Operations Research

September–October 1966

COMBAT MODELS AND HISTORICAL DATA: THE U.S. CIVIL WAR

Herbert K. Weiss*

Palos Verdes Estates, California

(Received September 21, 1964)

Doubt has recently been cast on the applicability of mathematical models in general and Lanchester-type equations in particular to real combat. A comparatively small amount of effort has been spent in the past in testing theory against data, and so the general validity of such modeling largely remains to be demonstrated. In this paper, combat data on the U.S. Civil War is examined to determine the extent to which it can be explained by simple mathematical relations, and some support, as well as problems, for such modeling is developed. A model for probability of winning is developed, depending solely on the ability of each force to continue to fight after sustaining specified fractional losses. Distributions of parameters describing battles, force ratios, and loss ratios are derived. Limited applicability of Lanchester-type equations is indicated. It is concluded that although these hundred-year-old battles may have limited application to future combat, the possibilities of deriving improved combat models by analyzing real life data are essentially unlimited, and that fact differs sufficiently from theory to make it unlikely that pure conjecture will hit upon the proper form for a combat model structure.

SINCE F. W. LANCHESTER propounded his 'N-square' law[1] (and its complementary linear form), dozens of papers have been written playing variations upon the mathematical theme which he expounded so clearly that it scarcely seemed to require experimental verification. Some justification for these elaborations was provided by the papers of ENGEL[2] and WEISS,[3] which indicated that the progress and outcome of certain battles in the Pacific Theater of Operations in World War II were consistent with the Lanchester 'square law.'† For the current state of mathematical

* This study is the result of the author's hobby. Selected preliminary results were presented at the 14th Military Operations Research Symposium, Coronado, Calif., 27–29 October 1964. The present paper includes additional data, extended analysis, and findings.

† Although Engel stated clearly that other forms of Lanchester's equations might also fit his data on Iwo Jima.

modeling deriving from the original Lanchester-type formulation, the reader is referred to DOLANSKY,[6] whose paper contains an excellent set of references to relevant investigations. Recently WILLARD,[5] in a significant paper examining some 1500 land battles of modern civilization, has concluded that "there is little value in a simple version of Lanchester's equations as a predictive tool, where the only known quantities are initial strengths."

For those who are interested in attempting to dispel some of the 'fog of war' by searching for quantitative relations that will assist in predicting the outcome of combat and aid in the design of forces, Willard's conclusion is both a discouragement and a challenge. As a minimum, it suggests a return to the data to determine whether there are, in fact, useful consistencies and patterns and to determine whether, if Lanchester's 'laws' are false, there may not be other relations as helpful.

The search for the necessary data is a thankless one, especially for recent wars. One would like to know, at least, the number of forces of each kind participating on both sides, and the losses. Willard's fifteen hundred points were taken from the battles of 1618 to 1905, tabulated in Bodart's "Kriegs-Lexikon"; for the battles of World Wars I and II, the writer has found no modern-day Bodart. The many spirited and colorful American histories of recent wars are particularly discouraging in this regard, however interesting they may be to read. British historians are less reluctant to include an occasional table of forces or casualties (rarely both); but the average military historian is particularly susceptible to the criticism aimed by VAGTS[18] at the "average military officer" of avoiding "bellometrics" "as something too materialistic and derogatory to military art."

There is, however, one war for which data is available in moderation, if not in abundance, the U.S. Civil War. It is not a modern war, but it stands at the threshold of the era of mechanized warfare. The Civil War was the first in which railroads were extensively used for the transportation of troops, aerial reconnaissance was effectively used, command and control was exercised through the electric telegraph, ironclad naval vessels engaged in combat, a multimanned submarine sank a naval vessel, rifled artillery came into general use, wire entanglements were used in field fortifications, and the repeating rifle was used by large troop units.[17]

Furthermore, it is to the Civil War Battle of Chancellorsville that KANE has appealed[7] in arguing against 'computerized war.' At Chancellorsville Lee, outnumbered two to one, defeated Hooker; and Kane argues that a computer forecast of the outcome would have predicted a Union victory.

As for the war itself, analysis of Civil War data may not be intrinsically valuable with regard to application to future wars. It may help to answer the more general question of whether models may describe any part of *any* war.

The paper begins with a brief review of the over-all characteristics of the Civil War with regard to size of forces involved, number of battles, total losses, and some distributions of force ratios and casualty ratios. It then examines the applicability of Lanchester's laws. It is shown that a categorization of battles into 'assaults on fortified lines' and 'other battles' is significant. A relation between casualty rate and the ability to continue a battle is then developed, and this is shown to describe fairly well the probability of winning a battle as a function of relative casualty rates. The influence of fortifications is demonstrated. Finally, some suggested areas for future analysis are indicated.

If, in this paper, the bare bones of statistics seem to have been stripped of the human flesh and blood, 'which really matters' in warfare, perhaps they compensate if they reveal a structure that has constrained, if not prevented, the individual acts of will that, as Kane[7] points out, "cannot be handled by scientific methodology."

SOURCES OF DATA

THREE PRINCIPAL sources of data were used for this paper. They are PHISTER's "Statistical Record,"[8] LIVERMORE's "Numbers and Losses in the Civil War,"[9] and BODART's "Kriegs-Lexikon."[10]

Phister records, by place of encounter, 2261 engagements of the war, and gives Union killed, wounded, and missing for 149 battles and engagements in which total Union loss was 500 or more. However, he gives only total Confederate losses; these are available for most of the 149 cases.

Livermore has made a meticulous effort to determine the total numbers of men involved on both sides in the war as a whole, and in all battles in which the number hit was 1000 or more on either side. He lists 64 battles, for most of which he is able to give the size of force involved on each side and the number hit (killed plus wounded) on each side. In addition, Livermore designates the victor and loser in most of the battles and identifies those which involved assaults on fortified lines. The latter characteristic turns out to be of great importance in analysis of the data, as will be shown.

The third source, Bodart's "Kriegs-Lexikon," gives data on 50 battles. The numbers given by Bodart are generally consistent with those of Livermore, but Livermore's estimates of numbers actually engaged seem to have been done more carefully and so are used in this paper. Bodart, however, gives in addition for some battles a breakdown of losses among general officers (by name), officers, and men, as well as cannon involved and captured, and the sizes of infantry, artillery, and cavalry units.

The distribution of battles and engagements by year, according to various criteria and sources, is given in Table I.

RESOURCES AND TOTAL FORCES

At the beginning of 1861 the population of the free Northern States was about 18,900,000. The four slave states that remained with the Union had a white population of about 2,600,000 and 430,000 slaves. These four states are reported to have furnished men to both sides in about equal numbers. The white population of the Confederate States in 1861 was about 5,500,000, with a slave population of about 3,500,000.

TABLE I

Battles and Engagements of the Civil War Listed in Various Sources

Year	Reference 17[(a)]	Reference 8[(b)]	Reference 11	Reference 8[(c)]	Reference 9[(d)]	Reference 10[(e)]
1861	NA	156	44	5	2	3
1862	NA	564	146	37	20	16
1863	NA	627	122	22	11	11
1864	NA	729	154	62	26	16
1865	NA	135	29	23	5	4
Total	6000	2261	495	149	64	50

(a) States that there were some 6000 engagements in the Civil War.
(b) "Chronological List of Engagements, Etc."
(c) Battles in which Union loss was 500 or more.
(d) Battles in which number hit was 1000 or more on either side.
(e) Casualties on both sides totalled at least 2000.

Livermore has estimated the average size of the Confederate Army as about 55 per cent that of the Union, the ratio decreasing from about 60 per cent in 1861 to 50 per cent in 1865. The peak strength of the Union force has been estimated at just under one million men by Livermore. Other sources place the value at almost exactly a million.

Estimates of the total number of men serving vary widely, ranging from two to three million for the Union and from 600 thousand to 1.5 million for the Confederacy.

Total killed and wounded were estimated by Livermore at 385,000 for the Union and 329,000 for the Confederacy. The fact that losses on both sides were nearly equal, although average force ratio was almost two to one favoring the Union, suggests that if the over-all force ratio carried over into individual battles and if Lanchester laws applied, the linear law (which predicts equal losses) is more likely to be appropriate than the square law (which would predict Union losses one-quarter those of the Confederacy). As will be shown, the Confederates almost invariably succeeded in obtaining better force ratios in individual battles than the relative over-all strengths would indicate, but the square law is not found to be justified.

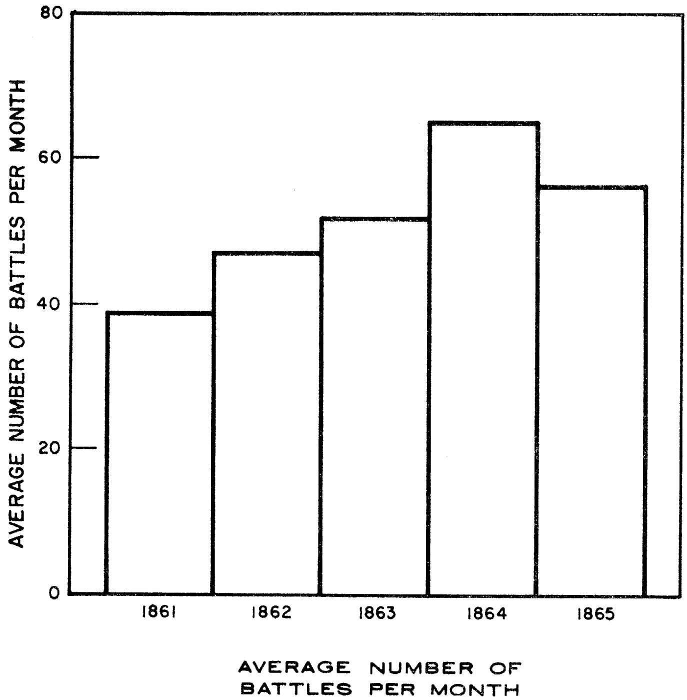

Fig. 1. Frequency of battles by year.

INTENSITY OF CONFLICT

Active conflict, as measured between the first and last battle in which several hundred men were lost, extended from the first Battle of Manassas,* July of 1861, to April of 1865—about $3\frac{3}{4}$ years. The growth in the intensity of conflict, as measured by the average number of battles per month (from Phister) and the cumulative losses (killed, wounded, and missing), is shown in Figs. 1 and 2.

It is notable that cumulative losses on both sides were almost equal. In computing Fig. 2 from reference 8, it was necessary to estimate Confederate losses for 16 of the 149 battles for which numbers were not given.

* Both sides had a fantastic 'overkill' capability in ammunition stocks for use in the first Battle of Manassas, allowing between 8000 and 10,000 bullets to be fired for every man killed or wounded, and inflicting an average of 5.5 per cent casualties.

They were for each of these battles assumed to be equal to Union losses; the correction was negligible except for 1864, for which year 57,000 was added to the Confederate total by this method. Since the Confederate loss

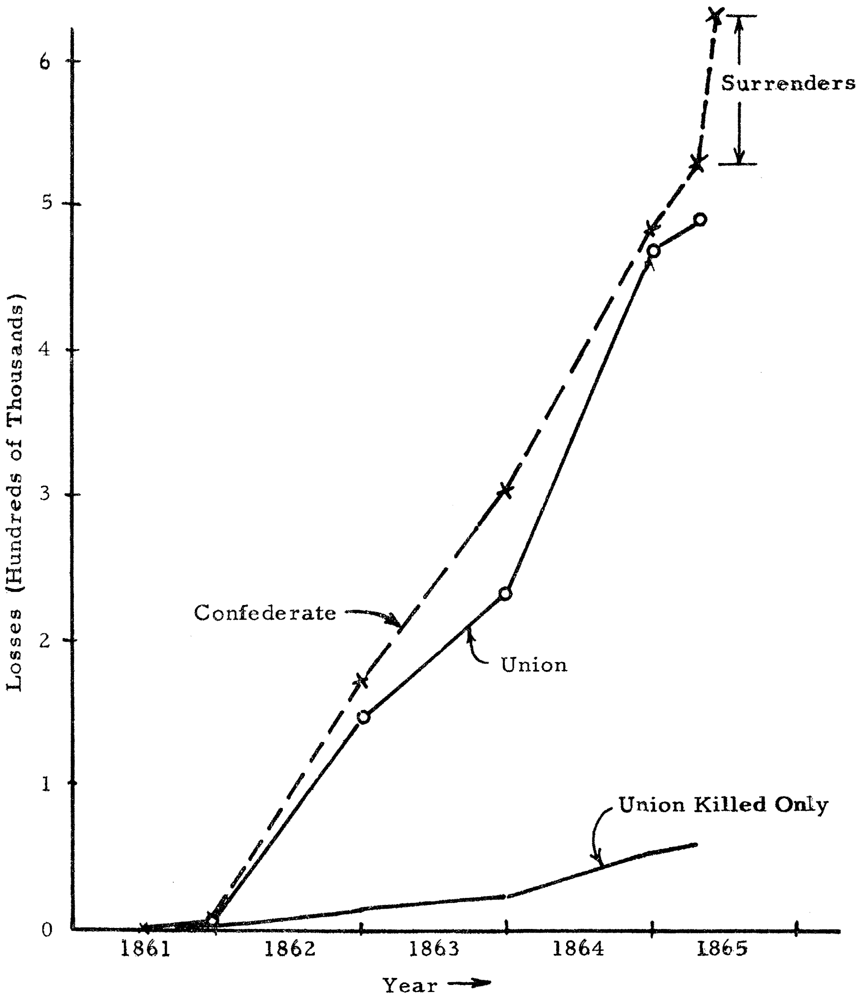

Fig. 2. Cumulative losses killed, wounded, and missing.

data from reference 8 are probably the least accurate of the Confederate loss data used, and since both sides include 'missing' as well as casualties, the small difference between the two cumulative curves should not be considered of significance.

As Table I shows, the number of engagements identified varies very strongly with the casualty rate used as a lower bound. Figure 3 shows the distribution of battles by size, i.e., the number of battles in which the Union loss was greater than a specified number. Also shown is a similar

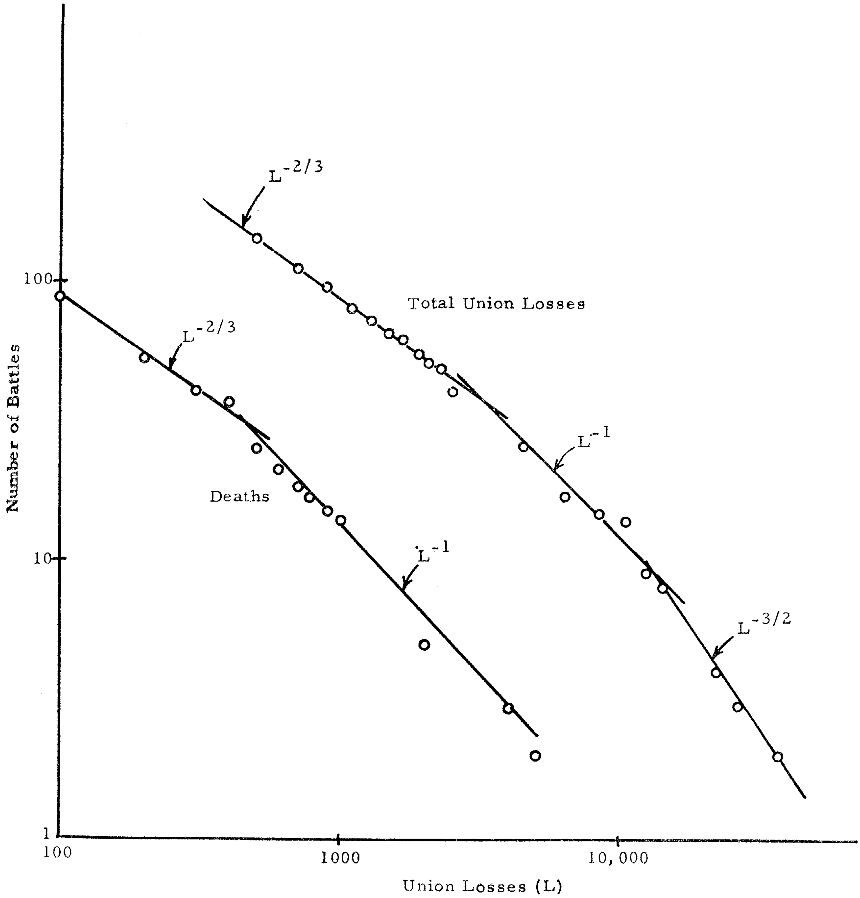

Fig. 3. Number of battles in which Union losses exceeded '*L*.'

distribution for deaths in the Union force. Since losses on both sides were on the average about equal per battle, total losses for both sides would follow these curves with abscissas doubled.

FORCE RATIOS IN BATTLES

THE DETERIORATION in the over-all ratio of Confederate to Union strength as the war progressed affected the force ratios in individual battles. Figure

4, based on Livermore's data, shows that the Confederates managed to obtain numerical superiority on the battlefield in the majority of the initial major battles, but slowly lost this capability as the war progressed.

It is notable that in spite of the fact that the Confederate Army numbered only from 50 to 60 per cent of the total Union Army at any time, according to Livermore's estimate, the Confederates almost invariably succeeded in securing a better ratio than this in the battles listed by Liver-

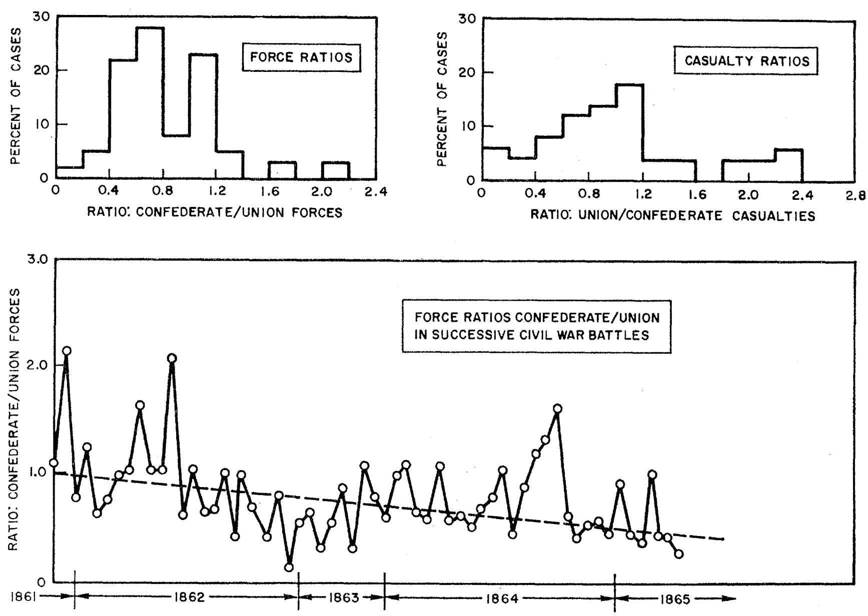

Fig. 4. Distribution of ratios by frequency and time.

more, as shown by Fig. 4. Figure 4 also shows histograms of force ratios and casualty exchange ratios.

To get some idea of variability of casualty rates across battles, the histograms of Fig. 5 have been constructed from Livermore's data. The mean Confederate casualty rate per battle was 12.3 per cent; the mean for the Union per battle was 10.9 per cent. Considering the ability of men to produce casualties (i.e., opposing casualties divided by own force size), the Confederate mean was 15.5 per cent as compared with the Union's 11.5 per cent. Union forces thus suffered a lower fraction of casualties per battle than the Confederates and inflicted slightly fewer casualties per man per battle.

EXCHANGE RATIOS AND FORCE RATIOS

We now wish to determine whether there is any discernible relation between casualty ratios and force ratios. Lanchester's linear law states that exchange ratio (casualty ratio) should be independent of force ratio; his

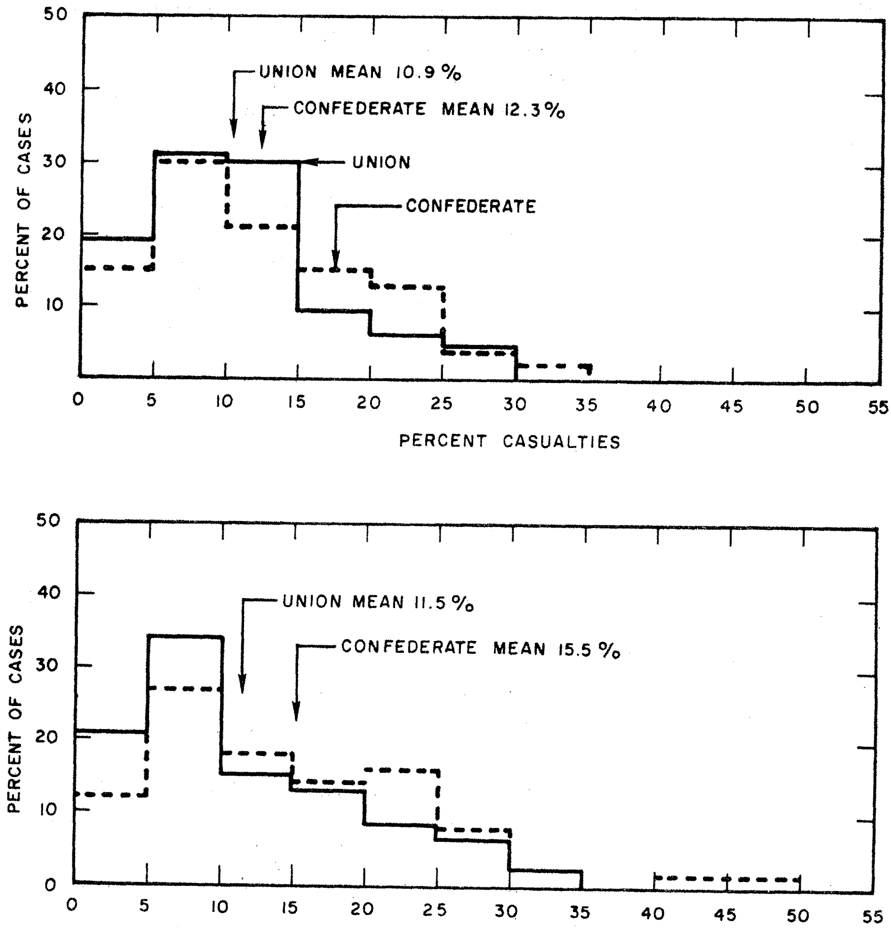

Fig. 5. Distributions of casualty rates received and produced.

square law, that it should vary inversely as force ratio. The Willard study indicates that for the land battles given by Bodart, neither Lanchester law holds and, in fact, the larger force tends to sustain the larger number of casualties.

For each battle for which the necessary quantities were given in the

Livermore list, casualty ratio and force ratio were computed. A scatter diagram locating each battle in these coordinates showed no apparent correlation of casualty ratio with force ratio. The points scattered very widely in the diagram.

At this point it was noted that Livermore singled out certain battles as 'attacks on fortified lines.' These battles were therefore aggregated in one class, and all other battles collected in a second class, called for convenience, 'meeting engagements.' Typical of battles involving assaults on fortified lines are Chattanooga and Vicksburg. 'Other' included Chancellorsville, Shiloh, and Antietam.

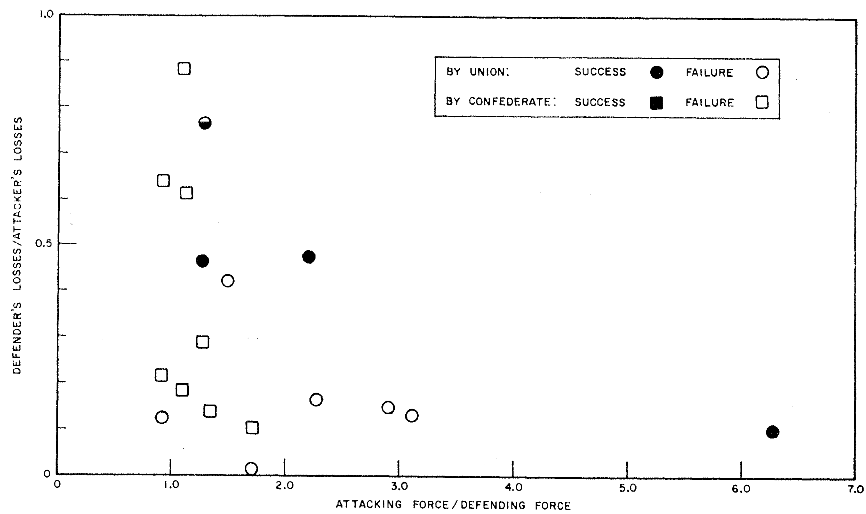

Fig. 6. Force and loss ratios in assaults on fortified lines.

Lifting of the Confederate siege of the Army of the Cumberland in Chattanooga required Union troops to overrun a series of fortified positions, including those on Tunnel Hill and Missionary Ridge. On Missionary Ridge, Cleburne's division, well-entrenched, held off six divisions under Sherman for the duration of the battle, which was decided when Thomas's troops overran the trenches in the Confederate center and continued without stop to capture the ridge.

At Chancellorsville, although both sides had entrenched, while Lee with a part of his force held the attention of Hooker, Jackson marched around the Union forces and attacked from the rear, rolling up the weak Union right flank.

At Shiloh, the Union force was neither entrenched nor prepared when Johnston burst upon the Union camp, catching the troops in their tents.

Vicksburg, when Grant attacked it in 1863, was heavily fortified with entrenchments that had been prepared long before the Confederate forces withdrew to them, and a series of Union assaults failed. Vicksburg finally capitulated to siege.

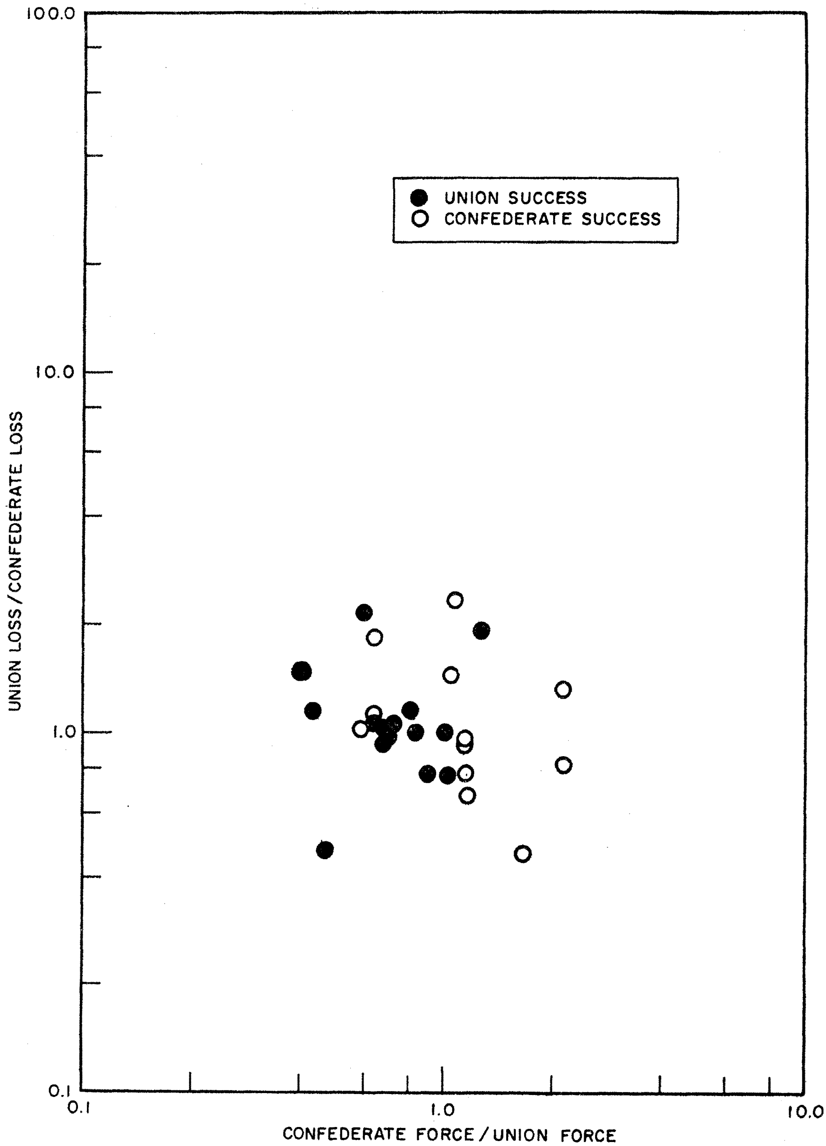

Fig. 7. Force and loss ratios in battles other than assaults on fortified lines.

Summarizing, without examining Livermore's categorization battle by battle, it appears that battles in which fortifications played a minor part, or were evaded, or in which natural terrain features played a principal part, were not designated as 'assaults on fortified lines.' An indication of the significance of Livermore's dichotomy is that it does appear to divide the battle data into two groups of differing characteristics. These will now be examined in turn.

Figure 6 shows casualty ratio versus force ratio for assaults on fortified lines. It is notable that there were few successes and very wide variations in loss ratio, with the attacker always sustaining greater losses than the defender. But since the defender fired from the protection of his fortification, it is clear that in this type of attack the ratio of defender's loss to attacker's can easily approach zero if the attacker does not have sufficient strength to overrun the fortification. For force ratios in which the attacker is unable to overrun the position, the casualty ratio is probably a function chiefly of the attacker's willingness to accept casualties before abandoning the attack. From Fig. 6, the required force ratio for a high probability of 'winning' appears to be at least 3.0.

Examination of battles other than assaults on fortified lines yields a more homogeneous picture, as shown in Fig. 7. There is no obvious correlation between casualty ratio and force ratio, suggesting that Lanchester's linear law might be justified here. The scatter in exchange ratio from battle to battle amounts to only a factor of two at the most.

When Union casualties are plotted against Confederate casualties, as shown in Fig. 8, for the battles previously used for Fig. 7, it is again observed that losses do tend to be fairly equal on both sides, i.e., there are no cases of very large casualties on one side while the other had few.

From Fig. 7 it is noted that when the force ratio favored the Union, the preponderance of battles resulted in Union victories. This characteristic suggested that further analysis of the data might be illuminating. This category of battles, 'other than assaults on fortified lines,' will be designated for convenience, 'meeting engagements.' It should be remembered, however, that the term 'meeting engagement' is used as a convenience and does not necessarily describe the manner of initiation of each battle.

MEETING ENGAGEMENTS

BECAUSE IT was felt that even in the absence of extensive fortifications the distinction between attacker and defender might be significant, narratives of almost all of the recorded battles were scanned[11, 13-16, 19] and a judgment made wherever possible as to which side was the attacker. Where given, Livermore's judgment as to the victor in each battle was accepted, although in at least one case he differs with Bodart, and it is known that the question of 'who won' in many battles is still a matter of controversy among Civil War buffs.

Of the battles considered, Peach Orchard to Malvern Hill and Williamsburg, according to Livermore, were considered to have no decisive victor since in the former case the Union, and in the latter, the Confederates, withdrew after repelling attack. In the case of Appomattox Camp, Livermore labels his designation of Confederate casualties 'estimate'; it is, however, included.

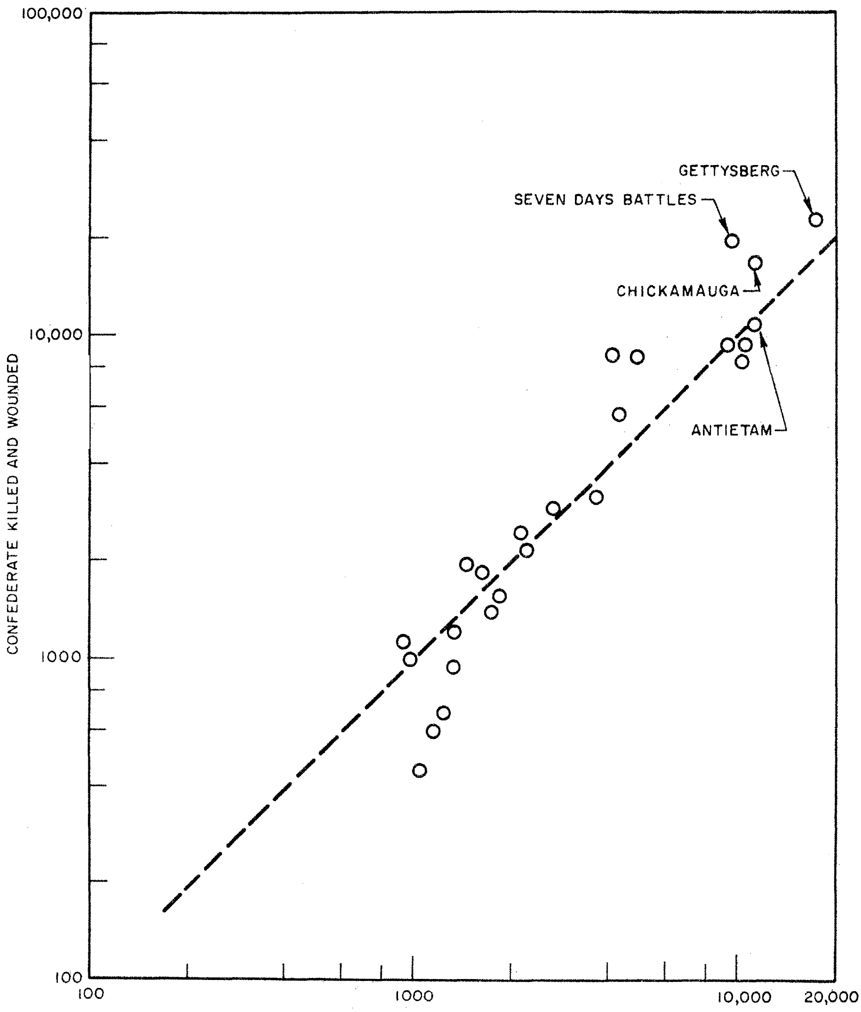

Fig. 8. Relative casualties.

This review yielded a total of 24 'meeting engagements,' with force and casualty information and identification of attacker and winner. Of these, the two battles noted above were considered to have an 'indecisive' outcome. Table II shows average casualty ratios, and Table III shows average force ratios for the 22 remaining battles. Numbers in parentheses indicate the number of battles constituting the average in each block.

TABLE II

AVERAGE CASUALTY RATIO; UNION/CONFEDERATES 'MEETING ENGAGEMENTS'

Winner	Attacker	
	Union	Confederates
Union	1.09 (4)[a]	1.06 (8)
Confederates	1.14 (4)	1.13 (6)

[a] ( ) Indicates number of battles.

It will be noted that for this collection of battles, segregated by winner, neither the force ratio nor the casualty ratio appears to vary significantly as a function of which side was judged to be the attacker. Possible implications are (1) the battles developed by entry of additional units on both sides at about the same rate, so that the effect of the initial attack was lost,

TABLE III

AVERAGE FORCE RATIO; CONFEDERATES/UNION 'MEETING ENGAGEMENTS'

Winner	Attacker	
	Union	Confederates
Union	0.62 (4)[a]	0.70 (8)
Confederates	1.14 (4)	1.28 (6)

[a] ( ) Indicates number of battles.

(2) each battle consisted of a series of attacks and counterattacks by both sides, (3) the designation of 'attacker' is faulty. More detailed map study of individual battles could resolve some of these possibilities.

It will be noted, however, that the winner had on the average a larger force ratio. If one considers that an advantageous force ratio may be a usual prerequisite to attack and compares average force ratios of each side as attacker regardless of outcome, for the 24 battles one finds (Table IV) that on the average the Union attacked with a 15 per cent force superiority, whereas the Confederates attacked with an average 5 per cent force inferiority. Looking at the 24 battles from a different point of view, one finds that the Union forces attacked in seven of the 15 battles where they had force superiority (47 per cent), while the Confederates attacked in six of the eight battles in which they had force superiority (75 per cent). For one battle the force ratio was unity, and the Confederates attacked. The

TABLE IV

AVERAGE FORCE RATIO OF ATTACKER

Attacker	Average ratio, Confederates/Union
Union	0.87 (9)
Confederates	0.96 (15)

difference in tactics that these averages indicate is a basic theme in qualitative analyses of the war.

To show the relation of force ratio to probability of winning, all 24 battles included in Tables II and III were pooled and four battles added for which no definition of attacker could be made. The two battles termed

TABLE V

FRACTION OF UNION WINS AS FUNCTION OF FORCE RATIO

Confederate / Union force	Cases	Fraction of Union wins		
		Average	50% Confidence limits	
0.40–0.49	3	1.00	1.00	0.63
0.50–0.79	11	0.68	0.80	0.54
0.80–1.25	11	0.50	0.64	0.36
1.26–2.00	1	0	0.75	0
2.01–2.50	2	0	0.50	0
	28			

'indecisive' were credited half to each side. Table V and Fig. 9 show the result. In Fig. 9 50 per cent confidence bands have been drawn. Also shown, as a smooth curve, is the function

$$P=(1+\mu^3)^{-1},$$

where μ = force ratio: Confederate/Union.

Taking 25 battles (2 draws and one unitary force ratio excluded from the sample of 28), it was found that when the Union force outnumbered the Confederates, it won 12 out of 15, or 80 per cent of the battles; when it was outnumbered, it won two out of 10, or 20 per cent.

Although Table II suggests a slight difference in casualty ratio, depending on which side wins, when the four battles for which the attacker was not designated were added to those of the table, the average ratios were as shown in Table VI; no dependence of casualty ratio on force ratio or who won is indicated.

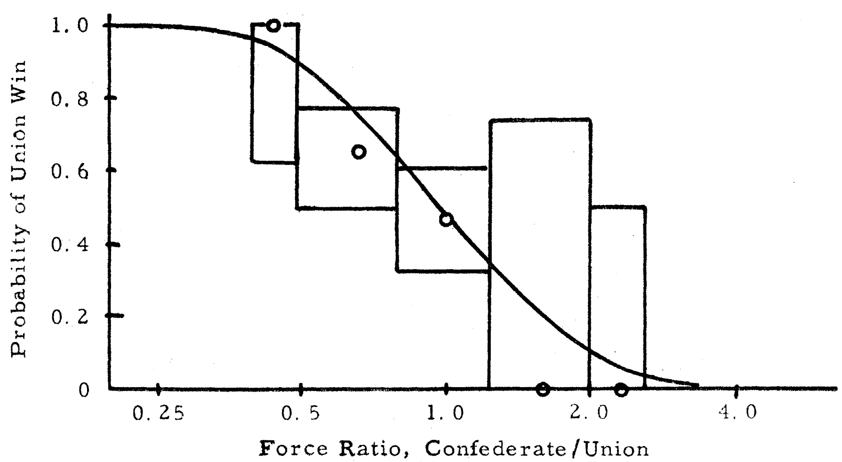

Fig. 9. Probability of Union winning vs force ratio (28 battles).

It is concluded, therefore, that for the complete sample of battles collected here as 'meeting engagements,' the casualty ratio is not obviously dependent on the force ratio and that casualties on both sides tend to be equal within a factor of about 2.0. However, more detailed study of the

TABLE VI

AVERAGE CASUALTY RATIOS, CONFEDERATE/UNION 'MEETING ENGAGEMENTS'

Force ratio:	>1.0	<1.0
Casualty ratio:	1.17 (10)[a]	1.14 (15)
Winner:	Confederates	Union
Casualty ratio:	1.15 (12)	1.13 (14)

[a] ( ) Indicates number of battles.

terrain and action of individual battles may help to explain differences from battle to battle.

In the 28 battles represented in Table VI, casualty ratio was less than unity in 12 battles, equal to unity in 2 battles, and greater than unity in

14 battles. Casualty ratio tends to follow a log normal distribution across battles; hence, the arithmetic averages shown in Table VI and prior tables are consistent with a median value of casualty ratio close to unity.

Since casualties tend to be about equal and the larger force tends to win, one would expect the winner's per cent casualties to be less on the average than the loser's. Table VII shows that this is the case. The average per cent casualty rate suffered by the loser was 15 per cent; that of the winner, 12 per cent.

This suggests a combat model based on ability to continue fighting as a function of sustained fractional losses, and such a model is developed in the following paragraphs.

TABLE VII

AVERAGE PER CENT CASUALTIES

	Force	
	Union, %	Confederate, %
Winner	11.6 (14)[a]	12.6 (12)
Loser	14.2 (12)	15.9 (14)

[a] ( ) Indicates number of battles.

A MODEL FOR PROBABILITY OF WINNING

TO RECAPITULATE the findings thus far, it is observed that

1. Probability of winning appears to be a strong function of initial force ratio.

2. Casualty ratios are independent of which side attacks, or who wins, or initial force ratio, and lie within the extreme values 0.46 to 2.33.

3. On the average the loser sustained 15 per cent casualties; the winner, 12 per cent.

The following simple model of the combat is therefore suggested:

The battle begins and, as it progresses and casualties mount on both sides, each side continuously evaluates its ability to continue to fight, using as a single criterion its own cumulative fractional losses to the moment of evaluation.

One may imagine a more complex criterion in which each side modifies its decision by knowledge of initial relative strengths (prebattle intelligence), or relative loss rates during the battle (combat intelligence), or relative force movement during the battle. At this point we prefer the simplest model, although further study of the data may suggest improvements along the above lines, or others.

It is assumed, further, that the casualty *ratio* at each battle's termination represents an exchange ratio characteristic of that battle and is sustained at a constant value throughout the battle. A stochastic modelling of relative casualty growth would be of interest, but the simpler choice will be made here.

Designate the two sides as ODD and EVEN and let

Δy_1, Δy_2 = ODD's, EVEN's cumulated casualties to time 't' in the battle,

Δx_1, Δx_2 = ODD's, EVEN's cumulated casualties at the termination of the battle,

x_{10}, x_{20} = ODD's, EVEN's initial force,

$$\begin{aligned} g_1 &= \Delta y_1/x_{10}; \qquad g_2 = \Delta y_2/x_{20}, \\ f_1 &= \Delta x_1/x_{10}; \qquad f_2 = \Delta x_2/x_{20}; \end{aligned} \tag{1}$$

let

$h_1(g_1)\, dg_1$ = probability that ODD gives up in dg_1, after having sustained a fractional loss g_1,

$h_2(g_2)\, dg_2$ = comparable probability for EVEN, (2)

$\phi_1(g_1)$ = probability that ODD continues to fight at least until it has sustained a fractional loss g_1.

Consider the probability that ODD continues to fight at least to g_1, EVEN not having given up either during this period. Divide g_1 into n small intervals, each of width Δg_1. The probability that ODD does not give up in the jth interval is very closely

$$1 - h_1(j\Delta g_1)\Delta g_1,$$

and the probability that ODD does not give up in any of the n intervals, since each decision is independent of prior decisions, is

$$\begin{aligned} \phi_1(g_1) &\cong \textstyle\prod_{j=1}^{j=n} [1 - h_1(j\Delta g_1)\Delta g_1] \\ &= \exp \textstyle\sum \mathrm{Ln}[1 - h_1(j\Delta g_1)\Delta g_1], \end{aligned}$$

or, allowing Δg_1 to become very small, expanding the logarithm, discarding terms in $(\Delta g_1)^2$ and higher, and replacing the sum by an integral

$$\phi_1(g_1) = \exp - \int_0^{g_1} h_1(g)\, dg,$$

and similarly

$$\phi_2(g_2) = \exp - \int_0^{g_2} h_2(g)\, dg. \tag{3}$$

Next, in accordance with the assumption of constant exchange ratio during the battle, define a constant R as

$$\begin{aligned} R &= g_1/g_2, \\ &= f_1/f_2. \end{aligned} \tag{4}$$

The probability ϕ that neither side has given up by the time losses g_1, g_2 have been sustained is

$$\phi = \phi_1 \phi_2. \tag{5}$$

The probability that, the battle having progressed to g_1, g_2, ODD then gives up in dg_1 is

$$dQ_1 = \phi h_1(g_1)\, dg_1,$$

and the probability that ODD gives up at all, summed over all g_1, is

$$\begin{aligned} Q_1 &= \int_0^1 \phi_1 \phi_2 h(g_1)\, dg_1 \\ &= \int_0^1 \phi_2\, d\phi_1, \end{aligned} \tag{6}$$

where the fact that $\phi_1 = \phi_1(g_1)$; $\phi_2 = \phi_2(g_2)$ and $R = g_1/g_2$ will allow ϕ_1 to be determined as a function of ϕ_2 and the integral to be evaluated.

The probability that ODD wins is

$$P_1 = 1 - Q_1, \tag{7}$$

and the probability that EVEN gives up is

$$Q_2 = \int_0^1 \phi_1\, d\phi_2. \tag{8}$$

As developed later (equation 20), it is possible to write

$$\phi_2 = \phi_1^{\,a}, \tag{9}$$

where a is a function of R and is constant during each battle.

Upon integrating (6) and (8)

$$P_1 = a/(1+a);\ P_2 = 1/(1+a). \tag{10}$$

EXPECTED LOSS FRACTIONS

THE EXPECTED value of the fractional loss in a battle, f, is obtained as follows: The probability that the battle ends in dg_1 is $-(\partial\phi/\partial g_1)\, dg_1$; hence, the expected value is

$$\begin{aligned} f_1 &= -\int_0^1 g_1(\partial\phi/\partial g_1)\, dg_1 \\ &= \int_0^1 g_1\, d\phi. \end{aligned} \tag{11}$$

Integrating by parts and noting that at $g_1=0$, $\phi=1.0$; $g_1=1.0$, $\phi=0$;

$$\begin{aligned} f_1 &= \int_0^1 \phi\, dg_1, \\ f_2 &= \int_0^1 \phi\, dg_2 \\ &= R^{-1}\int_0^1 \phi\, dg_1. \end{aligned} \tag{12}$$

APPLICATION TO COMBAT DATA

In order to apply the above expressions, it is necessary to define $h(g)$, preferably as an elementary function. We first determine $\phi(g_1)$ from the battle data, as follows:

All of the meeting engagements were ranked in order of increasing f_1 (the fractional losses suffered by ODD at the time the battle ended). Let N_0 = the total number of entries in the sample. Now if only ODD lost, we should expect to find $N_0\phi_1(f_1)$ entries indicating loss fractions in excess of f_1. In fact, choosing a particular f_1, we note that L_2 battles ended at lesser values of f_1 because EVEN lost. Had these battles not been terminated in ODD's favor, we should have expected $L_2\phi_1$ to have continued to f_1. From our table of battles we observe O_1 battles continuing to f_1; this is equated to the expected number less those terminated by EVEN, i.e.

$$O_1 = N_0\phi_1 - L_2\phi_1. \tag{13}$$

We therefore compute an estimate of ϕ_1 as

$$\phi_1 = O_1/(N_0 - L_2). \tag{14}$$

Continuing for EVEN, the same battles are ranked in order of increasing f_2, and ϕ_2 is computed as

$$\phi_2 = O_2/(N_0 - L_1). \tag{15}$$

Results are shown in Fig. 10, where ODD represents the Union and EVEN the Confederate forces. Plotting data in the form $\text{Ln}(1/\phi_j)$ against f_j; $j=1, 2$, we find that we can write

$$\text{Ln}(1/\phi_j) = kf_j^3; \qquad k = 150; \tag{16}$$

and that this fits the data for both sides, as shown in Fig. 10. Use of the same function for both ϕ_1 and ϕ_2, in the interest of simplicity, is indicated by Fig. 10 to be reasonable, although better fits to the separate data sets could be obtained by slightly different functions. Similarly, a slightly lower power of f than the cube would improve the fit, but increase the complexity of later computation.

Now since

$$\phi_1 = \exp -kf_1^3,$$

$$\phi_2 = \exp -kf_2^3,$$

$$\phi_2 = \phi_1^a,$$

where

$$a = (f_1/f_2)^3 = R^3.$$

Hence from (10)

$$P_1 = 1/(1+R^3); \qquad P_2 = R^3/(1+R^3). \tag{17}$$

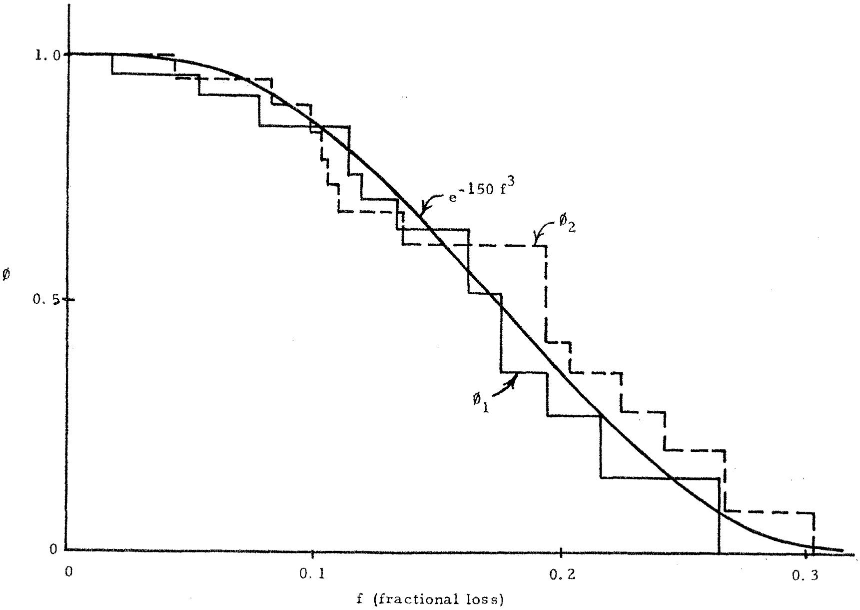

Fig. 10. Probability of continuing battle vs loss fraction (meeting engagements).

Equation (12) now becomes

$$f_1 = \int_0^1 \exp -kg_1^3(1+R^{-3})\, dg_1,$$

$$f_1 = \Gamma(4/3)k^{-1/3}(1+R^{-3})^{-1/3} - S, \tag{18}$$

where

$$S < t^{1/3}e^{-t}/3,$$

and

$$t = k(1+R^{-3}) > k.$$

Since $k = 150$, S is negligible for present purposes. Equation (18) may be written in more symmetric form

$$f_1R^{-1/2} = f_2R^{1/2} = \Gamma(4/3)/(kR^{3/2}+kR^{-3/2})^{1/3}. \tag{19}$$

Since R, the ratio of fractional losses, appears as a parameter in these equations, it was decided to subdivide the battle data by ranges of R into four equal groups and compare (17) and (19) with the averages of the data in each group. Results are shown in Table VIII.

Equation (17) describes the probability of winning as a function of fractional loss ratio R. But Fig. 9 indicated that force ratio μ was as good an indicator. The reason appears to be that, for these battles, casualty ratio is independent of force ratio; hence, for a given μ, writing r for casualty ratio and averaging (where $R=\mu r$) over a log-normal distribution function

TABLE VIII

Comparison of Data with Model Results 'Meeting Engagements'

Range of R	0.22–0.69	0.70–0.79	0.80–1.11	1.12–2.78
Average R	0.56	0.73	0.94	1.91
No. of battles	7	7	7	7
Fraction of Union (ODD) wins	0.79	0.57	0.50	0.29
Probability of win P_1 in av. R computed from (17)	0.85	0.72	0.55	0.13
Average Confederate fractional loss	0.17	0.19	0.11	0.08
f_2 computed from (18)	0.16	0.15	0.13	0.08

for r, centered on $r=1.0$,

$$P(\mu)=\int_0^\infty g(r)P(\mu r)\,dr$$

$$=\int_0^\infty g(r)[1/(1+\mu^3 r^3)]\,dr,$$

one would obtain a $P(\mu)$ curve not greatly differing from $P(\mu r)$. It is also possible that the dispersion in casualty data conceals more subtle effects of μ on $h(g)$ than has been assumed in these paragraphs.

ASSAULTS ON FORTIFIED LINES

The assaults on fortified lines differ from the battles considered in the preceding paragraphs in that the ratio of attacker's casualties to defender's casualties always exceeded unity and had an extremely wide scatter. Some of the elements of the previous model can be applied, however.

By computation, using (14) and (15), ϕ_j were obtained for the Union and the Confederates as attackers. As shown in Fig. 11, a single curve

$$\phi = \exp - 300f^3 \qquad (20)$$

fits both ϕ_1 and ϕ_2 fairly well for the attacking forces.

Comparing Fig. 11 with Fig. 10, we note that in an assault on a fortified position both the Confederate and the Union forces were less likely to accept a specified fractional loss before abandoning the attack than they were if

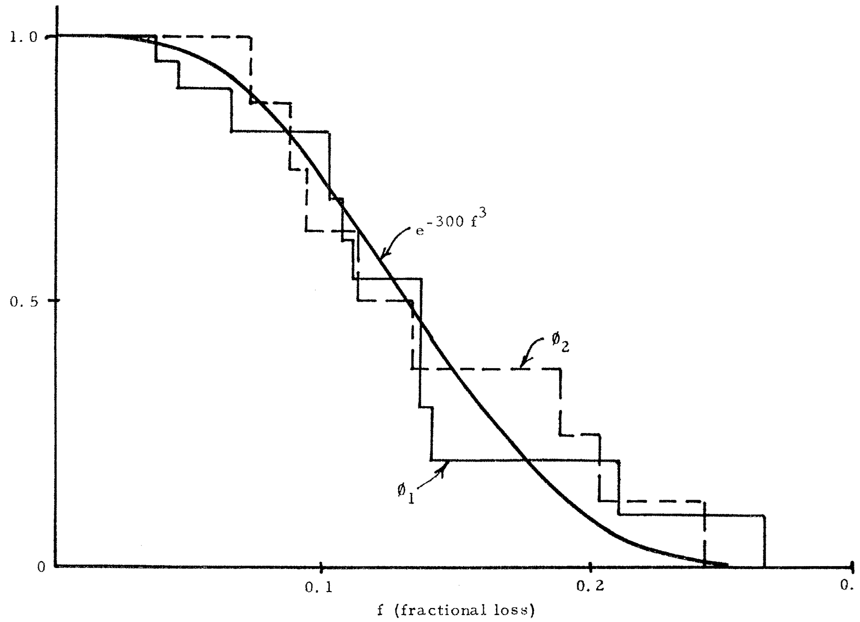

Fig. 11. Probability of continuing battle vs loss fraction (attacker in assaults).

meeting the enemy on more even terms. One may conjecture that the difference is explained by one of the following reasons:

1. Since the defender suffers fewer casualties than the attacker, losses in the assault that are too high will prevent the attacker from overcoming the defender when he overruns the position, i.e., the attacker does not really decide when to break off without regard for the remaining enemy strength; he breaks off because he feels that he has been weakened but the defender has not.

2. The psychological effect of attacking a fortified position has a deterring effect on the attacker. He has no way of telling what casualties he has inflicted on the defender. He may feel that even if he expended his whole force, he could not overrun the position.

The defender, too, seems to be in a different psychological position. There are only three assaults (Union) in the Livermore data in which the loss of the defender was given and in which the assault was a success. For these, the defender's fractional losses were 2.4 per cent, 5.5 per cent, and 12.3 per cent, compared with the attacker's losses of 3.6 per cent, 9.7 per cent, and 12.4 per cent. The defender, therefore, lost in two of these three cases with much lower casualty ratios than might have been expected in meeting engagements. However, this is an extremely small sample, and the data gives no convincing basis for not using equation (20) for the defense as well.

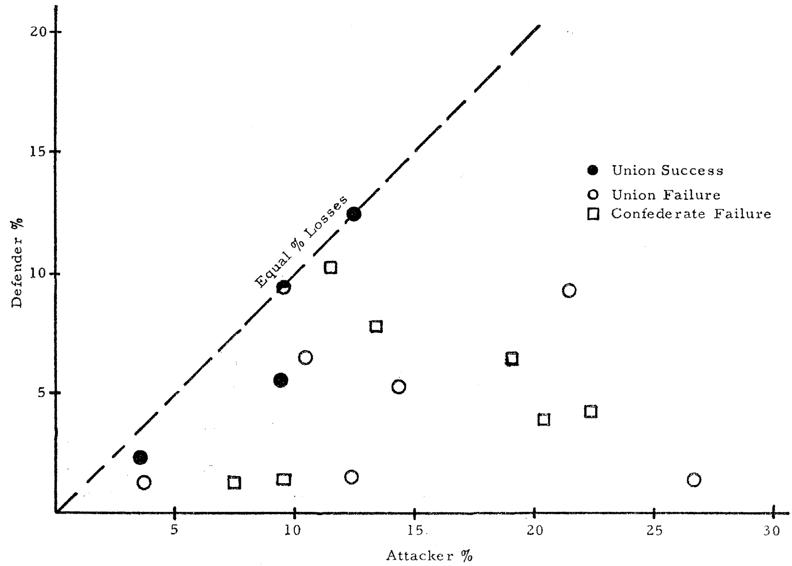

Fig. 12 Relative loss rates in assaults.

An indication of the requirement for a successful attack on a fortified position is given by Fig. 12, which compares fractional losses of attacker and defender. The few cases of successful attack lie near the line of equal fractional losses.

Comparing the ratio of fractional losses, f_1/f_2, for Union assaults, it is observed that in the four cases where this ratio was between 1.01 and 1.5, the Union had three successes and one partial success. For the six cases where the ratio lay between 1.61 and 20.5, the Union had no success. The ratio of fractional losses, f_2/f_1, in Confederate assaults lay between 1.1 and 5.9 and yielded no successes.

A model for an assault must probably consist of two sequential portions.

In the assault, the attacker's losses are comparatively high. If he persists until the position is overrun, defender's losses begin to rise. But the variability in strength of the defended position may be great.

In looking at the data, it is noted that, with one exception, Confederate casualties in response to Union assaults were small, unless the Union was victorious. This suggests that the model need look principally at the attacker's fractional loss and the initial force ratio. If the attacker's loss is high and his initial force ratio small, he will not have sufficient strength

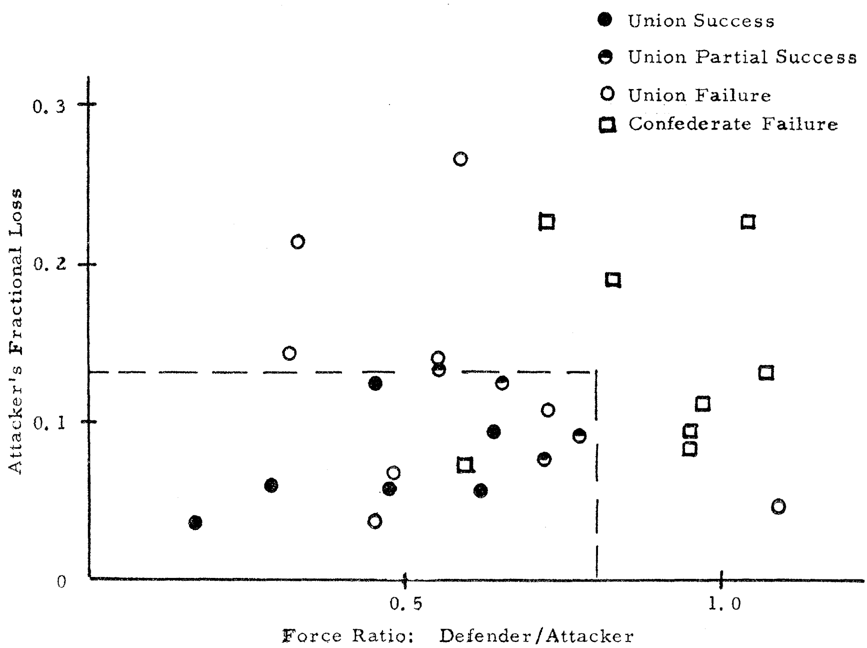

Fig. 13. Assaults on fortified line—attacker's loss vs force ratio.

to prevail after the assault. Furthermore, his losses in the assault should bear some relation to the strength of the defense.

Figure 13 shows the attacker's fractional loss plotted against the force ratio for all assaults. When the attacker lost less than 13 per cent of his force, and began with a force superiority of better than 1.25 to 1, he won in six battles out of 14 and had partial success in four, failing completely in only four. Outside the stated bounds, the attacker never succeeded (11 cases).

If the attacker's losses were proportional to the defender's strength, his loss ratio would be proportional to the force ratio; and we should expect that initial force ratio would be the principal determinant of victory, as was found in the case of meeting engagements. Figure 13 shows that this is not the case, and Fig. 14 indicates that the reason is the great variability of the loss ratio in small battles.

Figure 14 shows attacker's losses vs. defender's strength. In general, attacker's losses are proportional to defender's strength, with a dispersion about this mean independent of force size, i.e.,

$$\Delta x_a = c x_d + s; \qquad c \cong 0.2; \qquad s \cong 1500; \tag{21}$$

where s is a random variable across assaults. But this means that the

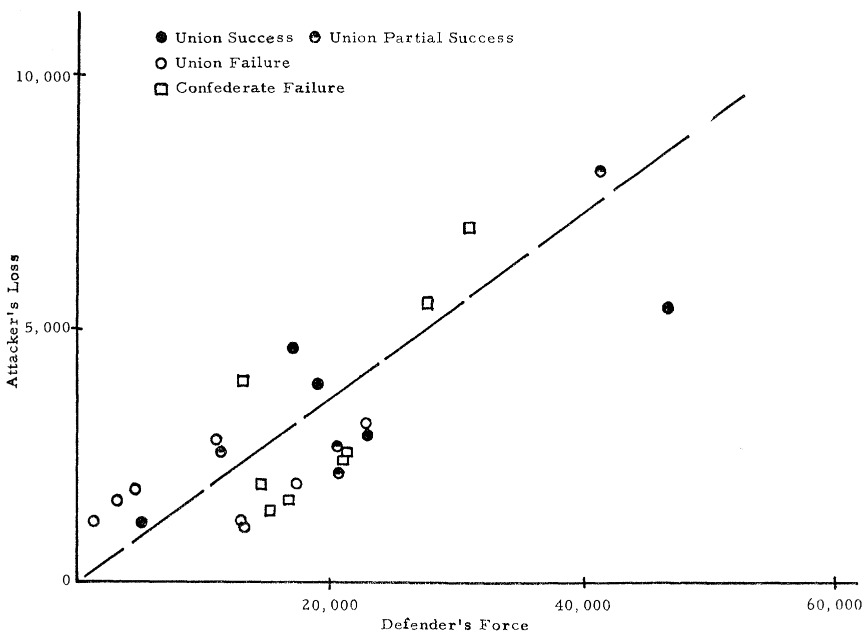

Fig. 14. Assaults on fortified lines—attacker's loss vs defender's force.

attacker's fractional loss has greater variability in small battles than large battles. (It could also reflect, in part, the uncertainties in all casualty data.)

Dividing the battles for which force ratio was less than 0.80 into two classes, according to magnitude of forces involved, we find that in those battles where the total force was less than 40,000 men, the attacker won three out of eight; when the force exceeded 40,000, the attacker won seven out of ten.

To summarize: in assaults on fortified positions, it was necessary for the attacker to have a force ratio of at least 1.25 to 1 to be successful; in the two cases where the attacker's force ratio exceeded 3 to 1, he was successful. When the total forces involved exceeded 40,000 men and force ratio was at least 1.25 to 1, the attacker was almost twice as likely to be successful as when the total force involved was less than 40,000.

THE STABILIZING EFFECT OF LARGE NUMBERS

A BRIEF review of the casualty ratios for meeting engagements indicates that the ratio of Union/Confederate losses showed much greater variability when the total losses were small than when the total was large. This is as indicated above for assaults. It has important implications for the prediction of outcomes of battles by mathematical modeling. Probabilistic formulations of Lanchester's equations are common, but their predicted outcome approaches certainty for a few dozens of troops involved. It seems that the proper way to introduce probability is not as the probability (generally assumed constant) that a particular man is killed, but rather in a variability of the exchange ratio as determined by variations in terrain, weapons, and movement, and the initiative and skill of the commanders in various parts of a battlefield, each of which contains part of the action. In a very large battle, such factors tend to even out, as evidenced by the stability of the exchange ratio in the very large Civil War battles. In small battles, it is possible for a small force to defeat one much larger; in large battles, chance works to the gambler's ruin.

It is likely that further study of the present data, separating large and small battles, will further clarify relations and lead to improved models.

A MODEL FOR AN ASSAULT ON A FORTIFIED POSITION

WE NOW construct a model of the assault, having two phases. In Phase I, the defender's losses are comparatively small; the attacker, however, loses Δx_{al} men. As a result of this loss, he may not succeed in overrunning the position. The probability that he will abandon the attack is, from equation (20)

$$\exp - k_a(\Delta x_{al}/x_a)^3, \tag{22}$$

where x_a is attacker's initial force size and $k_a = 300$, from (20).

In Phase II, the attacker succeeds in closing with the defender, and it is assumed that the incremental loss rates on both sides are now equal. At a particular time t during Phase II, the total attacker's losses are, where Δx_{dt} is defender's loss to t,

$$\Delta x_{at} = \Delta x_{al} + \Delta x_{dt}. \tag{23}$$

So that the attacker's loss fraction at t is

$$g_{at} = g_{al} + \eta g_{dt}, \tag{24}$$

where g_{al} is the attacker's fractional loss in the attack,

$$g_{dt} = \Delta x_{dt}/x_d; \qquad \eta = x_d/x_a, \tag{25}$$

and x_d is defender's initial force size.

Applying equation (6) for the probability that the defender loses

$$Q_d = -\int_0^1 e^{-k_a g_{at}^3}[\partial(e^{-k_d g_{dt}^3})/\partial g_{dt}]\, dg_{dt}. \tag{26}$$

The integral can be evaluated approximately by replacing the upper limit by infinity (acceptable because of the very rapid decay of the exponentials), expanding $\exp(-kg_{at}^3)$ in a series about the point

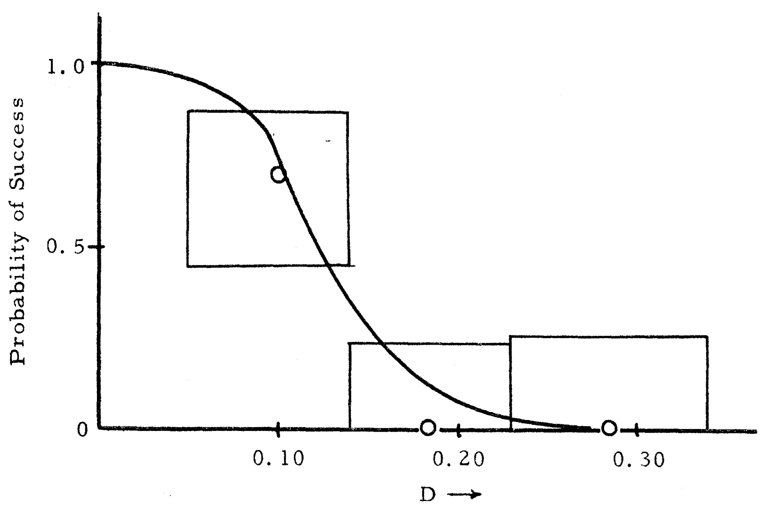

Fig. 15. Probability of success in an assault on fortified lines (18 battles).

$g_d^* = [2/(3k_d)]^{1/3}$ and integrating term by term. The first term is, assuming $k_a = k_d = 300$,

$$Q_d = \exp - 300(g_{al} + 0.13\eta)^3. \tag{27}$$

We do not know g_{al} for individual battles. We may, however, infer it for each battle by inserting the recorded total fractional loss rates f_a, f_d into equation (24). We expect, then, that the expression

$$\begin{aligned} D &= g_{al} + 0.13\eta, \\ D &= f_a - \eta f d + 0.13\eta, \end{aligned} \tag{28}$$

will be a discriminant of success or failure in an assault on fortified lines. Computation of D for the eighteen assaults for which data is available reveals that this is indeed the case. For the five assaults (all Union) for which $D<0.14$ there were three successes and one partial success. For the thirteen assaults for which $D>0.14$ (both Union and Confederate) there were *no* successes. Taking the probability of a success as

$$P_a = e^{-kD^3}, \tag{29}$$

these averages are compared with equation (29) in Fig. 15.

The losses sustained by the attacker in each assault will depend on the strength of the fortifications which, as Fig. 14 implies, is variable. We have in equation (27), however, an expression for estimating the probability of success in an assault, given an estimate of the fractional loss sustained by the attacker in overrunning the position.

CONCLUSIONS

The following general conclusions may be drawn with regard to this examination of forces, force ratios, and loss ratios in the Civil War.

1. Total losses on both sides cumulated at a fairly uniform rate after the first half year of active hostilities.

2. On the average, the Confederate forces were able to secure a much more favorable local force ratio in individual battles than the over-all sizes of the respective armies would suggest. This battle force ratio deteriorated as the war progressed.

3. Casualties on both sides were remarkably equal, both on the whole and in battles other than assaults on fortified lines.

4. In battles other than attacks on fortified lines the casualty ratios appeared to be independent of force ratios. The probability of winning was clearly a function of the force ratio in these cases, a 2:1 advantage corresponding to about an 0.87 probability of winning. Casualties on both sides were equal within a factor of two. As a result, the winner tended to have smaller per cent casualties than the loser.

5. The larger the battles, measured in total casualties, the lower the variability in casualty ratios observed across many battles.

6. In assaults on fortified lines, the losses of the attacker were proportional to the force size of the defender. In meeting engagements, battles of all sizes included in the class, casualty ratio showed no dependence on force ratio.

7. In attacks on fortified lines, there was great variability in the casualty ratio. The probability of a successful attack increased with the ratio of attacker's force to that of the defender. With a favorable force ratio, the probability of a successful attack increased with the size of the total force involved. The principal determinant of success was the fractional loss sustained by the attacker.

SUGGESTIONS FOR FURTHER RESEARCH

The possibilities of the data here analyzed have not been exhausted with regard to meaningful relations. A few topics for further study that turned up in the course of the present effort, and of which time has prevented exploration, are listed below in hopes that other investigators may find them of sufficient interest to pursue.

1. *Exchange ratio vs. force composition:* The few meeting engagements for which Bodart gives the number of artillery pieces on both sides show a tendency for casualty ratio to be proportional to the ratio of number of artillery pieces. If true, this would allow model improvement.

2. *High casualty fractions on both sides at unity force ratio:* There was an indication that when the forces were nearly equal, combat continued to substantially higher casualty fractions on both sides than is indicated by equation (19). If true, this suggests possible dependence of decision to continue fighting on relative losses as well as own losses.

3. *Different k factors for assaults and meeting engagements:* A more general model is desired to avoid the necessity for using $k=150$ in meeting engagements and $k=300$ in assaults. A more careful examination of a possible interrelation between k and μ might allow a reconciliation.

4. *Large battles vs. small battles:* Division of the class of 'meeting engagements' into one containing battles involving at least 15,000 troops on each side and one containing smaller battles indicated a possibly significant phenomenon with regard to the relation of casualty ratio to force ratio. For the small battles, high force ratios appeared to be associated with low casualty ratios (Lanchester's Square Law). For the large battles, high force ratios appeared to be associated with high casualty ratios (a side tends to lose men in proportion to the number committed, regardless of enemy strength). When all battles are lumped together, the apparently uniform scatter of Fig. 7 results.

The phenomenon of losses increasing with force committed was observed by Richard H. Peterson at the Army Ballistic Research Laboratories, in about 1950, in a study of tank battles. It was again observed by Willard,[5] and the present author has noted its appearance in the Battle of Britain data.[20] It does not follow from either the linear or square Lanchester 'laws.' As mentioned by Weiss,[20] it can be explained by recognizing that individual units of an apparently homogeneous force are not equal in capability and that any organization becomes less efficient as it becomes larger. It is suggested that the vulnerability of a force as a target increases directly with its number, but that its effectiveness in delivering fire power increases less rapidly than the number of units comprising it. A differential equation for the effectiveness of an N force against an M force might then be

$$dM/dt=-cM\log_e N.$$

Peterson called this the 'logarithmic law.' It is hoped to pursue this development in a later paper.

5. *Command structure and size of battle:* A large force has more redundancy in its command structure than a small force. Hence, there should be less disper-

sion in the casualty rate to which a large force can fight than in the case of a small force. This effect is susceptible to modeling and may be used to interpret equation (16).

A war like the U.S. Civil War may never again be fought. The connection between the relations noted above and battle with nuclear weapons is tenuous indeed. But nonnuclear conflict still continues. It is hoped that military operations analysts will find sponsors to support the development of Vagts' science of 'bellometrics' by examination of real data from World War II, Korea, Vietnam, and other contemporary conflicts, and that such studies will be directed to deriving more fundamental combat relations than the usual probabilities of hitting and destruction of specific targets. At best, such work will provide substantiation for the analytic representation of the Principles of War and indicate the prospects and uncertainties of such tools. At worst, it will reveal that there is no predictable pattern to modern warfare and free the analysts for other work. Based on the analysis presented herewith, it appears that there is scope for analysis and that battle data can furnish a justification for mathematical modeling—that it is, in fact, indispensable for the justification of such models.

REFERENCES

1. F. W. Lanchester, *Aircraft in Warfare: The Dawn of the Fourth Arm*, Constable, London, 1916.
2. J. H. Engel, "A Verification of Lanchester's Law," *Opns. Res.* **2,** 163–171 (1954).
3. H. K. Weiss, "Lanchester-Type Models of Warfare," *Proc. First Internat. Conf. on Operational Res.* 82–99 (December 1957).
4. R. L. Helmbold, "Lanchester Parameters for Some Battles of the Last Two Hundred Years," Combat Operations Research Group, CORG-SP-122 (1961), cited in reference 5.
5. D. Willard, "Lanchester as Force in History: An Analysis of Land Battles of the Years 1618–1905," Research Analysis Corp., Technical Paper RAC-TP-74 (November 1962).
6. L. Dolánsky, "Present State of the Lanchester Theory of Combat," *Opns. Res.* **12,** 344–358 (1964).
7. Francis Xavier Kane, "Security Is Too Important to be Left to Computers," *Fortune* (April 1964).
8. F. Phister, *Statistical Record of the Armies of the United States*, J. Brussel, pub., The Blue & The Gray Press, distributors, New York.
9. T. L. Livermore, *Numbers and Losses in the Civil War in America, 1861–65*, Civil War Centennial Series, Indiana University Press, Bloomington (1957). (First published in 1900.)
10. G. Bodart, *Militär-historisches Kreigs-Lexikon (1618–1905)*, C. W. Stern, pub., Wien und Leipzig, 1909.
11. *The Army Almanac*, U.S. Superintendent of Documents, Washington (1950).

12. F. B. Heitman, *Historical Register and Dictionary of the U.S. Army*, published in 1903, cited in reference 11.
13. *The Army Almanac*, Stackpole, Harrisburg, Pa., 1959.
14. Excellent summaries of most of the major battles of the Civil War, with background, maps, and photographs, are available in the *Historical Handbook Series* of the National Park Service, available from the Superintendent of Documents, Washington, D.C..
15. J. B. Mitchell, *Decisive Battles of the Civil War*, Putnam's, New York, 1955.
16. "A Capsule History of the War," *Life* (23 January 1961).
17. *Facts About the Civil War*, The Civil War Centennial Commission, Washington, D.C., 1960.
18. A. Vagts, "Battle and Other Combatant Casualties in the Second World War," *J. Politics*, Part I, **7**, 256–294 (1945); Part II, **7**, 411–438 (1945).
19. V. J. Esposito, *The West Point Atlas of American Wars, Vol. I, 1689–1900*, Praeger, New York, 1964.
20. H. K. Weiss, "Influence of the Guidance Designer on Warfare," chapter in *Air, Space, and Instruments, Draper Anniversary Volume*, Sidney Lees (ed.), McGraw-Hill, New York, 1963.