

“PERCEPTIONS” of EFFECTIVENESS

Barry D. Watts
Senior Analyst, Northrop Corporation

U.S.-FRG Bilateral Talks
Rosslyn, Virginia
1988

OVERVIEW

- COMPARISON of EXISTING “MOE” SYSTEMS
 - Motivations, Costs, Problems of MOEs
 - Force Comparisons Using Selected Systems
 - Strengths and Weaknesses of Current “MOEs”
- ENDURING LIMITATIONS of “MOEs”
 - “Intangibles”
 - The Nonlinearity of Combat Processes

“Measures of Effectiveness”

“The **choice** of a measure of effectiveness is far from a trivial problem. Here **intuition** is as likely to be **wrong as right**: one needs to think through very carefully and in detail the **probable and possible uses** to which a system will be put, the **advantages** which will derive from its employment, and the associated **costs**, and to express these in a quantitative form which is **consistent with** the objectives of system aggregates at ‘**higher**’ levels, and **interpretable** as measures for constituent subsystems at ‘**lower**’ levels.”

— Herbert K. Weiss, 1966

Motivations for “MOEs”

- DECISIONS on force structure, weapons procurement, operational planning, etc. UNAVOIDABLE
 - A major input to these decisions are ASSESSMENTS of MILITARY BALANCES
 - Balance assessments can be EXPLICIT & ANALYTIC or IMPLICIT & HUNCHES
- AGGREGATE PREWAR “BEAN COUNTS” INADEQUATE
 - France, May 1940
 - Bekka air-to-air results, 1982
- MOEs—A FIRST STEP TOWARD GETTING BEYOND “MERE” BEAN COUNTS
 - Depict Static TRENDS in Opposing Forces
 - INPUTS for War Gaming, Modeling, Simulations, etc.

Slide 5 8/3/88 "Numbers, Strategy, and the European Balance," *Interna. Security*, Spring 88, pp. 177, 175 & 182; Hines, "Soviet Front Operations in Europe — Planning for Encirclement," Apr 86, pp. 18-20; *AFJ*, November 87, p. 32.

COSTS of MOEs

JOHN MEARSHEIMER's 1988 ASSESSMENT of the BALANCE of CONVENTIONAL NATO and WP FORCES in central Europe:

- "Severe force-to-space ratio constraints" mean that WP ground forces "can only achieve temporary local superiorities in the vicinity of 1.6:1" (i.e., cannot concentrate more than 1 attacking "brigade" on 7 kms of front)
 - Assumes ADEs are "the best existing measure of combat power"
- Since 1.6:1 is well below the "widely accepted rule of thumb" that an ADE advantage of 3:1 or more is needed to break through, a WP conventional ATTACK WILL FAIL

HOWEVER:

- Concepts Analysis Agency disowned WEI/WUV years ago
- Actual Soviet norms are:
 - 3- or 4-to-1 WITHOUT "massing of fires"
 - Regularly **achieved by** National Training Center OPFOR
 - Regularly **practiced by WP forces** (e.g., 2 tank regiments attacking on a 3 km front in Druzhba 1986)

Theater Force Ratios

France, May 1940

Ground Force Categories

Allied/German WEI/WUV Ratios

Small Arms	1.1 to 1
Tanks	1.4 to 1
APCs	2.3 to 1
Armored Systems (Tanks + APCs)	1.5 to 1
Antitank Guns	1 to 2.5
Armor/Antiarmor (Tanks + APCs + Antitank Guns)	1 to 1.5
Artillery	1.6 to 1
Mortars	1 to 1.3
<u>Firesupport (Artillery + Mortars)</u>	<u>1.1 to 1</u>
Theater Ground Forces Total	1.03 to 1

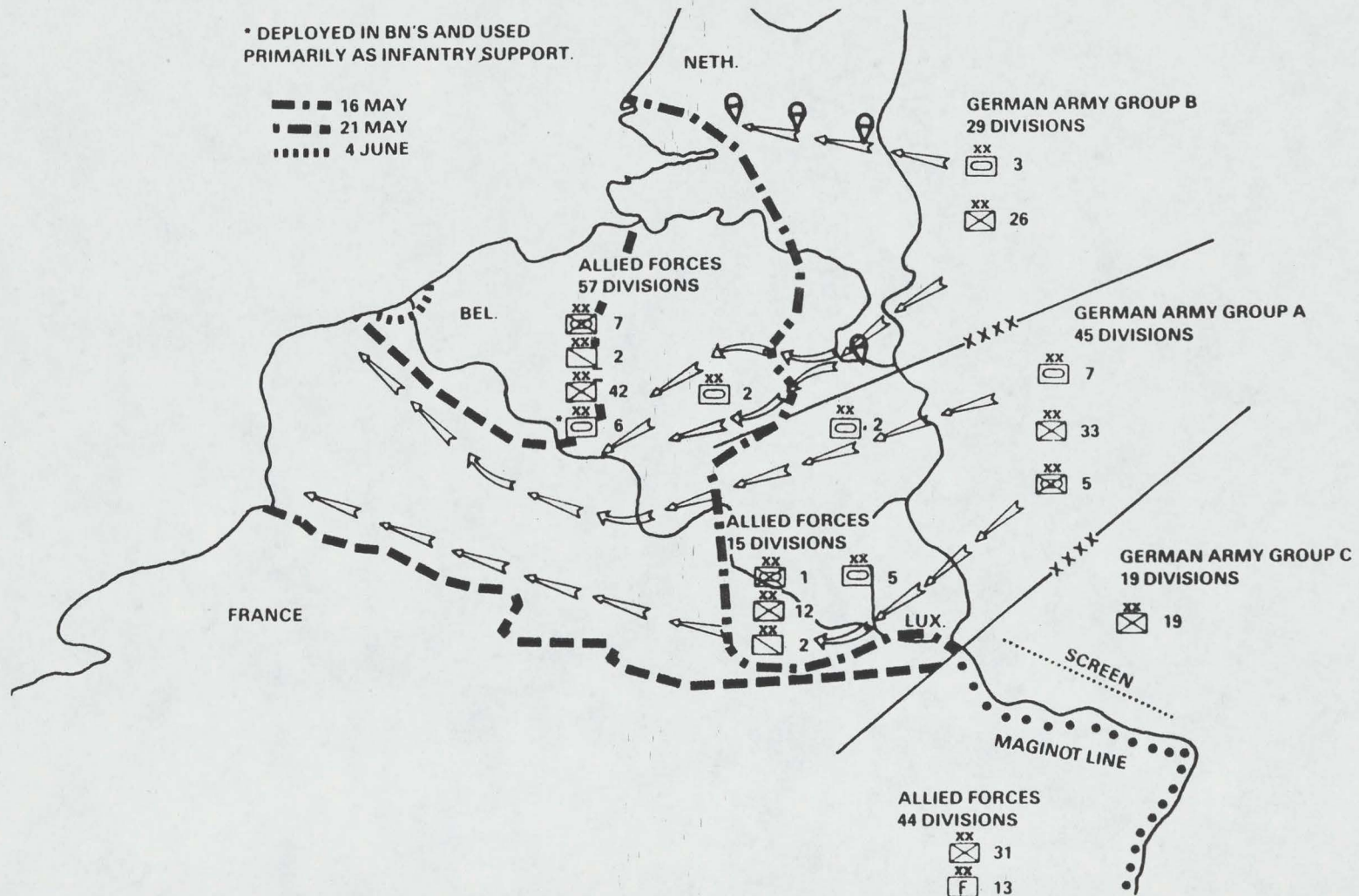
Air Force Categories

Allied/German Numerical Ratios

Aircraft	1 to 1.9
Antiaircraft Artillery	1 to 2.1

"Relatively evenly matched"
at the theater level

Campaign Results May 1940



Below-Theater (Army Group) Force Ratios France, May 1940

Northern Sector (Army Gp B—Belgium & Holland)

Allied/German Divisions (56 vs 29):	1.96 to 1
Allied/German WEI/WUV Ratio:	1.52 to 1

Central Sector (Army Gp A—Ardennes)

Allied/German Divisions (15 vs 54):	1 to 3.00
Allied/German WEI/WUV Ratio:	1 to 4.20

Southern Sector (Army Gp C—Maginot Line)

Allied/German Divisions (44 vs 19):	2.32 to 1
Allied/German WEI/WUV Ratio:	1.69 to 1

Reflections on May 1940

WEI/WUV FORCE ASSESSMENT:

- Dealt with INPUTS only
- Looked at one precombat "snapshot" in time

BDM/Project-186 (Phil Karber, et al.) CONCLUSIONS:

"WEI/WUV applied at the theater level of aggregation provided little insight into the conduct of operations or the outcome of battle."

At the army-group level, however, it yielded "insights into the dynamics of the campaign" and "possible outcomes."

QUESTIONS:

- Do WEI/WUV ratios *really* say more than division ratios?
- How can you predict the right level of aggregation?
- Is the outcome of theater combat *really* "the cumulative result of numerous smaller battles" along the front?

OVERVIEW

- COMPARISON of EXISTING “MOE” SYSTEMS
 - Motivations, Costs, Problems of MOEs
 - *Force Comparisons Using Selected Systems*
 - Strengths and Weaknesses of Current “MOEs”
- ENDURING LIMITATIONS of “MOEs”
 - “Intangibles”
 - The Nonlinearity of Combat Processes

MOTIVATIONS for Comparing "MOEs"

- After DECADES of ONGOING USE, NO CONSENSUS ON:
 - Adequacy and Acceptability
 - Range of Applicability
 - What MOEs *really* Measure
- Yet MOEs ADDRESS REAL ANALYTIC NEEDS
 - Theater-level **gaming, modeling, simulations** (MTM, RSAS, etc.) **need MOEs** of some sort
 - **MOEs permit summarization** of the aggregate firepower, modernity, combat-potential of disparate weapons/forces
- By 1985, ENOUGH NON-NUCLEAR MOEs with ENOUGH OVERLAP TO ATTEMPT COMPARISONS
 - Limitations of nuclear-force aggregate measures and dynamic models evident since late-1970s (Foster & Bennett, RAND/WN-9976-1-NA, Oct 77)

OSD/NA-CIA/ASG "MOE" Project

1985-86 Effort

- PA&E Contributed Blue Data, OSD/NA and CIA/ASG money
- Contractor: BDM

Systems Evaluated:

- CAA's **WEI/WUV** (Weapon Effectiveness Indices/Weighted Unit Values)
 - WEI II and WEI III
 - **GWEI** (Generic Weapons Effectiveness Indices)
- **TASCFORM** (The Analytic Sciences Corporation's Technique for Assessing Comparative Force Modernization)
 - **AWSP** (Adjusted Weapon System Performance)
- Data Memory Systems' **QJM** (Quantified Judgment Model)
 - **OLI** (Operational Lethality Index)

Used Actual NATO-WP Force Data (1975-1985)

- NATO: BOAR, FRG III Corps, US V Corps
- Warsaw Pact: 3rd Shock Army in GSFG

MOE Comparison Caveats

EACH "MOE" SYSTEM HAD VERY DIFFERENT ORIGINS
and AIMS

- WEI/WUV
 - Weapon and Force Effectiveness (Firepower + Mobility + Survivability) for NATO/WP ground forces
- GWEI
 - Political Viability in NATO (M-1 = Leopard II = Chieftain) based on WEI II
- TASCFORM
 - Force Modernization (*not* gaming, predicting outcomes, etc.) starting with Tacair
- OLI
 - Land Weapons Lethality derived from Historical Data

Comparisons are, strictly speaking, "Apples and Oranges"

- But all four systems focus on air and land forces

OVERVIEW

- COMPARISON of EXISTING “MOE” SYSTEMS
 - Motivations, Costs, Problems of MOEs
 - Force Comparisons Using Selected Systems
 - *Strengths and Weaknesses of Current “MOEs”*
- ENDURING LIMITATIONS of “MOEs”
 - “Intangibles”
 - The Nonlinearity of Combat Processes

Areas of MOE Comparison

- APPLICABILITY
 - What range of conventional forces does the system cover?
- DATA AVAILABILITY
 - How much information is needed to score new weapons?
- TRANSPARENCY
 - How simple is a new weapon to score? How intuitively correct is the result?
- SENSITIVITY and BIASES
 - What drives outputs for units/forces?
- INCLUSION OF OPERATIONAL and ENVIRONMENTAL FACTORS
 - What's left out?

Applicability

QJM and TASCFORM were highest in applicability to air and ground forces

- QJM covers all ground-based weapons, but no air-to-air
- TASCFORM omits small arms

WEI systems omitted fixed-wing air, surface-to-surface missiles, and, to different degrees, air defense

- WEI II: No air, no SSMs, and some air defense (divisional)
- WEI III: No air, no SSMs, some air defense (slightly more than WEI II)
- GWEI: No air, no SSMs, no air defense

No coverage of naval forces by any MOE systems examined

Data Availability

QJM's OLI scores require the least data (fewest technical characteristics) to score weapons

GWEI is being updated by PA&E and so has scores for most current weapons (assuming WEI II)

WEI II, WEI III, and TASCFORM have rather substantial requirements for data on technical weapons characteristics in order to compute scores

- Hard to get data for threat weapons as well as for newer allied systems

Transparency

QJM is the most transparent of the MOE systems

- Only 13 factors are needed to score a tank
- Multiplicative structure means a 10% change in a factor changes the score 10%

TASCFORM and the WEI systems are opaque

- TASCFORM only needs 24 values for a tank score, but the calculations are complex
- Calculation WEIs is straightforward, but 65 factors are need for a tank score

Sensitivity & Biases: QJM & TASCFORM

QJM:

- OLI's are essentially the product of lethality and rate of fire
- Heavily favors large combat vehicles, large basic loads, and full-track suspension
- Heavily favors weapon calibers of 130-155 mm
- Does reflect the ability to kill soft targets (expressed as casualties for men in the open per hour)

TASCFORM:

- Tacair origins lead to "tank payload" and "tons on target" outlook
- Conversion of all weapons to F-4 equivalent fails to account for differences in targets or employment
- Uniquely sensitive to obsolescence (which complicates trend-line analyses and data-base requirements)
- System survivability represented by a single coefficient that affects scores by $\pm 20\%$ (± 40 for armor)
- Treatment of SAMs (normalized to HAWK) and SSMs (as one-way, non-maneuvering aircraft) inconsistent

Sensitivity & Biases WEI/WUV

WEI/WUV:

- Assumes mid-intensity combat but has regional factors (Europe, NEA, Mideast)
- Raw scores limited to 1.00 maximum, which may tend to undervalue new or projected systems
- Attack helos valued by ability to kill armored vehicles
- Cannot generate new scores same way as original scores (i.e., collective judgments of groups of experts/analysts in 1974 & 1978)
- Antiarmor system scores tied to a specific tank threat

GWEI:

- Suffers same sensitivities & biases as WEI
- But less sensitive to minor differences between same-generation weapons

Inclusion of Operational & Environmental Factors

	<u>OPERATIONAL</u>	<u>ENVIRONMENTAL</u>
QJM:	55 Including mobility, surprise, air superiority effects, logistics, intelligence, and momentum	77 Including 17 terrain types, 12 weather states, and 12 seasonal variations
TASCFORM	1 TASCFORM simply does not consider operational and environmental factors (EFP undeveloped)	1
WEI II/III	2 Offense/Defense	3 Eur, NEA, Mideast

Conclusions

- “MOEs” CAN “SUPPORT” ANALYSIS (that is, provide INSIGHTS)
 - IF you understand purposes, limitations, biases of system(s) used
- “MOEs” Do NOT Measure EFFECTIVENESS
 - Cannot determine/predict outcomes
 - WEI, OLI, AWSP/AASP had other purposes
- “MOEs” Studied Are Basically FIREPOWER Scores
 - Focus on firepower, mobility, survivability of major weapons
 - Emphasize tanks & artillery; devalue infantry
 - Operational/environmental factors ignored or not handled convincingly
- “MOEs” ARE U.S. PERCEPTIONS OF MILITARY VALUE
 - Of various weapons, mostly in Europe
- All “MOEs” Studied HAVE TROUBLE “BRIDGING”
 - BETWEEN weapons categories (tanks vs artillery; air-to-air vs air-to-ground; etc.)
 - ACROSS service boundaries (air vs land)

OVERVIEW

- COMPARISON of EXISTING “MOE” SYSTEMS
 - Motivations, Costs, Problems of MOEs
 - Force Comparisons Using Selected Systems
 - Strengths and Weaknesses of Current “MOEs”
- ENDURING LIMITATIONS of “MOEs”
 - *“Intangibles”*
 - The Nonlinearity of Combat Processes

“INTANGIBLES”

Terrain, Weather, Tides, Time and Timing

Morale, Unit Cohesion, Training, Proficiency, Combat Experience, “Men in the Loop”

“Fog and Friction” of War

- Chance, Uncertainty, Danger, Exertion

Surprise, Maneuver, Leadership, Command and Control, Logistics, Organizational Factors

Doctrine, Missions, Operational Style

Plans, Strategy, Operational Art, Tactics

Political Aims versus Military Means

Human Variability, Individual Skill, Situation Awareness

Aircrew "FIGURES of MERIT" TASC, 1981

An ATTEMPT to QUANTIFY "INTANGIBLES" in air warfare

BASED ON:

- Cultural Factors
- Flight and Combat Training
- Proficiency Training

FACTORS:

Israel:	1.3
U.S., FRG	0.9
UK	0.8
USSR	0.6
Egypt	0.4

Variation: 3.25

1981 APPLICATION of TASC Aircrew "Figures of Merit"

	<u>DEP</u>		<u>EEP</u>	
IAF	5350		6950	
		$= \frac{1.62}{1}$		$= \frac{5.27}{1}$
EAF	3300		1320	
IAF flying EAF force	3300		4290	
		$= \frac{1}{1.62}$		$= \frac{2.0}{1}$
EAG flying IAF force	5500		2140	

Do such numbers explain IAF results in 1982?

OVERVIEW

- COMPARISON of EXISTING “MOE” SYSTEMS
 - Motivations, Costs, Problems of MOEs
 - Force Comparisons Using Selected Systems
 - Strengths and Weaknesses of Current “MOEs”

- ENDURING LIMITATIONS of “MOEs”
 - “Intangibles”
 - *The Nonlinearity of Combat Processes*

UNIT VARIABILITY

8th AAF Fighter Groups (2 Oct 42-31 May 45)

Fighter Group	Numbered Missions	CLAIMS			LOSSES "MIA"	EXCHANGE RATIO
		A-A	A-G	TOTAL		
56th	447	674½	+ 311	= 985½	128	7.70-to-1
339th	264	239½	+ 440½	= 680	97	7.01-to-1
352nd	420	519½	+ 287	= 806½	118	6.83-to-1
479th	351	155	+ 279	= 434	69	6.29-to-1
357th	313	609½	+ 106½	= 716	128	5.59-to-1
353rd	447	330½	+ 414	= 744½	137	5.43-to-1
355th	?	365½	+ 502½	= 868	175	4.69-to-1
4th	?	583½	+ 469	= 1,052½	241	4.37-to-1
78th	450	338½	+ 358½	= 697	167	4.17-to-1
20th	312	212	+ 237	= 449	132	3.40-to-1
364th	342	256½	+ 193	= 449½	134	3.35-to-1
359th	346	253	+ 98	= 351	106	3.31-to-1
55th	?	316½	+ 268½	= 585	181	3.23-to-1
356th	413	201	+ 75½	= 276½	122	2.27-to-1

$$\frac{7.70}{2.27} = 3.39$$

The Problem of "Non-Shooters"

"Later when the companies were interviewed at a full assembly and the men spoke as witnesses in the presence of the commander and their junior leaders, we found that **on an average not more than 15 per cent of the men had actually fired at the enemy** positions or personnel with rifles, carbines, grenades, bazookas, BARs [Browning automatic rifles], or machine guns **during the course of an entire engagement**. Even allowing for the dead and wounded, and assuming that in their numbers there would be the same proportion of active firers as among the living, the figure did not rise above 20 to 25 per cent of the total for any action. The best showing that could be made by the most spirited and aggressive companies was that one man in four had made at least some use of his fire power."

"In the whole of the initial assault landings on the Omaha Beachhead, there were only about five infantry companies [out of two divisions] which were tactically effective the greater part of June 6, 1944. In these particular companies an average of about one fifth of the men fired their weapons during the day-long advance from the water's edge to the first tier of villages inland—a **total of perhaps not more than 450 men firing consistently with infantry weapons in the decisive companies.**"

— S. L. A. Marshall, 1947

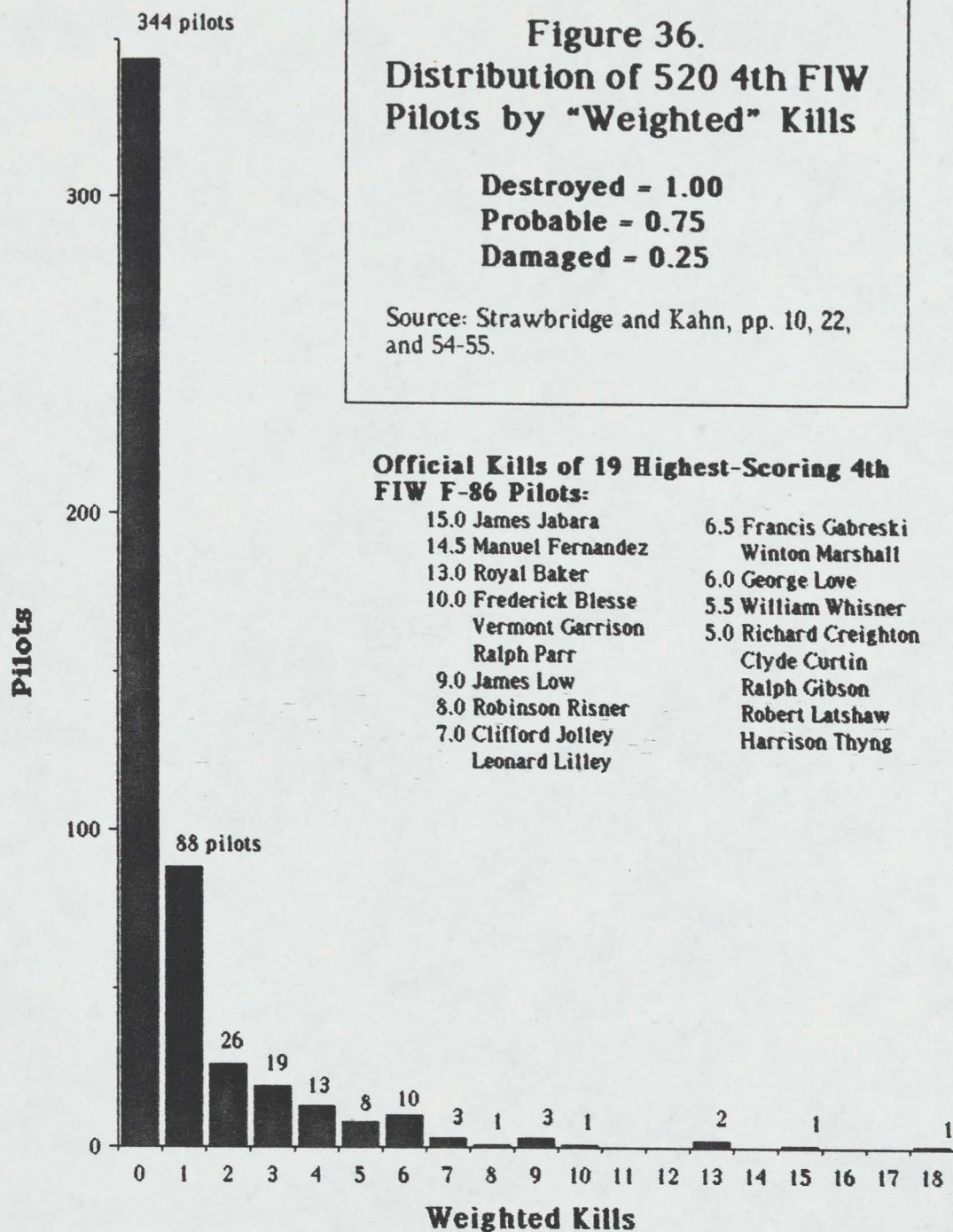


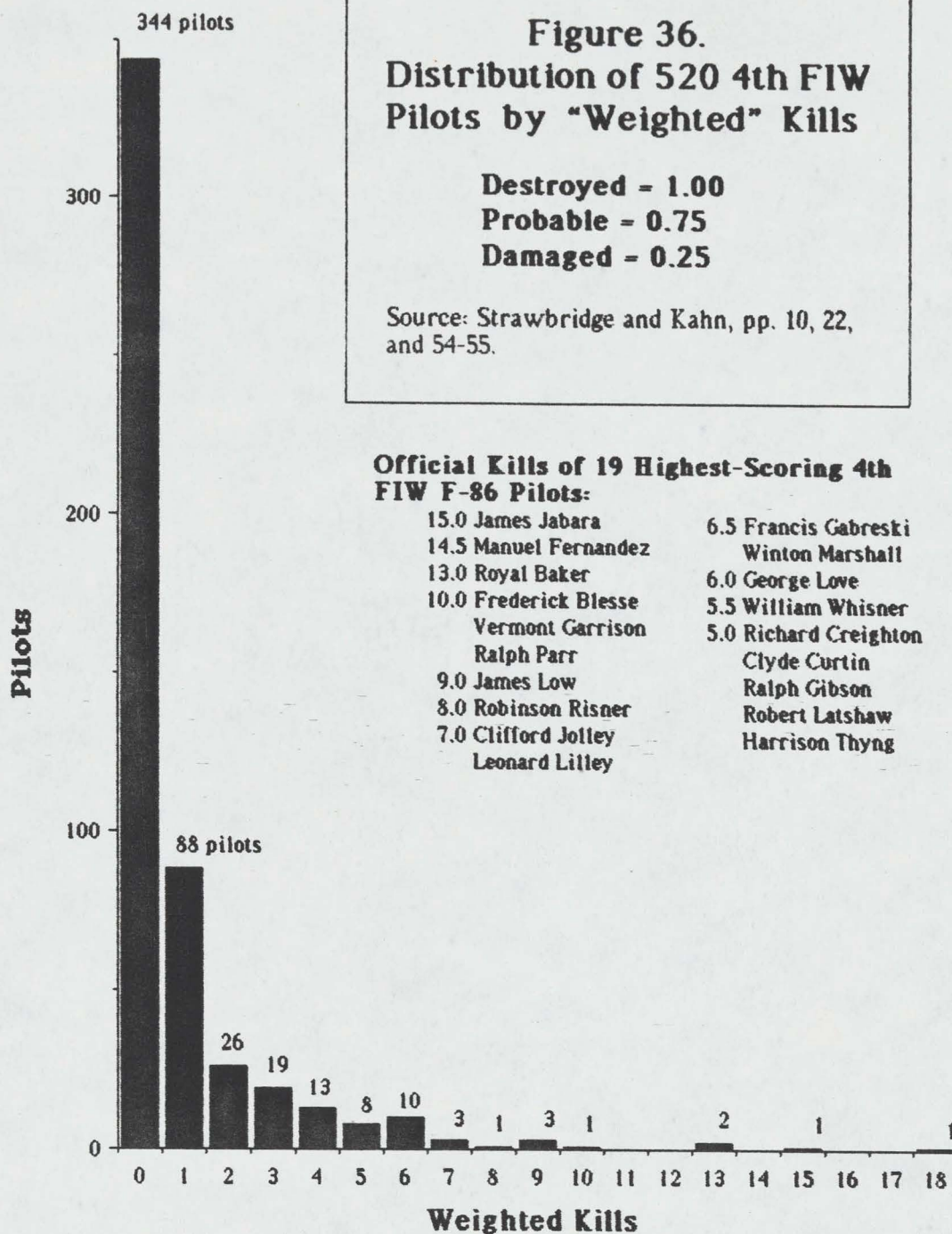
Figure 36.
Distribution of 520 4th FIW
Pilots by "Weighted" Kills

Destroyed - 1.00

Probable - 0.75

Damaged - 0.25

Source: Strawbridge and Kahn, pp. 10, 22,
 and 54-55.



INDIVIDUAL VARIABILITY

4th FIW F-86 Pilots (Dec 50-Jul 53)

*Interquartile Performance Range (middle 50% of the distribution)
for all 520 pilots:*

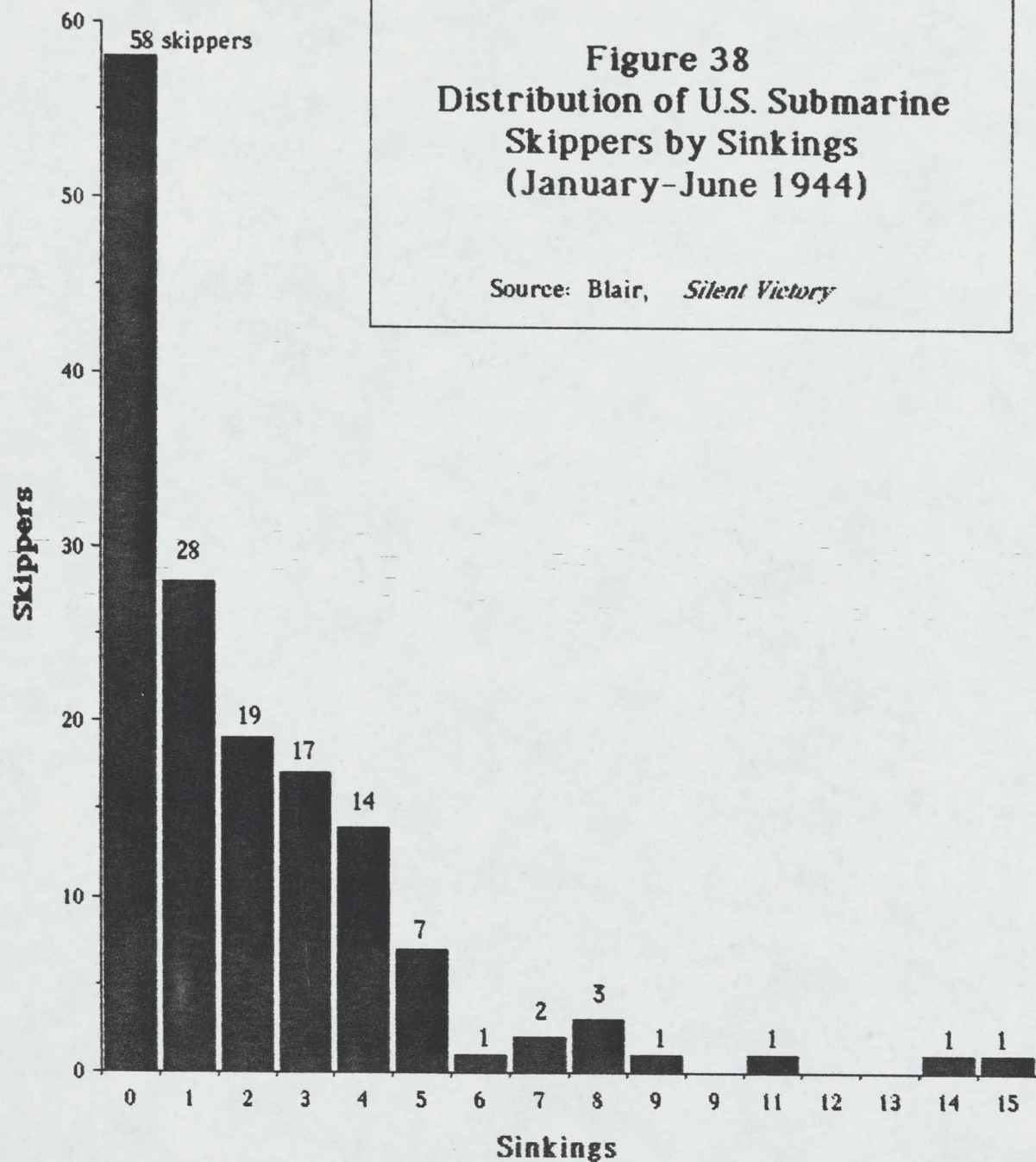
Sightings/Mission	.03 - .17
Encounters/Sighting	.57 - .84
Firing Passes/Encounter	.31 - .58
Kills/Firing	.18 - .48

$$\frac{17\% \times 84\% \times 58\% \times 48\%}{3\% \times 57\% \times 31\% \times 18\%} = 42$$

Variation for 69 pilots with 15 or more encounters as leaders:

Firing Passes/Encounter	.12 - .88	Highest/Next-to-Lowest:
Kills/Firing Pass	.00 - .70	$\frac{88\% \times 70\%}{43\% \times 3\%} = 48$

19 of the 4th FIW's 520 PILOTS PRODUCED
153 of the wing's 484 KILLS (31.6%) in "MiG Alley"



Is Individual Combat Performance Predictable?

Strawbridge and Kahn (1955):

"In the aggregate, these points all indicate that the qualities necessary for success in [air-to-air] combat are not well understood....It seems evident that reliable prediction of combat effectiveness must await the development of much more subtle psychological and physiological tests."

Torrance, Rush, Kohn and Doughty (1957):

"The present data give an interesting picture of Korean aces, and they suggest hypotheses for future work on selection and training of aces, but the observed differences cannot be confidently used as bases for prediction."

Clay Blair (1975):

"The truth was, there was absolutely no way to tell in advance whether a new submarine skipper would be aggressive or nonaggressive. What he did, how he performed, was the result of deep inner motivation that could not be quantified or assessed in port or even while he was an exec. An aggressive exec might turn cautious when left alone and handed responsibility for a ship and the lives of sixty men. There was only one solution to the problem: Give everybody who deserved it and was qualified a chance to see how he made out; then, if he failed to perform, relieve him after two patrols."

What Are the Real Drivers in Outcomes?

1979 Retrospective on ACEVAL (1977):

"On the short term basis, though, **numbers account for only about 10 to 20% of the variation in results. The rest of the variation is due to the things that people do** [that is, things like situation awareness, mutual support, and getting a tally]."

AMRAAM OUE Results (1981-82):

In multi-ship air-to-air engagements, **human interactions outweighed hardware interactions. Situation awareness**—the individual aircrew's "perception of the whole picture, not only location but also likely future activity, both friendly and enemy"—**was statistically the single most important factor in determining engagement outcomes.** All other test variables—including performance differences between opposing aircraft, avionics, air-to-air weaponry, force ratios, and the tactical environment (GCI, BVR vs WVR, etc.)—proved less important.

— S. R. "Shad" Dvorchak, 1985

Are Combat Processes Really Expressed by Simple, Linear "Laws"?

Analysis of the Role of Casualties in 80 Battle Outcomes (1941-1982):

"No matter how casualties are measured, battles have been given up as lost when casualties ranged from insignificant to overwhelming....The outcome of battle modeled by the Lanchesterian equations postulates a development of combat in response to casualties incurred. **During the last 50 years**, however, **battles appear to have been resolved largely on the basis of other considerations** [primarily due to being outmaneuvered]."

— Robert McQuie, 1987

"... my assertion that **there is no discernible direct relationship between strength ratios and advance or attrition rates** is not a matter of interpretation. As the graphs show, it is simple historical fact, undeniable for the vast majority of 601 battles in the [Land Warfare] data base between the years 1600 and 1973.

"For several years our operations research analysts and model designers seem to have recognized that our models are wrong in their representation of the fundamental combat processes of attrition and movement. But it has not been easy for them to accept this fact."

— Colonel Trevor N. Dupuy, August 1987

Slide 42 8/24/88 "Simple Mathematical Models with Very Complicated Dynamics, *Nature*, 10 June 1976, p. 459; "Chaos," James P. Crutchfield, J. Doyne Farmer, Norman H. Packard, Robert S. Shaw, *Scientific American*, Dec 86, p. 46.

Determinism ≠ Predictability

"First-order difference equations arise in many contexts in the biological, economic and social sciences. Such equations, **even though simple and deterministic, can exhibit** a surprising array of dynamical behavior, from stable points, to a bifurcating hierarchy of stable cycles, to **apparently random fluctuations.**"

—Robert M. May, 1976

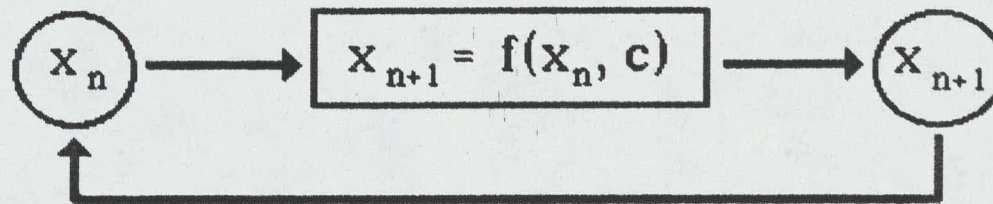
"... until recently there was little reason to doubt that precise predictability could in principle be achieved. It was assumed that it was only necessary to gather and process a sufficient amount of information.

"Such a viewpoint has been altered by a striking discovery: **simple deterministic systems with only a few elements can generate random behavior.** The randomness is fundamental; gathering more information does not make it go away. Randomness generated in this way has come to be called chaos [meaning 'deterministic disorder' and 'sensitive dependence on initial conditions']."

—Crutchfield, Farmer, Packard, and Shaw, 1986

Nonlinearity

The **essence of nonlinearity** lies in “**simple feed-back processes** in which the same operation is carried out repeatedly, the output of one iteration being the input for the next one.”



$f(X_n, c)$ is a nonlinear function (that is, more than a simple proportionality, $X_{n+1} = kX_n$)

c is a parameter (a complex number in the case of the Mandelbrot set)

“The smallest differences are blown up and have far-reaching effects. Causality holds at every instant, but it does not carry over a sequence of branching. Sooner or later the initial knowledge of the system becomes irrelevant.”

In simple terms, a “chaotic” system is one in which “there is no way to predict its long time behavior.” The process is *determined* by its initial value, yet it cannot be *predicted* “other than by letting it run.”

— H.-O. Peitgen and P. H. Richter, 1986

DOs and DON'Ts

DO:

- Use Common Counting Rules
- Explain Counting Rules and Assumptions
- Analyze the Details
- Use a Variety of Tools
- Examine Trends
- Brief "Results" to Experts and Get Their Opinions
- Look at Alternatives
- Recognize that the "Doing" is of Major Value

DON'T:

- Overlook the Value of Military Expertise and Judgment
- Overlook the "Intangibles"
- Rely only on "Analytic" Tools
- Overlook the Value of "Bean Counts"
- Believe Models Represent Reality (Especially Lanchester Models)
- Forget to "Think Red"
- Lose Objectivity
- Be Afraid to Make Synthetic Judgments