill ratios, probabilities, and indices reflecting the pilot's firing capabilities (Brecke and Miller 1991, Schreiber and Bennett 2005). The AAMI is a composite index developed by Vreuls Research Corporation and Logicon (1987) and was intended as an estimate measure of the Theoretical Instantaneous Probability of Weapons Intercept (TIPWI) (Schreiber and Bennett 2005, Portrey 2005).

1.1 Theoretical Instantaneous Probability of Weapons Intercept

TIPWI is a construct coined by Brian Schreiber (Schreiber and Bennett 2005) that represents the depth of WEZ penetration, i.e., the probability that a weapon will intercept its target if fired at that moment. The TIPWI is a function of opposing aircraft ranges, weapon type, altitude, aspect, and airspeed, taking into account the current intercept geometry of one fighter aircraft against one threat aircraft for a given weapon (i.e., the physics-based envelope parameters). Consistently maintaining a TIPWI advantage directly contributes to the theoretical probability of kill (Pk). Up to and including the moment of weapon launch TIPWI is a continuous estimate of Pk. After deployment of the weapon, TIPWI no longer applies to the weapon deployed, but rather to the weapons remaining on the aircraft. Basically, TIPWI combines the aircraft's maneuvering capabilities and the weapon's capabilities to form a sensitive estimate of Pk.

The instantaneous geometry between two aircraft defines their current situation, and includes, among other factors, velocity vectors, X, Y, Z positions, relative headings, and altitude. Given this instantaneous inter-aircraft geometry, the probability of intercept is then determined by the weapons onboard. TIPWI calculations from the fighter to a threat are considered offensive estimates, while all the same calculations from a threat are considered defensive estimates of TIPWI. Current developmental research at the AFRL/Mesa is seeking to provide a better estimate of TIPWI by adding to and refining the original ideas from the AAMI (Schreiber and Bennett 2005, Portrey 2005).

1.2 Pairwise Escape-G Metric

For a pilot to achieve an offensive advantage he has to learn to control the interplay between energy management and positional maneuvers (Breidenback, Ciavarelli, and Sievers 1985). This study will analyze and discuss a performance metric called Pairwise Escape-G, which was derived as an estimate of the TIPWI using measures collected via a performance measurement system called the Performance Effectiveness Tracking System (PETS), developed by Schreiber et. al (2003). PETS collects measures that encompass all indices used in all three approaches: positional, energy, and discrete event metrics. Some of the positional and energy measures calculated in the Pairwise Escape-G metric are shown in Table 1.

Table 1: Pairwise Escape-G Metric Components

Metric	Measure
Airspeed	Energy
Lat/ Lon/ Alt	Positional
Velocity	Energy
Acceleration	Energy
Relative Range	Positional
Orientation	Positional
Angular rates	Positional
Air density ratio	Energy
Mach value	Energy
Turn rate	Energy

Note: All **bolded** measures were used in previous studies

The Pairwise Escape-G is a calculation of both positional and energy metrics using the same algorithms the aircraft uses to display the Dynamic Launch Zone (DLZ), a current display used by pilots to judge WEZ penetration of a threat. However, the calculation of Escape-G is more in-depth and continuous, involves weapons fly-out models that use instantaneous relative aspect angels, ranges, attitudes, closure velocity, and type of missile – all to determine the precise degree of WEZ penetration by the adversary. The metric is called the Escape-G value because it estimates the amount of G-force a pilot must pull (to either 0 or 180 aspect, whichever is more appropriate) in order to defeat the weapon fired. In essence, the metric indicates how quickly the pilot must turn in order to survive. The metric is also defined as Pairwise because calculations are made for a pilot's "defensiveness" and "offensiveness" against other aircraft thus creating a matrix of Escape-G values. The focus of this study is to investigate offensive Escape-G values and their relationship to outcome events.

Pairwise Escape-G values can be utilized as a measure of how well a pilot is managing his WEZ during the engagement. The WEZ is a relatively simple way for a pilot to think about how far a weapon has to travel to a target. Figure 1 is a graphical representation of friendly and threat WEZ.

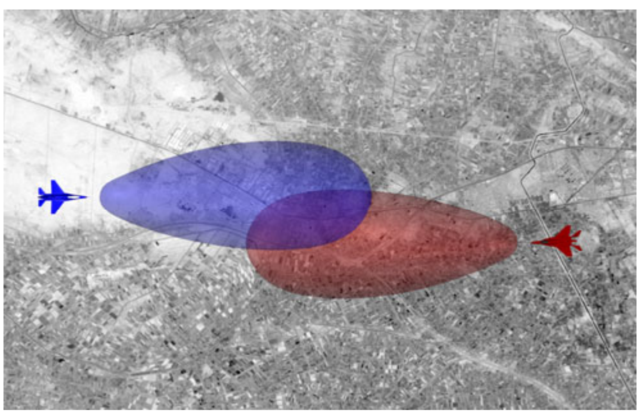

Figure 1: WEZ of an F-16 and a Threat Aircraft

Escape-G values can assist in the assessment of whether or not a targeted threat aircraft is within a vulnerability zone in order for the F-16 pilot to engage. The WEZ, being a mental model, does not account for what type of threat aircraft is the target, nor does it account for any maneuvering the threat aircraft may do to defeat a missile shot. It is purely based on the capabilities of the weapon the F-16 chooses to employ. The WEZ introduces an idea of the weapons intercept, while Escape-G is a dynamic calculation of the weapons interception probability.

Putatively, a pilot who is successful in managing WEZ will have the threat aircraft in positions requiring high Escape-G maneuvers to survive. Similarly, their adversaries desire favorable WEZ geometries and attempts to manipulate the inter-aircraft geometries such that the friendly pilots are in states requiring high G maneuvers. By choosing the direction and rate of turn, the target can exert tremendous influence on the WEZ. Whichever aircraft maintains the geometric advantage and a high Escape-G value will have the advantage over the adversary. As Escape-G values depend on weapon type, different weapons loads may suggest different tactics, so as to favorably impact WEZ penetration (i.e., Escape-G) and the depth of penetration required deciding to shoot—an intriguing question specifically explored in the current work.

Most current tactics emphasize using medium-range missiles that are deployed beyond-visual-range (BVR). A technologically complex series of events must occur for a missile to detect, track, and fire on a target, assuming a launch in a proper weapons envelope. The primary goal of this current study is to confirm the utility and predictability of the Escape-G metric in its sensitivity to shot opportunities.

- Hypothesis 1 – The offensive Escape-G value should be significantly higher around offensive events - shots and shots resulting in a kill – than during periods of no-shot events.
- Hypothesis 2 – The weapons load of the offensive aircraft will significantly impact the relationship between the Escape-G values and shots. Different weapons loads call for different combat tactics which, in turn, dictates firing doctrine. Standard weapons load ((4) medium and (2) short ranged missiles) will have higher Escape-G values than a medium-ranged missile load.

2 METHODS

2.1 Participants

Four Air Force F-16 pilots, all instructor pilot qualified, were selected from the Air Force Research Laboratory at Mesa, Arizona. Experienced pilots were selected for this initial investigation of Pairwise Escape-G metrics because pilots experienced in the aircraft should perform with higher skill and exhibit the more realistic flying behavior, thus providing the most realistic and valid data. This group of pilots has a $M = 19.75$ years of service, a $M= 1794$ hours F-16 flight time, $M= 4149$ total fighter hours, and a $M= 161.25$ hours in the AFRL/Mesa simulators.

2.2 Apparatus

Simulation System. The Pairwise Escape-G scenarios were run on the high fidelity simulation system at AFRL Mesa Research Site. This facility conducts research on technologies and training applications in a Distributed Mission Operations (DMO) test bed comprised of four high-fidelity networked F-16 Block 30 simulators. The basic components of a single-ship system within DMO include an F-16 multi-task trainer (MTT) cockpit, a rear-projection display system, image generator hardware and software, and a detailed terrain database.

The Mobile Modular Display for Advanced Research and Training (M2DART) is a rear-screen, real-image, display system that uses commercially-off-the-shelf cathode-ray (CRT) projectors to provide out-the-window visual imagery to the user with a full 360˚ field of regard. The M2DART has eight flat projection screens linked together to display eight channels of full color imagery. The F-16 MTT uses existing Air Force-owned operational flight trainer computer code and aircraft operational flight program software from the aircraft systems' line replacement units provided by the aircraft logistics depot. This software was converted to run at the same 50 Hz rate of the aircraft microprocessors. The F-16 MTT cockpit is functionally equivalent to its respective aircraft; it has full-fidelity instrumentation and controls. Another component of the simulation system is the Instructor/Operator Station (IOS), which is designed to control all operations of the four linked F-16 MTTs. The IOS controls the set-ups, initializations, and configurations of the linked MTTs.

Performance Effectiveness Tracking System. PETS is interfaced with the IOS and F-16 MTTs via ethernet connection. PETS is a software tool that enables multi-platform, multi-level measurement ability at the F-16 individual and team level in distributed environments. Relevant to the current work, PETS captured demographical, objective data such as outcome events, and Escape-G variables in both time-stamped, tab-delimited files and shot summary files (i.e., unit of analysis is every shot).

2.3 Trial Scenario

The initial set-up for each trial began with each aircraft at an altitude of 25k feet, at an airspeed of 350 knots, and each aircraft was paired off facing each other at a starting range of 40nm. The pilots were told that they had full fuel tanks with no external tanks, no Tactical Awareness Display (TAD) or Situational Awareness Data Link (SADL), and a specific weapons load (announced before each trial).

2.4 Data Collection Procedure

Prior to data collection, all pilots were briefed on the background and purpose of the study. The briefing included the objective of the study and a description of the task to be performed. After the initial briefing, the pilots were instructed to fill out a pilot demographic questionnaire. The pilots and the console operator were given a protocol sheet explaining the procedures during the trials.

Since all the pilots routinely work in the simulation environment and are familiar with the F-16 MTTs at AFRL/Mesa, no familiarization flight was necessary. Each pilot flew a total of thirty 1v1 trials in one day. Since all the pilots have the same qualifications and similar experience, they flew against each other in a balanced matrix. Random pairings of pilots were determined prior to the start of the study. Each pilot flew fifteen 1v1 trials with four medium-ranged, radar missiles and fifteen trials with a standard weapons load of four medium-ranged, radar missiles and two short-ranged, heat seeking missiles. Table 2 shows the trial setup given to the instructor operator, with (1) being the medium-ranged missile load only and (2) being the standard weapons load.

Table 2: Arrangement of Trials

Sess	**Pilots**	**T1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
A	**1v2** **3v4**	1	1	2	1	2	2	1	1	2	2
B	**1v3** **2v4**	2	1	1	2	2	2	1	2	1	1
C	**1v4** **2v3**	1	2	2	1	1	2	1	2	2	1

The trials ended when a pilot was killed, both aircraft were out of weapons, or three minutes had elapsed. Three minutes was decided upon to reflect prior research (see Brecke and Miller 1991, Dixon 1990), and to ensure enough data has been collected. If both pilots were killed simultaneously, then both shots were analyzed as individual kills.

Data recording for the trials was started automatically by an initializing global start from the IOS station. The data were recorded on CD. The saved trials were then played back through PETS and all the data from the trials was collected at a 5 Hz update rate. During this time the trial outcomes were coded for each trial.

2.5 Raw Data Collected

Two kinds of data were collected during the study. These were: a) pilot demographical data and b) means and selected scores from continuous measurement of the performance measures. The demographical data were collected by questionnaire before the subjects finished the initial briefing. The most comprehensive data were taken by the automatic recording of the various performance measures. This data was then calculated into raw means scores for further analysis.

2.6 Experimental Design

To capture the relationship between Escape-G values and missile shots taken, the trials were summarized by computing the mean Escape-G values. The data was partitioned according to two factors – Load (defined by the missile load) and Interval (defined by the size of the time segment near a missile shot.

3 DATA ANALYSIS

3.1 Time Interval Analysis

Due to the possibility of multiple shots occurring in a trial, a frequency distribution was used to select the optimal interval times to represents the likely interval during which a pilot recognizes a missile shot opportunity and takes it. Analysis of the distribution showed that a majority of the multiple shots were taken within ten seconds, the smallest pause between shots was .8 of a second. Since the data was collected every 200 milliseconds, we decided to segregate the time into 10 one second intervals – five seconds before and five seconds after a shot event. Interval had twelve levels beginning with level 1 being the baseline estimate (no-shot).

Level 1 was analyzed by taking the Escape-G mean of all the times not within the ten seconds surrounding a shot. Levels 2-6 are the pre-shot cluster intervals, five seconds before a shot. Levels 8-12 are the post-shot cluster intervals, five seconds after a shot. Level 7, the mid-interval, is the Escape-G mean at shot event. For engagements with multiple shots being fired within ten seconds of each other, by the same fighter, we calculated these situations as shot clusters. If the multiple shots had been coded individually, then some post-shot intervals of one shot would have overlapped with the pre-shot intervals of the next shot, thus creating a violation of independence. Mean Escape-G values were computed for each category defined by the different combinations of these factors.

The study was a repeated measures design with two factors, including 1)Weapons Load (all medium-range missiles vs. standard load), and 2) Interval (one through five seconds). The dependent variable of interest was mean Escape-G, with means computed by collapsing across trials and opponents. The data was analyzed using the GLM procedure of SPSS version 12.0.

4 RESULTS

The analysis was a 2x12 completely-crossed, within subjects design. The means and standard deviations found for each combination of conditions are displayed in Table 3.

Table 3: Descriptive Statistics

Load	Interval		Mean	Std. Dev.	N
	Level	Description			
4 Medium-Range Missiles	1	No-shot baseline	.07	.03	4
	2	Five secs before	.26	.10	4
	3	4 Secs	.29	.12	4
	4	3 Secs	.32	.14	4
	5	2 Secs	.35	.15	4
	6	1 Sec (Immedi-	.36	.15	4
	7	Midinterval	.35	.19	4
	8	1 Sec (Immedi-	.38	.12	4
	9	2 Secs	.36	.11	4
	10	3 Secs	.34	.10	4
	11	4 Secs	.33	.11	4
	12	5 Secs	.30	.11	4
Standard Weapons Load	1	No-shot baseline	.07	.04	4
	2	Five sec before	.28	.20	4
	3	4 Secs	.32	.19	4
	4	3 Secs	.36	.21	4
	5	5 Secs	.39	.23	4
	6	1 Sec (Immedi-	.40	.23	4
	7	Midinterval	.67	.56	4
	8	1 Sec (Immedi-	.63	.44	4
	9	2 Secs	.62	.43	4
	10	3 Secs	.51	.25	4
	11	4 Secs	.51	.26	4
	12	Five secs after	.47	.21	4

4.1 Interval

With respect to the hypothesis that Escape-G metric is significantly related to missile shot-taking behavior, the Interval factor showed a significant linear increase reflecting the pilot's intercept geometry management leading up to a shot; Escape-G values increase from zero when the pilot takes a shot, $F(1,11) = 13.84$, $p<.05$ (see Figure 2). Analysis showed that mean Escape-G is significantly higher in the shot intervals than in the no-shot intervals, $F(1, 3) = 11.952$, $p<.05$. Equally, when shots are separated into categories with Escape-G values at either zero or above zero, the conditional probability of a shot resulting in kill prediction is 2.58 times higher in the latter category [i.e., p(kill|Escape-G = 0) = .0203 versus p(kill|Escape-G>0) = .0523].

4.2 Weapons Load

The second hypothesis posited that the weapons load of the aircraft would affect pilot behavior in a manner that would have an affect on the relationship between Escape-G values and outcome of a shot. However, in this analysis, there was no significant difference between weapon loads or for the interaction involving Interval and Load.

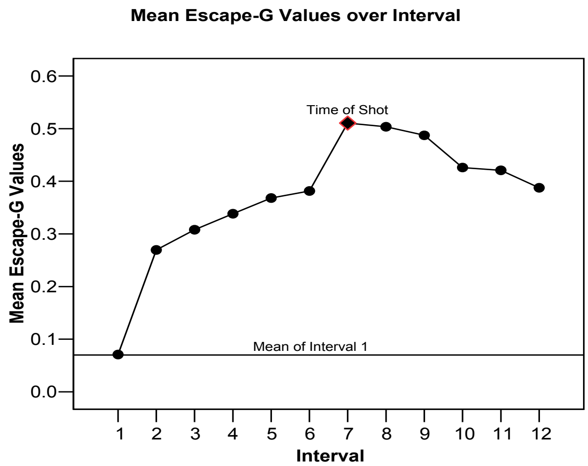

Figure 2: Mean Escape-G Values over Interval

5 DISCUSSION

The objective of this study was to examine the Escape-G metric and its relationship to important offensive events under various conditions of weaponry and interval. The main effect of significantly higher Escape-G values around shots than no-shot intervals clearly confirms that Escape-G is a sensitive metric related to offensive events such as shots. The data shows that there is an increased probability of a kill as Escape-G values increase from zero. The data also shows that, after a shot, Escape-G values slowly decrease as time progresses; this is an effect of medium-ranged weapon tactics. Tactics for medium-ranged weapons are designed for beyond-visual-range (BVR) fighting. Due to the distances in BVR fighting, pilots are not required to go straight into evasive maneuvers after firing a missile like in close air conflict, thus corresponding Escape-G values gradually decrease. Overall, this finding of linking a TIPWI-derived metric to important sortie events such as kills accomplishes two important research objectives. First, it validates the Escape-G metric as a sensitive measure of TIPWI. Second, the association with important sortie events provides justification for further exploring the modeling and predicting of military sortie events, not only for just offensive events during small force employments like that used here (i.e., 1 v 1), but also possibly for larger force employments in the context of offensive and/or defensive events.

The second hypothesis dealt with the correlation between weapons load (Load) and Escape-G values. While there was no significant difference between weapons load and the Escape-G values, this may be reflecting a bias in missile preference. In the standard weapons load condition the participants were given a weapons load consisting of four medium-range, radar missiles and two short-range, infrared missiles. Over all trials the participants never fired a short-range missile; all missiles fired were medium-range. This is a representation of the preference for BVR warfare (Houck, Whitaker, and Kendall 1993). Participants were risk-averse, therefore preferring to only take shots close to the limit of the WEZ. Infrared missiles would have placed them well within the WEZ. Another interesting result was that Escape-G values tended to be low overall, even during shot periods. We attribute this to the experienced participants being knowledgeable in weapons employment and easily able to recognize when they (and their adversary) were at the limits of the WEZ. This is further supported by the fact that there were few kills overall and that some shots were taken with a zero Escape-G value.

5.1 Potential Applications

The validation of this performance metric will provide information that could be utilized in many applications. For example, this metric could provide improved feedback to pilots concerning their performance of basic flight maneuvers and advanced combat maneuvers and associated tasks. Second, it could provide pilots with information about their intercept geometry and WEZ management. In an assessment capacity, this metric combined with other significant metrics could be used to determine the performance of a pilot at each stage of a sortie. A final use for this metric could be used in tactics assessment; new tactics could be tested and the results evaluated based on the progression of this metric's scores over time throughout the sortie. This would aid in determining which new maneuvers would provide the best results without the need for actual aircraft time (Dixon 1990).

5.2 Diagnostic and Assessment Display

In order to serve as a diagnostic tool, the metric must be able to indicate a student's strengths and weaknesses (Dixon 1990). This could be achieved by examining an individual's metrics against an opponent's and noting the differences over time via a graphical display. The design of this display would have to be investigated to ensure that the instructor pilots and evaluators understood all information and that the display followed human factors interface guidelines. Besides reflecting the Escape-G values, the display could include outcome events and other significant metrics such as minimum abort range (MAR) violations. All of the information used would be graphed against a timeline. A simple example of Escape-G values plotted against a timeline is shown below in Figure 3. This graph shows the Escape-G values for both fighters. Aircraft B is shown as firing when the Escape-G values are high, while Aircraft A is firing when there is no advantage.

This display would allow pilots to isolate individual differences from the sortie with focus on important outcome measures such as shots and kills. Once differences are found the sortie can be more closely scrutinized using the debrief system. This display, being graphed along a

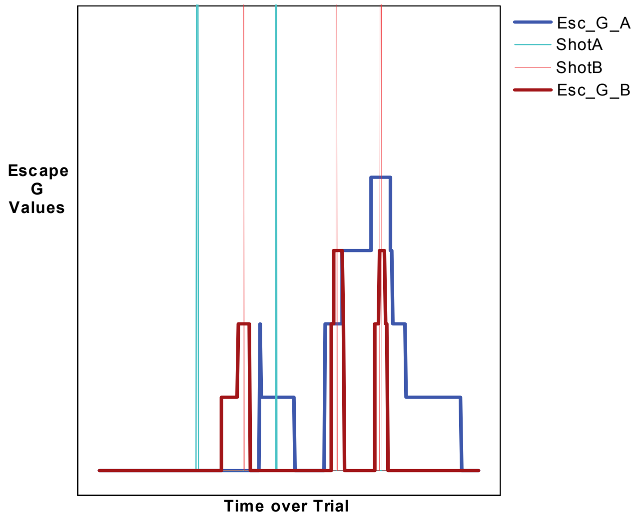

Figure 3: Diagnostic Display Example

timeline, could provide feedback to the pilot on actions that would provide better performance in combat maneuvering and positional advantage.

The value of using the above example for training applications is based on several requirements. One requirement is that the proposed screen designs would have to be studied. There are many factors that are involved in designing screens for a certain population and working environment (Proctor and Van Zandt 1994; Rogers, Sharp, and Preece 2002). Some of the more relevant factors are position of the information, complexity level, and intuitiveness. Studies in screen layout would have to be accomplished using proven human factors screen display design principles and input from the user population. Another requirement is that instructors would have to be trained to use all the metrics available along with the Escape-G measure, which deviates from their current non-linear metrics and display references of WEZ penetration.

5.3 Tactics Assessment

Another future research path would be the utility of the using Escape-G metric in the study of tactics assessment. Several air combat maneuvering (ACM) tactics could be developed using data generated by the measures and metrics. For instance, new maneuvers could be designed that maximize the offensive Escape-G measure against a threat over time and in specific areas and compare results against those generated by traditional maneuvers. In this manner simulated maneuvers could be developed without the cost of actual aircraft time and manpower. In addition, air combat tactics already in operational use, and those on the planning board, could be examined before actual implementation. A further application would be testing tactics against enemy aircraft simulations. The simulated sorties could provide invaluable information for ways to improve air combat performance in these areas.

5.4 Methodological Limitations

The current work contained several limitations. Only four highly experienced, relatively homogenous pilots were used, thereby limiting the range of possible pilot behaviors and responses. Only F-16s with two types of weapons were used, limiting generalization across platforms, mission types, and weapon type. Only US F-16 pilots were used and F-16s rarely fly or use medium-ranged missiles against each other, thereby exercising a restricted range of tactical options. There were Radar Warning Receiver (RWR) limitations in that the RWR would not track the other F-16. The trials were three minutes in duration; this was noted as too short a duration for a BVR sortie. Studies on dissimilar weapon loads or short-range radar missiles were not possible due to actual training operations, availability of subjects, and simulator scheduling problems.

5.5 Final Recommendations and Conclusions

This study represents the initial development and analysis of data collected during 1v1 simulated air combat sorties. The data collected and the developed metric offer effective avenues to enhance training of air combat maneuvering. Given that Pairwise Escape-G validly estimates TIPWI and has shown here to correspond with important offensive events, a number of research initiatives could be undertaken to further understand its generalization and predictive utility across weapons, mission types, important mission event types, and force employment size.

ACKNOWLEDGMENTS

This research was performed at the Air Force Research Laboratory, Warfighter Readiness Research Division in Mesa AZ, under Air Force contract #F41624-97-D-5000, Principal Investigator Dr. Winston Bennett, Jr.

REFERENCES

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

Brictson, C. A., A. P. Ciavarelli, K. W. Pettigrew, and P. A. Young. 1978. Performance assessment methods and criteria for the Air Combat Maneuvering Range (ACMR): Missile envelope recognition. Special Report No. 78-4 (Confidential). Pensacola, FL: Naval Aerospace Medical Research Laboratory.

Dixon, K. W. 1990. The development and validation of air combat maneuvering outcome measures. Unpublished master's thesis, Arizona State University, Tempe, Arizona.

Gunston, B. and M. Spick. 1983. *Modern air combat: The aircraft, tactics, and weapons employed in aerial warfare today*. New York: Crescent Books.

Houck, M. R., L. A. Whitaker, and R. R. Kendall. 1993. An information processing classification of beyond-visual-range air intercepts. (AL/HR-TR-1993-0061). Williams Air Force Base, AZ: Aircrew Training Research Division, Armstrong Laboratory.

McGuinness, J., J. M. Forbes, and J. E. Rhoads. 1984. Air combat maneuvering performance measurement system design. (AFHRL-TP-83-56) Williams Air Force Base, AZ: Operations Training Division, Armstrong Laboratory.

Oberle, R. A. 1974. An air combat maneuver conversion model. (CRC 274) Office of Naval Research.

Portrey, A. M. 2005. The Escape-G metric: A concise measure for air combat maneuvering performance. Unpublished master's thesis, Arizona State University, Tempe, Arizona.

Proctor, R. W. and T. Van Zandt. 1994. *Human factors in simple and complex systems.* Boston: Allyn and Bacon.

Pruitt V. R. 1973. Energy management display system for a tactical fighter (Confidential) (AAFDL-TR-73-38) Wright-Patterson Air Force Base, OH: USAF Flight Dynamics Laboratory.

Pruitt, V. R. 1979. Energy management training aid for the Navy's Air Combat Maneuvering Range (ACMR) (Contract N00123-78-C-1371). St. Louis, MO: McDonnell Aircraft Co.

Rogers, Y., H. Sharp, and J. Preece. 2002. *Interaction design: Beyond human-computer interaction.* New York: John Wiley & Sons, Inc.

Schreiber, B. T. and W. Bennett Jr. 2005. Distributed Mission Operations within-simulator training effectiveness baseline study. Volume II: Metric development and objectively quantifying the degree of learning. Manuscript in preparation.

Schreiber, B. T., E. Watz, W. Bennett Jr., and A. M. Portrey. 2003. Development of a Distributed Mission Training automated performance tracking system. In *Proceedings of the Behavioral Representations in Modeling and Simulation (BRIMS) Conference.* Scottsdale, AZ.

Shaw, R. L. 1985. *Fighter Combat: Tactics and maneuvering.* Annapolis, ML: United States Naval Institute.

Vreuls Research CORP. 1987. Air combat maneuvering performance measurement system for SAAC/ACMI, Volume II, Appendices 5 & 6. Wright-Patterson AFB, OH: Air Force Systems Command.

Wooldridge, L., R. W. Obermayer, W. H. Nelson, M. J. Kelly, D. Vreuls, and D. A. Norman. 1982. Air combat maneuvering performance measurement state space analysis (AFHRL-TR-82-15).

AUTHOR BIOGRAPHIES

ANTOINETTE PORTREY is a Research Engineer with Lockheed Martin at the Air Force Research Laboratory Human Effectiveness Directorate, Warfighter Training Research Division, located at Williams-Gateway Airport, Mesa AZ. She completed her M.S. in Applied Psychology - Human Factors from Arizona State University in 2005. She is the Team Lead for the Performance Effectiveness Tracking System development team. Her email address is antoinette.portrey@mesa.afmc.af.mil.

BRIAN T. SCHREIBER is a Senior Research Scientist with Lockheed Martin at the Air Force Research Laboratory Human Effectiveness Directorate, Warfighter Training Research Division, located at Williams-Gateway Airport, Mesa AZ. He completed his M.S. in Human Factors from the University of Illinois at Urbana-Champaign in 1995. He has been actively involved in military aviation research since 1994. His email address is brian.schreiber@mesa.afmc.af.mil.

WINSTON BENNETT, JR. Ph.D is a Senior Research Psychologist with the Air Force Research Laboratory Human Effectiveness Directorate, Warfighter Training Research Division, located at Williams-Gateway Airport, Mesa AZ. He is the team leader for training systems technology and performance assessment research and development. He received his Ph.D. in Industrial Organizational Psychology from Texas A&M University in 1995. His email address is winston.bennett@mesa.afmc.af.mil.