SOLDIER CAPABILITY - ARMY COMBAT EFFECTIVENESS (SCACE)
HISTORICAL COMBAT DATA AND ANALYSIS

ACN 64024

Trevor N. Dupuy and Gay Hammerman

Prepared for
US Army Soldier Support Center

Prepared under Contract DAAG29-76-D-0100,
Battelle Columbus Laboratories, for Delivery Order No. 1731
December 1980

HISTORICAL EVALUATION AND RESEARCH ORGANIZATION
8316 Arlington Blvd.
Fairfax, VA 22031

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ACN 64024	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Soldier Capability - Army Combat Effectiveness (SCACE): Historical Combat Data and Analysis		5. TYPE OF REPORT & PERIOD COVERED Topical October-December 1980
		6. PERFORMING ORG. REPORT NUMBER 2756
7. AUTHOR(s) Trevor N. Dupuy and Gay M. Hammerman		8. CONTRACT OR GRANT NUMBER(s) DAAG29-76-D-0100
9. PERFORMING ORGANIZATION NAME AND ADDRESS Historical Evaluation & Research Organization A Division of T.N. Dupuy Associates, Inc. P.O. Box 157, Dunn Loring, Virginia 22027		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Mr. R.K. Thatcher, Program Manager Battelle Columbus Laboratory, Durham Operations 200 Park Drive, Research Triangle Park, N.C. 27709		12. REPORT DATE December 1980 13. NUMBER OF PAGES 58
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Mr. Juri Toomepuu, Senior Management Analyst US Army Administration Center Fort Benjamin Harrison, Indiana 46216		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Troop quality, troop capability, combat effectiveness of units, human factors, war games

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This study included a survey of historical literature related to troop quality and capability; a survey of psychological literature dealing with the relationship of individual aptitude to group performance, with special attention to the group performance of low-aptitude individuals; a statistical comparison of quantified ratings of the combat effectiveness of 17 national military forces since 1945 with demographic, educational, and other statistical characteristics of the nations involved; and a survey of collections of combat data for the purpose of identifying units with unusually high combat effective-

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ness, and also of identifying detailed combat data on relatively small units (regiments and battalions) that would make possible quantitative analysis of unit combat effectiveness at those levels.

The study report includes recommendations for further research in the quantification of troop capability and its effect on combat. A fully annotated bibliography accompanies the report.

This report was prepared by Trevor N. Dupuy and Gay Hammerman, with the research assistance of Dee Allyson Horne. It was reviewed by Grace P. Hayes, HERO Director of Research. Abraham Wolf and Angus M. Fraser participated in the study as consultants.

SOLDIER CAPABILITY - ARMY COMBAT EFFECTIVENESS
HISTORICAL COMBAT DATA AND ANALYSIS

Table of Contents

Executive Summary. iv

Report 1

Introduction 1

Research and Analysis 1

Conclusions. 8

Recommendations for Follow-on Research. 10

Bibliography 17

Appendix A: Individual Capability and Group Performance. A-1

Appendix B: Combat Effectiveness and Characteristics of Society . . B-1

EXECUTIVE SUMMARY

This study included a survey of historical literature related to troop quality and capability; a survey of psychological literature dealing with the relationship of individual aptitude to group performance, with special attention to the group performance of low-aptitude individuals; a statistical comparison of quantified ratings of the combat effectiveness of 17 national military forces since 1945 with demographic, educational, and other statistical characteristics of the nations involved; and a survey of collections of combat data for the purpose of identifying units with unusually high combat effectiveness, and also of identifying detailed combat data on relatively small units (regiments and battalions) that would make possible quantitative analysis of unit combat effectiveness at those levels.

The study concluded that superior leadership and training are required to compensate for low troop quality, with leadership of prime importance; that there is a strong statistical association between the combat effectiveness of armies during the past 40 years and national characteristics of male literacy, household size (negative), birth rate (negative), and temperature of the capital city in the hottest month (negative); that the high level of correlation found indicates the probability that a formula could be derived to estimate the effectiveness of any two forces in a war game; and that detailed data on combat experience is available for units of exceptionally high combat effectiveness and for both sides in combat between battalions and regiments.

The study report includes recommendations for further research in the quantification of troop capability and its effect on combat. A fully annotated bibliography accompanies the report.

SOLDIER CAPABILITY - ARMY COMBAT EFFECTIVENESS (SCACE)
HISTORICAL COMBAT DATA AND ANALYSIS

Introduction

In fulfillment of the objectives of the SCACE Study, HERO was asked to provide a literature overview of historical and other related materials, to conduct such relevant historical analysis as was possible within the brief assigned time frame (45 days), and to identify data gaps and recommend follow-on studies for which a need was found.

For purposes of HERO's work, the following definitions have been established, and an effort has been made to use the terms thus defined consistently in this report:

- Quality is the basic aptitude the soldier brings with him when he enters the Army. In the US Army today it is measured by the amount of prior schooling and scores on standard aptitude tests.
- Capability is the skills and aptitudes the soldier brings into battle, after quality has been tempered by training and/or experience.
- Combat effectiveness is the capability of units. It is in part dependent upon troop capability, but is also dependent upon leadership, tactics, and other factors.

Research and Analysis

The following research efforts were carried out:

1. Literature overview
   a. Historical materials

Following consultation with staff military historians, the following categories of relevant literature were selected for investigation:

- classical military studies dealing with troop combat performance
- memoirs of commanders known to have treated the subjects of training and combat performance

- accounts of combat from the point of view of the individual soldier that might be expected to provide insights on troop quality and troop performance
- materials dealing with the performance of troops regarded as of very high or very low quality
- analytic accounts relating training to combat performance

In carrying out this research, the US Army Military History Research Collection at Carlisle Barracks, Pennsylvania, and the guidance of its staff historians were used, as was the Army Library collection at the Pentagon; a computer-assisted search of Library of Congress holdings was also carried out.

The literature examined, which stretches from the early 16th Century (Niccolo Machiavelli*) to the most recent Arab-Israeli war (Luttwak and Horowitz), produced a clear consensus on a number of points related to troop behavior:

- leadership is crucial to combat success
- unit cohesion and loyalty are crucial to combat success
- unit training under realistic conditions, and/or combat experience, is extremely important to combat success
- discipline and drill are valuable in forming capable soldiers and cohesive units
- the factors listed above can outweigh opposing superior numbers
- panic in combat is a function of the group environment rather than of the individual's personal qualities, but the action of a few individuals can start or stop it.

Very little is said specifically about troop quality or capability in this literature. It may fairly be stated that most of the writers reviewed assumed a normal distribution of inherent quality as a given for combat troops, and also assumedthe desirability of at least basic general and military education. For works that deal with quality more explicitly than most, see Baynes, Chuikov, Machiavelli, Marshall 1956, and Truscott.

---

* Names inserted in the text within parentheses refer to entries in the attached bibliography.

Since it appeared that a reasonably normal distribution of troop quality was an assumption for most authors, HERO sought extreme cases in which a force of acknowledged low-capability troops fought another representing a national demographic cross-section; HERO also sought examples of individual units regarded as made up of elite troops or very poor troops in order to compare their combat performance with their assumed quality. A number of examples of both were identified. For example, most armies during the Thirty Years' War of the 17th Century were composed of either mercenaries or impressed troops, or both, units made up of what is usually referred to as the "dregs of society," while the highly successful Swedish army of Gustavus Adolphus, in the same wars, was a truly national army.

Another area in which extreme cases were sought was that of prolonged and especially deadly wars that severely drain manpower. Cases identified were the case of Paraguay in the Lopez War of the mid-19th Century, of Germany near the end of World War II, and of both Union and Confederate armies near the end of the Civil War. There was not time within the limitations of the present study to develop appropriate bibliography for these cases, much less to conduct research, but further study might well prove rewarding.

Another, probably more promising, area for exploring the impact of troop quality and capability appeared from the literature review. The Israel Defence Forces induct almost all Israeli young men (and many women), and serve an important educational role for those without elementary education, especially Israelis of non-Western origin. Luttwak and Horowitz, Appendix 5, includes some relevant material on troop capability problems. For example, the IDF discovered that providing concentrated elementary-school education in the first months of military service does not improve the chances of military success for those who take it, and consequently reverted to its earlier practice of giving this training at the end of the three-year period of service. Further research in IDF reports and through interviews would appear desirable.

b. Results of training and unit readiness evaluations and weapons tests

A preliminary survey of material in this area, consultation with Army Research Institute personnel engaged in evaluation of training and

proficiency, consultation with the COTR for this study, and examination of the extensive bibliography for the SCACE study, prepared by the COTR, confirmed that the COTR had investigated all readily available material in this area. It appeared that a study of the relationship between success of individual soldiers in field training exercises on the one hand, and the quality of individual soldiers as measured at induction and their capability as measured by job-performance tests (SQT scores) would be highly desirable. Such a study should help validate SQT scores and make a contribution toward establishing a predictor for individual and crew performance in combat.

c. Theoretical data pertaining to the objective of the SCACE Study

A research psychologist was given the task of providing an overview of relevant theoretical material relating individual performance to aggregate group performance. He conducted a preliminary survey of the literature and found:

- The most relevant research is military related. It was determined that this research has been fully explored by the COTR.
- A search of the massive literature on the relationship of individual to small-group (crew-size) performance would require more time than was allocated to this investigator under the contract. It also appeared doubtful that such an effort would yield significant amounts of relevant material.
- In general, research studies in this field tend to use college students as subjects, so that the question of a high proportion of low-aptitude (low-quality) subjects (as measured by standard tests or educational level) is not dealt with.
- Industrial studies, which deal with the relationship between training and performance and might be expected to be useful, are not a promising source, because they rarely include a sizable proportion of low-aptitude personnel.

It was therefore decided, in consultation with the COTR, that the most useful theoretical contribution for this task would be an overview of the relationship of cognitive skills to training and performance, with special reference to man-machine systems, and the problems that still remain to be solved in this area. A summary of this work is attached as Appendix A.

2. Assessment of correlation between historical data and data on combat effectiveness of soldiers, and relevant analysis

a. Preliminary investigation of the relationship between national demographic, educational, economic, political, military, and climatic statistics and troop capability

This task was designed as a first step toward establishing a possible link between troop quality and the performance of units in combat. The first exercise carried out in fulfillment of this task was the scaling (ranking) of 17 armies of the past 40 years as to the combat effectiveness of their forces in battle.

Ideally, it would have been desirable to scale the average capability of the individual soldiers, since troop capability is the independent variable of this study. However, there is no way that troop capability can be seen in isolation when one examines the performance of soldiers at the national-force level. Performance at this level inevitably reflects additional factors, including especially quality of leadership. Combat effectiveness presumably includes troop capability, and is the only way we now have of examining troop capability at the national force level. Therefore, these 17 national forces (armies) were scaled according to the quality of their combat effectiveness.

The scaling described above was done by a group of highly qualified military historians, using the method of paired comparisons, a simple, but widely accepted, and mathematically rigorous, procedure. The resulting scalings of armies were then compared with a group of national demographic, economic, and other statistics in an effort to find possible associations between national armed force capability and national statistics. Appendix B presents a discussion of the results, together with tables showing the scalings and associations. Some of the conclusions of these exercises are these:

- The military historians who ranked the 17 armies judged the Germans in 1943-44 to have had the most effective forces, with the Israelis during the period 1967-73 ranked second. The least effective forces were those of the Iraqis and Syrians during the latter period.
- US forces during the period 1966-70 were evaluated as somewhat lower than US troops of 1943-44 and 1951, but higher than Japanese or

Soviet troops of 1944. (No effort was made, of course, to rate US forces of today, since judgments were made only on the basis of combat performance.)

- Two characteristics related to demography were very strongly negatively associated with national combat effectiveness. These were
  - -- household size
  - -- birth rate
- One characteristic related to education was very strongly positively associated with combat effectiveness:
  - -- male literacy
- One characteristic related to climate was strongly negatively associated with combat effectiveness. This was
  - -- temperature in the hottest month
- A number of other factors tested were associated with combat effectiveness less strongly or not at all. (See Appendix B for details.)
- The strong associations noted above indicate that it may be possible to derive a formula for estimating the combat effectiveness (including troop capability) of any given national force being analyzed or war gamed.

As part of this same task, a review of the findings of HERO's Quantified Judgment Model (QJM) on the combat effectiveness of national forces was carried out. Analysis of a large number of combat engagements -- analysis carried out during the past 10 years -- has shown that there is a quantifiable factor in addition to numbers of men and firepower that helps determine the outcome of battles, and that this factor differs from division to division and from national force to national force. This factor is termed by HERO the relative combat effectiveness value (CEV). The CEV is assumed to include troop capability, along with leadership training, and tactics. Thus far it has not been possible definitively to isolate any of these subfactors, including troop quality, from the total CEV, but the unquestionable existence of CEVs does show objectively that numbers and weapons alone do not determine battles. To put it another way, it can be readily demonstrated that numbers alone do not win battles, and CEV is the term used in the Quantified Judgment Model for the quantifiable qualitative difference between two forces of equal size and weaponry.

When the scale of force effectiveness derived from the judgment of military historians, which has been described above, was compared with a scale based on the CEVs of the national forces under consideration, the

two scales agreed closely.

b. Review of data in HERO Engagement Data Base for relevant material

A search was made in the HERO Engagement Data Base for units that had unusually high CEVs, in the hope that demographic and other relevant data about the troops composing these units might help establish a link between individual capability and unit performance. The division identifed as having a particularly high CEV is the 88th Division, in its performance against German units in Italy in 1944. It was not possible within the time frame of this study to carry out any research on the relationship of the relative combat effectiveness of divisions and demographic or other statistics, but follow-on research would appear promising. See Recommendations for Follow-on Research, below, for a fuller discussion.

c. Search for detailed combat data on US regiments and battalions

In pursuit of this task, a very promising source of data was discovered in the US Army Military History Research Collection at Carlisle Barracks, Pennsylvania. This is a collection of raw combat data compiled by the US 2d Division, an outstandingly successful World War I combat division. The 2d Division, following the war, collected the war diaries of all the German units that had opposed it in battle in 1918 and also compiled its own war diaries and those of all its subordinate units. This data will make possible the kind of small-unit combat analysis for which this task of the current study sought data.

In addition, there are some groups of information on US regiments and battalions in the Federal Records Center. The chief difficulty in determining the combat effectiveness of these units, and thus providing a link between individual performance and division-and-higher performance, is the task of reconstructing the combat data for opposing forces. This work would require a large investment of time, but it is believed to be feasible, and could have extremely useful results.

Conclusions

The following conclusions emerged from a consideration of the military historical literature reviewed for this project, the preliminary statistical analysis described above, and HERO's past work on the relative combat effectiveness of large units and national forces:

- There is almost certainly some tradeoff between the factors that compose combat effectiveness; that is, leadership and training almost certainly can compensate to some extent for low troop quality. However, it is obvious that superior leadership and training would be required if quality were low, not just average leadership and training.
- There also seems to be some synergistic effect among troop capability, training, and leadership. Good commanders raise the capability of their troops, and with good troops a commander performs better. Leadership is especially important, because of its impact on training, and the impact, in turn, of training on troop capability. This does not mean that low troop capability does not have a degrading effect on combat performance, and it does not mean that a predictor for degraded combat performance cannot be found, but it does complicate the task.
- National combat effectiveness appears to change very slowly over time. For example, despite the Bolshevik revolution and many years, Soviet CEVs with respect to the Germans were only slightly higher than those for Russia in World War I. German CEVs relative to the Western Allies for World Wars I and II are also close. These facts indicate that while a country may expect to coast for some time on the intangibles of troop quality, leadership, discipline, training, and tactics, a high level of combat effectiveness, once lost, may be hard to restore.
- Preliminary exploration of the relationship between the combat effectiveness of armies during the past 40 years and a variety of national demographic, economic, and other factors indicates a strong association between a number of these factors -- male literacy, household size (negative), birth rate (negative), and temperature of capital city in the hottest month (negative) -- with national combat effectiveness. Two points may be made about this

preliminary finding. One is that some of the statistics that characterize an industrialized society are also associated with high troop quality in the middle-to-late 20th Century. Another is that one of these factors, male literacy, is one that can be identified in the individual soldier and thus controlled for the army. Also, it is possible that it could become the basis for a predictor of army combat effectiveness. It is important to understand, however, that the findings so far do not give any reason to believe that increasing literacy in a given army will significantly increase combat effectiveness.

- It seems possible that a national army that represents an economic, social, and educational cross-section of the nation will generally be more combat effective than one that includes low-capability individuals in sharp disproportion to the general population. This appears to be a hypothesis worth testing. Both volunteer and conscript armies vary in their troop quality, with no clear pattern emerging, but there is some reason to believe that cross-section armies are of higher quality. The German armies of World Wars I and II, the Israeli army of the present, and the US and British armies of World War II, almost certainly the most combat effective armies of the 20th Century, have all been cross-section, universal-military-service armies. There is some evidence that the US Army that fought in Korea, which (for a peacetime army) included a very high proportion of the US population, was of somewhat higher capability than the less representative, although also conscripted, army that fought in Vietnam. S.L.A. Marshall believed it to be better in some ways than the US Army of World War II (Marshall 1956), an observation which does nothing to support (or deny) the present hypothesis but does indicate how fine the army in Korea was. (See Moskos 1980 for an excellent discussion of the strengths of a representative army.)
- The work carried out thus far has yielded no method for quantifying the impact of troop quality or capability upon weapon effectiveness, on the basis of historical data and analysis, but several research tasks have been defined that offer good probability of contributing to this quantification.

Recommendations for Follow-on Research

The SCACE Study seeks to clarify and quantify the links between the quality or capability of the individual soldier and the effectiveness of the weapons system, small unit, larger unit, and force. Some of these links are considerably clearer than others. In particular, there is a good deal of data on the link between the capability of the soldier before combat and his performance in combat. There are also some well-supported hypotheses about the quantitative role of human factors, or relative combat effectiveness values (CEVs) at the division and force level, based on HERO's past work in this field. The chief need, therefore, is more quantifiable information about the role of human factors at the small-unit level and up through the regimental level, and a methodology for relating these quantities to each other.

Strongly Recommended Research

HERO has carefully considered a number of possible follow-on studies suggested by the present work, and has selected three to recommend strongly, on the basis of their contributions toward filling these identified needs, as well as of the availability of data.

A. A Comparison of the Combat Effectiveness of Selected World War I and World War II Divisions

HERO's QJM methodology for assessing, consistently and reliably, the relative combat effectiveness of opposing forces in battle has shown that in the Italian Campaign of 1944 the relatively unsung 88th Infantry Division (a National Army organization) consistently performed better than any other American division involved in the campaign. Interesting corroboration of this quantitative analytical assessment comes from the war diary of the German Tenth Army, in which the 88th Division is referred to as "shock troops"; the record further shows that whenever the 88th Dvision was committed to combat the Germans shifted reserves to that sector of the front. The extent to which the superior performance of the 88th was due to better leadership, better training, better quality of manpower, or some other reason or reasons cannot be determined without further research. With that

research it may be possible to tell whether or not superior manpower quality was a contributor to the success of the 88th Division.

There are other possibilities for assessment of relative combat effectiveness of American divisions in Italy and Northwest Furope in World War II, divisions for which HERO has already assembled some data, including the 3d, 34th, 36th, and 45th Infantry Divisions and 1st Armored Division in Italy, and the 1st, 2d, 29th, 26th, 35th, and 90th Infantry Divisions and the 4th, 5th, and 6th Armored Divisions in Northwest Europe.

In World War I, the US 1st and 2d Infantry Divisions have consistently been judged by informed military opinion to be almost certainly the best divisions in General Pershing's AEF. Statistics on casualties, ground gained and prisoners taken, and other measures of combat effectiveness compiled by the Statistical Division, War Department General Staff, under the direction of Col. Leonard P. Ayers, appear to confirm this assessment.

A more reliable and precise method of confirming the superior combat effectiveness of the 2d Division is now possible. In the course of the present study, HERO discovered a collection of raw combat data on that division that will make possible an analysis of the 20 or more engagements of the 2d Division in World War II by HERO's QJM methodology. The data collection consists of the war diaries of all the German units that opposed the 2d Division, together with its own written combat orders and those of its subordinate units. Carrying out this analysis will quantify the almost certain combat effectiveness superiority of the 2d Division.

With the high standard of combat effectiveness of two divisions from two separate wars established and quantified, the reasons for this superiority, in comparison to performance of other divisions engaged at the same time, can be sought, and the possible role of troop quality and capability, and of various demographic and other national statistical characteristics, as predictors of combat effectiveness can be assessed.

The objective of this follow-on study would be to use the promising data available on these two divisions to seek a precombat predictor of combat effectiveness. This study is very strongly recommended.

B. Small-Unit Combat Effectiveness Assessments

As noted above, one of the greatest needs in SCACE Study research is for quantified data on unit performance in small-unit actions. This kind

of information has been difficult to obtain, since combat records at levels below the division are rarely complete. If complete records were available, it would be possible to carry out QJM analyses for small units. It might then be possible to relate individual soldier capability to small-unit performance, and also to develop aggregation factors relating small-unit performance to the performance of divisions.

A very promising resource is now available for carrying out this analysis. The 2d Division collection of records, compiled after World War I, includes not only all written division combat orders and the war diaries for units opposing the division, but also the orders for subordinate units of the division. This will make possible low-level QJM assessments, and since the corresponding division records are also available, comparisons can be made and aggregation factors derived.

Since preliminary research has indicated that low-level data of quality adequate for QJM analysis can also be found -- with difficulty -- in the World War II records, the relevance of World War I experience to that of World War II can be tested, and at the same time the nature of trends in low-level combat can be assessed.

This research could be carried out in two stages:

- The QJM analysis of the World War I performance of small-unit components of the 2d Division, and a comparison of the results with findings for the performance of the division.
- The compilation of combat data for selected small units in World War II; the reconstruction of data for enemy forces facing them; the QJM analysis of the results; and the comparison of QJM findings for these units and for the divisions of which they are components with those for the World War I components of the 2d Division, and analysis.

The objective of this study would be to quantify the relationship between small-unit performance and division-level performance. The study is strongly recommended.

C. National Manpower Quality and Combat Effectiveness

In the present study, a beginning has been made in exploring the relationship between national demographic, economic, and other statistics,

on the one hand, and the combat effectiveness of armies, on the other. (See Appendix B.)

In another approach to the question of comparative combat effectiveness, HERO has made consistent, and demonstrably reliable, comparisons of the combat effectiveness of opposing national armies in World War I, World War II, and the Arab-Israeli wars. It is clear that to a large extent the differences in national combat effectiveness from one country to another in these wars are a reflection of professionalism, leadership, and training. At the same time, it appears likely that cultural and technological differences in the societies have influenced the quality of military manpower and the comparative battlefield performance of individuals and units to some thus far unquantifiable extent.

If it should be possible to find some way to compare the quality of manpower of the various armies of these wars, and also to compare their military training systems and level of leadership, the significance of manpower quality could be assessed and at least roughly quantified.

Following are the next steps to be taken in this effort:

- Preparing plots of data already gathered and examining them for relationships that may not be detectable by the method already carried out, i.e., calculating a correlation coefficient.
- Refining the scaling of national combat effectiveness and establishing a broad consensus on the scaling by obtaining the judgments of a large number of both military historians and experienced combat troop commanders.
- Considering other possible statistical indicators of national combat effectiveness.
- Broadening the statistical base. Preparing reliable statistics for the earlier periods involved (pre-World War II) will require an appreciable investment of time.
- Examining the effect of time lags on the correlation between force effectiveness and demographic, economic, and other national characteristics. It may be that the characteristics of a society that produces effective combat forces are effective only after a delay in time.

The objective of this study would be to derive and quantify a predictor for national troop capability, or national troop combat effectiveness, from national demographic, economic, and/or other statistics. The study is strongly recommended.

Other Desirable Research

The following research projects were judged desirable but of lower priority:

D. Investigation and Analysis of Relevant Israeli Experience

The Israel Defense Forces (IDF), generally acknowledged to comprise a highly effective combat force, has a great deal of experience in providing basic education and training to its members, and in making effective use of troops with low educational levels and psychomotor skills. Some IDF reports on these subjects are known to exist, and some interviews with appropriate officers should be easy to arrange in Washington, or in Israel if necessary. The first step in such a study would be to determine whether this material has already been exploited and made readily available in this country. If it has not, this would appear an extremely useful project requiring a relatively low investment of time and money, even allowing for the need for translation.

The objective of this study would be to gather data on Israeli experience with troops lacking literacy, elementary education, and/or psychomotor skills required for many current military occupations. The feasibility portion of this study is strongly recommended, with follow-on as appears appropriate.

E. Volunteer vs. Conscript Army: the Franco-Prussian War

The question of the quality of US Army soldiers today is often related to the issue of the relative capabilities of a volunteer military force and a conscript army. Probably the closest comparison of battlefield performance of a volunteer army against a conscript army can be found in the Franco-Prussian War campaigns of 1870. The French Army was nominally conscript but actually consisted mainly of long-service volunteers; the ranks of the Prussian army were filled with short-service conscripts. In general, the French soldiers were of higher capability than the Prussians, but the Prussian noncommissioned officers and officers were superior to the French. There appears to have been no assessment of the relative performance of these armies in terms of the quality of their soldiers, or of the way in which this quality was affected by the volunteer or conscript nature of the army.

F. Comparison of Marine Corps and Army Troop Capability

The Marine Corps has consistently been somewhat more selective in recruitment than the Army, and to some extent it is regarded as an elite military organization. Comparison of the performance of comparable Marine and Army units in World Wars I and II, and in Korea, together with comparisons of their manpower quality and training and doctrine, could be expected to provide useful findings on the role of troop quality and capability in combat effectiveness. The fact that the US 2d Infantry Division in World War I -- an outstanding division for which detailed records are available -- was composed of one Marine Corps brigade and one Army brigade enhances the possibility of carrying out a significant Army-Marine Corps comparison.

G. Comparison of Troop Quality and Capability with Performance in Simulated-Combat Training Exercises

The Army Research Institute carries out research on the effectiveness of various field exercises as training devices, with quantified results. Enough information on the individual soldiers participating in these exercises is available from ARI's work to permit analyzing the relationship between soldier quality (as measured at induction), soldier capability (as measured by SQT scores), and the performance of individual soldiers and small units in simulated-combat performance. Since quantified data on these relationships is important for validating degrading factors for weapon effectiveness, this research should be carried out.

H. Historical Survey of Manpower Quality in US Army Combat and Support Units in the 20th Century

Whereas there is considerable agreement that the overall quality of manpower in the All Volunteer Army (AVA) is substantially lower than it was when Selective Service was supplying manpower by conscription, there are two almost-contradictory opinions as to the effect on the combat arms of the Army.

The first view is that the lower quality has a deleterious effect on the combat arms, for it exists at a time when modern technologically sophisticated weapons make it very difficult for low quality soldiers to operate them.

The second view is that actually the quality of manpower in the combat arms today is comparable to that in the combat arms in peace and war over the past 50 to 80 years, and that a return to conscription would not be likely to improve it. The admitted reduction in quality is held to have affected the support arms and services only.

A study that would evaluate the quality of manpower over this period, its implications in terms of combat readiness in the 1980s, and the possible effects of a return to conscription is highly recommended.

I. Differences between Elite and Poor Units

There have long been two schools of thought about the creation and use of elite units. One school contends that elite units, created either from highest quality men or from average men very highly trained, make the best use of available manpower. The other school maintains that any contribution of elite units to overall military combat effectiveness is more than offset by reducing the motivation of men in average or mediocre units.

A comparison of the relative combat effectiveness of elite and non-elite, or admittedly poor units, and of the soldiers in each, would make a direct and significant contribution to the objective of the SCACE Study. A feasibility study could rapidly determine whether sufficient data for such a study would be available.

BIBLIOGRAPHY

Introduction

As indicated in the body of the report, there is virtually no historical bibliography dealing specifically with the quality or capability of troops. The present bibliography is made up of those works that could be found that included substantial and useful references to troop quality, or capability, or to the relationship between quality, training, and combat performance. Whenever possible, specific page references to useful material have been given.

Ardant du Picq, C.J.J.J. Battle Studies: Ancient and Modern Battle. Translated by John N. Greely and Robert C. Cotton. New York: Macmillan, 1921.

This classic, whose author was killed in the Franco-Prussian war, deals at some length with combat behavior. Ardant du Picq says nothing specifically about troop quality, assuming a soldier who is "strong, apt, vigorous, trained...." (p. 99), but stresses leadership and especially unit cohesion. Ardant du Picq speaks of the discipline that is not imposed by superiors but comes from "the mutual supervision of groups of men who know each other well." (p. 96) He stresses the importance of keeping the personnel of combat groups as unchanged as possible, and thus producing "brotherhood, professional knowledge, sentiment, above all unity." (p. 96) The last chapter deals with the author's views on national characteristics of various countries and what impact he believes they have on military matters.

Baldwin, Fred Davis. The American Enlisted Man in World War I. Doctoral dissertation, Princeton University. Ann Arbor, Mich.: University Microfilms International, 1965.

This useful work gives the socio-economic background of enlisted men in World War I, describes the aptitude testing carried out and gives score results, describes training, and discusses soldier attitudes and adjustment. It does not deal with combat performance.

Baynes, John. Morale: A Study of Men and Courage: The Second Scottish Rifles at the Battle of Neuve Chapelle, 1915. London: Cassell, 1967.

This is an excellent case study on combat performance. Baynes defines morale as the soldier's determination to do his duty to the best of his ability in any circumstances, so that morale and good combat performance are almost synonymous in this work. Pride in doing a good job and in oneself are the most important elements of morale, Baynes says. He stresses the importance of training, and especially the inculcation of esprit de corps. He believes that the feeling of belonging, the trust in the group, are of the greatest importance. The unit he studied suffered 70% casualties in its five-day battle, but carried out its missions as ordered and held its position until relieved. The book includes much detail on the battalion's individual and unit training, and on the social, economic, and educational background of its officers and men.

Chuikov, Vasili I. The Battle for Stalingrad. Translated by Harold Silver. New York: Holt, Rinehart and Winston, 1964.

Chuikov, commander of the 62d Army in the defense of Stalingrad, stresses throughout his narrative the skill and courage of the individual soldier and the impact these factors had on this desperate battle. Chuikov believes that the quality of the Soviet troops, plus superior leadership and tactics, made the difference against initially superior German numbers,

and he gives numerous convincing examples. (See especially pp. 70, 97-99, 107-08, 146-47, 153-54, 157-58, 192, 276-77, and 308). The role of the individual under conditions of street fighting was apparently especially important: "It was the soldiers of the 62d Army who understood more quickly than anyone else what city fighting means, and learned more quickly and better than the enemy to make use of streets, buildings, basements, staircases, factory chimneys, and the roofs of houses." (p. 308) Chuikov attributes the soldiers' motivation to patriotism, Communist training, and the propaganda efforts of the political workers attached to the army.

Coll, Blanche D., Jean E. Keith, and Herbert H. Rosenthal. The Corps of Engineers: Troops and Equipment, pp. 355-390. US Army in World War II: The Technical Services. Washington: Government Printing Office, 1958.

A good case study on the recruiting and training of officers and enlisted personnel for highly specialized combat tasks, in this case, the conduct of amphibious landings. Vigorous and successful efforts were made to recruit men with marine and other appropriate experience for the Engineer Amphibian Command, but the command also had to absorb large numbers of low-quality personnel. Of the first 2,788 men assigned to the EAC from replacement training centers, only 49% scored at Grade III or better, whereas a normal distribution would have placed 69% in this category. (p. 366) Training programs had to be adjusted accordingly. The importance of recruiting good trainers is stressed, along with the importance of realistic training generally, and the special importance of ongoing training following combat experience.

Crane, Stephen. The Red Badge of Courage. New York: Random House (Modern Library), 1951. (First published in 1895).

Although a work of fiction written before its author had any personal experience of military combat, this book is widely praised by military men for its realistic portrayal of combat experience. Crane's sources were published and oral reminiscences of Civil War soldiers. The following characteristics of combat relevant to the current study emerge: the sense of security and identity provided by one's own unit; the nonabsolute nature of courage--a man will generally run or fight effectively depending on how the rest of his unit is behaving rather than from any innate cowardice or bravery; the importance of familiarity with combat conditions for effective fighting.

Duffy, Christopher. The Army of Frederick the Great. New York: Hippocrene Books, 1974.

Frederick's army was partly native (Pomeranians were highly regarded), and partly foreign mercenaries. Its commander emphasized the importance of drill and tactics to attain absolute obedience and loyalty. After two to three years of basic training in drill and tactics, officers and men received further training in accordance with their ability. Frederick regularly held drills and maneuvers to maintain morale, assure readiness,

and, in the case of maneuvers, to experiment with new tactics and train commanders. Frederick believed in esprit de corps, but stressed discipline above all, asserting that the officer should be feared more than the enemy.

_____. The Army of Maria-Theresa: The Armed Forces of Imperial Austria, 1740-1780. New York: Hippocrene Books, 1977.

A chapter on "The Men" (pp. 47-62) describes the recruiting practices; backgrounds of recruits, including ethnic groups, occupations, and ages; the organization of an infantry company; and the ranks and duties of non-commissioned officers. Duffy suggests that Frederick the Great's army was better on offense, Maria Theresa's better on defense. He states that Frederick's harsh discipline led to a good many Prussian desertions to Austria.

Firth, C.H. Cromwell's Army: A History of the English Soldier During the Civil Wars, the Commonwealth, and the Protectorate, 3d edition. London: Methuen; New York: Barnes & Noble, 1921.

A valuable work, based on much research in contemporary records and literature. Includes much detailed information on recruiting practices, quality of troops, and training.

Fraser, Angus M., Nehemiah Jordan, and John K. Moriarty. 1966. Notes on an Amnesty Program in Vietnam.... Institute for Defense Analyses, Economic and Political Studies Division. IDA Internal Note N-394. Arlington, Va., October 1966.

These papers, based on 59 interviews with Vietnamese who came over from the Communist forces in the Chieu Hoi (Open Arms) program, provide some useful information on motivation and morale in Communist Vietnamese forces. The role of the three-man cell in infantry units, with its family-type bonds, is mentioned.

George, Alexander L. The Chinese Communist Army in Action: The Korean War and its Aftermath. New York: Columbia University Press, 1967.

Chapters 8-10 (pp. 152-96) of this work, which is derived from extensive interviewing of captured Chinese Communist soldiers in the Korean War, explicate in detail the reactions of the indoctrinated People's Liberation Army soldier to battlefield conditions which seemed to contradict what he had been taught. While not immediately relevant to the present study, George's work provides an Asian perspective on the question of troop behavior in battle.

Gough, Hubert. Soldiering On. New York: Robert Speller and Sons, 1955.

In his memoirs, marginally useful for the present study, General Gough stresses the importance of military education and training, and also the

character of the British soldier as he sees it--marked by calm and endurance on the battlefield. He believes that loyalty to the nation and willingness to sacrifice for it is bound up with, and partly results from, the individual's loyalty to his town and region.

Hall, Grover G., Jr. 1,000 Destroyed: The Life and Times of the 4th Fighter Group. Montgomery, Ala.: Brown Printing Co., 1946.

This is unit history by a member, anecdotal, but with useful information and insights. The US 4th Fighter Group was made up of US volunteers who had joined the RAF before Pearl Harbor, many of them after being rejected by the US Army Air Corps for failing the physical examination or lack of college education, or after having washed out of pilot school.
The unit had the most successful record, in enemy planes downed, of any fighter group based in Britain. Throughout, Hall attributes the success of the group to the superb leadership of its commander, Colonel Blakeslee, and the unusually high motivation of his men. An interview with a World War II fighter pilot and military historian, conducted in the course of the present study, has confirmed the exceptional air combat leadership of Blakeslee. This historian also pointed out that another fighter group based in England, the 56th, also under exceptional leadership but made up of pilots whose backgrounds were indistinguishable from those of other US pilots, compiled a record virtually as good as the 4th's, with the two groups clearly outperforming all other fighter groups. Leadership appeared to make the difference.

Heavey, William F. Down Ramp: The Story of the Army Amphibian Engineers. Washington: Infantry Journal Press, 1947.

This book covers some of the same material as Coll and others 1958, and also describes the performance of the Special Brigades of the Engineer Amphibian Command in combat. Again, training is stressed, and described in considerable detail. The special efforts to recruit officers with sailing, power-boat, and shipbuilding experience, the careful efforts to tailor training to the specific tasks the units would perform, and the descriptions of combat results, make this material relevant for the present study.

Kippenberger, Howard. Infantry Brigadier. London: Geoffrey Cumberlege, Oxford University Press, 1949.

Major-General Kippenberger's account of commanding New Zealand infantry and armored troops in Greece, Crete, Libya, Syria, Africa, and Italy in World War II. While he does not discuss quality, as measured by education and aptitude, per se, Kippenberger does discuss soldier capability, emphasizing the importance of realistic unit training in achieving it. Kippenberger stresses the importance of battalion unit cohesion and esprit de corps. He emphasizes the roles of unit history, drill, ceremonial parades, games, and physical appearance in building soldier morale and unit cohesion, and cites examples of poor discipline, training, and morale leading to poor combat performance (pp. 48, 55, 175, 345). For force combat success, he stresses above all coordination of infantry and armored forces; for the

success of the individual units, he stresses unit cohesion and ongoing, realistic unit training. Kippenberger places tremendous empnasis on unit cohesion, loyalty to the unit (platoon, company, battalion), and loyalty to New Zealand. He stresses the value of drill discipline in building unit cohesion.

Lifton, Robert J. "Advocacy and Corruption in the Healing Professions." In The Social Psychology of Military Service, edited by Nancy L. Goldman and David R. Segal. Beverly Hills, Calif.; London: Sage Publications, 1976.

A study of the way in which attitudes and conduct were formed under the pressures of combat experience in Vietnam. Useful for the treatment of attitudes and reactions in combat-experienced men.

Luttwak, Edward, and Dan Horowitz. The Israeli Army. London: Allen Lane, 1975.

This book provides considerable information on the makeup and training of a universal-service national army generally ranked high among 20th Century armies. The esprit of the elitist Palmach during World War II (pp. 19-21) is discussed, as are the demographic makeup of the Haganah in 1947 (p. 23); decentralized command procedures (p. 54); leadership (pp. 61-62, 85-86); training, including "black-box" vs. technical-understanding approaches to high-technology training (pp. 190, 200-01); factors characteristic of Arab and Israeli societies which the authors see as greatly affecting their relative combat effectiveness (pp. 283-86); and military pros and cons of the elite youth-group-based Nahal (Appendix 1).
Appendix 5 (pp. 438-46), extracted from 1962 Israeli Defense Forces report on education in the IDF, includes much highly relevant information: The IDF is an important educational agency, especially for Israelis of oriental Jewish background, but it has been found that literacy education given at the beginning of military service does not improve the soldier's military capability; this education is therefore given at the end of the three-year term of service, with much better success and resulting benefits for the soldier in civilian life. There is a high correlation between success in the IDF and prior educational level. Although success is achieved disproportionately by the better-educated soldiers of Western origin, and although there have been charges that discrimination against oriental Israelis exists in other areas of Israeli society, almost no charges of discrimination have been made against the IDF. Military jobs requiring technical skills are most successfully filled by soldiers who score high in tests of psychotechnical aptitude, including manipulation of mechanical tools, and when these tests were changed to give more weight to the rapid-hand-movement skills of the Yemenite Israelis and less weight to tool manipulation, the tests were no longer good predictors of success in technical jobs.

Machiavelli, Niccolo. The Discourses. In The Prince and the Discourses. New York: Random House (The Modern Library), 1950.

Machiavelli, better known for his diplomatic career, had also organized, trained, and commanded a citizen army when he wrote the Discourses, and

they are packed with insights and observations on tactics, leadership, training, and discipline. Especially relevant to this study is Book 3, Chapter 13, in which Machiavelli deals with the question of whether a good commander with a feeble army or a good army with a poor commander is most likely to be successful. He quotes Caesar, who gave examples to show that neither was worth much, and then goes on to give numerous additional examples of good armies that were successful despite being poorly led, and also of outstanding commanders who achieved success with poor troops.

Marshall, S.L.A. Commentary on Infantry Operations and Weapons Usage in Korea, Winter of 1950-51. Chevy Chase, Md.: Operations Research Office, The Johns Hopkins University, 1952.

Unit cohesion, pride in unit, especially company and regiment, are stressed as factors in orderly withdrawal and recovery. Marshall found many more men firing than in World War II, with the chronic nonfirer the exception. There is nothing on quality of troops as a predictor of combat performance.

_____. Men Against Fire: The Problem of Battle Command in Future War. New York: William Morrow, 1947.

The theme of this analytic work is that it is the soldier who fights who wins battles, that fighting means using one's weapons, that what is needed for greater combat effectiveness is "more and better fire" (pp. 22-23). Marshall stresses the need for realistic combat training; the soldier's absolute need for physical support from other men in the unit; evidence that no more than one man in four actually fires his weapon in combat (in World War II) almost regardless of time duration. Which men did fire could not be predicted before combat. Neither good discipline nor drill perfection was an effective predictor (p. 60). Some fighters were good soldiers before battle, while some were not.
Combat seasoning consists of learning to do something better (e.g. fire a rifle at an appropriate time and target). Men do not become more willing, less afraid, but rather, if anything, more afraid, after combat experience.

_____. Pork Chop Hill: The American Fighting Man in Action, Korea Spring 1953. New York: William Morrow, 1956.

A battle narrative based on Marshall's after-combat interview methods. His conclusions: the US troops he observed in Korea were superior in junior officer quality and initiative of enlisted men to any he had observed earlier (World War II); the policy of frequent rotation of troops in and out of the combat zone was highly undesirable, as it robbed troops of the chance to acquire combat and deception skills, putting them at a disadvantage to the enemy. Marshall reports on the personal characteristics of 73 decorated soldiers whose superior courage and achievements he had personally observed or established through interviews with witnesses: none was an only child, and most were from families of three or more children; most volunteered positive statements about their families and home life; 52 of the 73 spoke voluntarily with warm admiration of their fathers (pp. 19-20).

_____. The River and the Gauntlet. New York: William Morrow, 1953.

An account of the November 1950 retreat in Korea, focusing on a single company of the 2d Infantry Division. This well-led, racially integrated unit (60% white, 30% black, 10% attached Korean troops with mention also of some Japanese-Americans) fought extremely well.
This is a detailed narrative account with little analysis. It includes examples of the contagiousness of fear (pp. 47-48, 74-75). There is nothing relevant to the relationship of precombat quality or capability to combat performance.

Mason, Philip. A Matter of Honour: An Account of the Indian Army, Its Officers and Men. London: Jonathan Cape, 1974.

Contains much material, especially pp. 341-61, 449-53, on recruitment of troops in British India, and the question of whether certain ethnic groups made especially good soldiers or whether effective training could make any troops effective.

Moskos, Charles C., Jr. The American Enlisted Man: The Rank and File in Today's Military. New York: Russell Sage Foundation, 1970.

Moskos describes US enlisted men's attitudes, showing that conventional military attitudes and sources of motivation are apparently no longer adequate. He suggests that the military has become increasingly isolated from the civilian population and from civilian attitudes and values. Discussion of attitudes and behavior in Vietnam is included, but there is no treatment of combat performance.

_____. "How to Save the All-Volunteer Force." The Public Interest, fall 1980, pp. 74-89.

A very useful article analyzing the current all-volunteer army and stressing the need for an army that represents a cross-section of the country in aptitude and education. Moskos urges greater education benefits (GI Bill) as the most useful incentives for enlistment by high-aptitude people. He also favors a "two-track" system, with short-term, lower-pay, high-education-incentive citizen soldiers, on the one hand, and long-term, higher-pay, career soldiers, on the other.

Remarque, Erich Maria. All Quiet on the Western Front. Translated by A.W. Wheen. Boston, Mass.: Little, Brown, 1928.

Like The Red Badge of Courage, a fictional classic that gives a realistic picture of combat at the level of the individual soldier. All Quiet provides a good supplement to the World War I material presented in Baynes's Morale. Remarque stresses the importance of training, discipline, and drill, not only for their own impact on combat performance but for building the all-important unit cohesion. The importance of keeping the same group of men together is stressed.

Saxena, K.M.L. The Military System of India, 1850-1900. New Delhi: Sterling, 1974.

This work includes much material on the ethnic, class, and caste composition and structure of the Indian Armies in British India.

Senger und Etterlin, Frido von. Neither Fear nor Hope. Translated by George Malcolm. New York: E.P. Dutton, 1964.

A highly regarded German memoir of World War II, with good material on tactics, strategy, and the relationship of political factors to military affairs, this book makes only passing references to troop quality. There are some references to declining troop quality in the last days of the war (pp. 259-60, 264).

Stouffer, Samuel A., Arthur A. Lumsdaine, Marion Harper Lumsdaine, Robin M. Williams, Jr., Brewster M. Smith, Irving L. Janis, Shirley A. Star, and Leonard S. Cottrell, Jr. The American Soldier: Combat and Its Aftermath. Studies in Social Psychology in World War II, vol. 2. Princeton, N.J.: Princeton University Press, 1949.

This work, and its three companion volumes, contain a wealth of quantified and analyzed data on combat behavior. This volume is especially relevant to the present study. One study it reports found that intelligence score, level of education, mechanical aptitude score, and positive precombat attitudes about combat all correlated strongly with good combat performance (pp. 30-41).

Sun Tzu. The Art of War. Translated by Samuel B. Griffith. London: Oxford University Press, 1963.

While this classic work may be regarded as folk wisdom, it does codify enduring Chinese military principles. There is useful material on the care of troops and means of motivating them, much of it tending to demonstrate the longevity of universal priciples. This work also contains comment on discipline, including Sun's famous story of the beheading of the emperor's favorite concubine.

Truscott, Lucian K., Jr. Command Missions. New York: E.P. Dutton, 1954.

Lieutenant General Truscott's memoirs of World War II show the importance of academic (e.g., tactical/strategic), physical (e.g., speed marching), and psychological (e.g., combat experience) training to combat performance. Truscott shows the importance of integrating men with combat experience with inexperienced men in training programs, thereby providing the knowledge gained through combat experience to training programs. He also describes "realistic" training programs and stresses the importance of combat rehearsals in assessing and adjusting training procedures (pp. 84, 88, 303).

Truscott claims that training must be continuous (pp. 85, 90, 352, 397, 456-57) and suggests that intensive training programs should be held, when possible, just prior to combat (p. 360). Truscott also shows how morale, leadership, discipline, esprit de corps, and planning play essential roles in determining the capability of the individual soldier and of the unit.

Throughout the book, Truscott stresses the importance of the capability of the individual soldier and shows how the capability of the individual often reflects the combat effectiveness of the unit (pp. 97-98). He also provides specific examples of training correlating with combat effectiveness (see pp. 81, 87, 115, 212, 214, 376, 382, 400-01, 415, 465, 474, 539-40). Truscott also shows how education and social and national attitudes influence not only basic troop quality, but also the ability to profit from training. Truscott compares the national characteristics of the American, British, and German soldier (pp. 555-56) and shows how these characteristics affect combat performance. For example, Truscott concludes that the British soldier is better at defense while the American soldier is better at offense. Truscott also provides examples of inexperienced troops with a poor level of training showing low levels of combat performance (pp. 83, 93, 178, 212).

US Army. Center of Military History. "Fact Sheet: 442d Infantry."

The 442d Infantry Regiment is of interest as a crack unit composed of highly motivated, well led, ethnically homogeneous troops that performed in combat with great distinction. The regiment, made up of Japanese-Americans, won four Presidential Unit Citations in World War II fighting in Italy and France, and is acknowledged by historians as a superb regiment, probably the most decorated in the campaigns in which it participated. It should make a good candidate for any study of elite units. This fact sheet gives the basic data on the regiment.

US Army. Historical Division, Headquarters US Army Europe. Guide to Foreign Military Studies, 1945-54: Catalog and Index. 1954.

Following World War II, a number of captured German general officers were asked by the US Army Historical Division to write reports of their wartime commands for historical purposes. Scattered through these reports is some information on troop quality and capability. A historian on the US Army Center of Military History staff has suggested that the report of General Blumentritt, Chief of Staff of OB West and commander of Army Group Blumentritt, may be an especially good source for such material. This guide indexes the German generals' work.

US National Archives. Captured German records of World War II.

A specialist on these records who was consulted about their relevance to the present study states that there is some material on troop quality and capability scattered through them, especially in records dealing with replacements. There is data on quality deterioration late in the war, and some material relating quality deterioration to operational problems.

APPENDIX A

INDIVIDUAL CAPABILITY AND GROUP PERFORMANCE, WITH SPECIAL REFERENCE TO MAN/MACHINE SYSTEMS

The research psychologist who formed part of the HERO team for this SCACE project was asked to prepare an overview of literature dealing with the relationship between individual performance and group performance. It was understood that, because of the nature of the project, the focus should be on material relevant to a military environment and military tasks. It was further understood that the project was especially concerned with the effect that the aptitude of the subjects, as measured by standardized tests, might be expected to have on the level of group performance, and that there was a special need to determine what differences might be found between the performance of groups composed of individuals representing a normal distribution of aptitudes (as measured by standardized tests) and groups composed of predominantly low-aptitude individuals.

An initial assumption of this investigator was that the effects of stress upon performance, because combat performance is performance under extreme stress, would be a significant area for investigation. Stress is known to introduce factors that can be expected to affect group performance differently from non-stress environments. Consultation with the COTR indicated, however, that it was not necessary to restrict the investigation to performance under stress, especially since the COTR had already reviewed much of this material.

The investigator then noted that group performance could be expected to differ as a function of several other variables as well. In particular, he called attention to the following variables and their joint effects:

- Task. Some tasks require coordinated performance by a number of individuals; others require isolation of carefully prepared operators; still others require neither cooperation nor isolation.
- Group organization. Some groups are loosely organized and require no central task leader or emotional leader; other groups may be extremely

hierarchical, or may allocate functions in one of many different ways.

- Group leadership. Some task leaders also attempt to function as emotional leaders; other distance themselves from that role.

The investigator believes that recent psychological literature shows that the effectiveness and morale of groups is greatly dependent on the match among these variables, as well as the talents and abilities of the group members.

The following major compendia and journals in the social sciences were reviewed for possibly relevant material:

- Journal of Applied Psychology
- Journal of Personality and Social Psychology
- Journal of Political and Military Sociology
- Sociology of Work and Occupations
- Personnel Psychology
- Advances in Experimental Social Psychology
- Handbook of Social Psychology (both editions)
- Handbook of Small Group Research
- Handbook of Group Leadership

A large number of texts, reference materials, and library catalog materials were also reviewed.

Nothing was found of particular relevance as a result of this search that had not already been considered by the COTR. A number of avenues of investigation, suggested by other HERO team members, or the COTR, and appearing to offer promise of relevant data, were not productive for a variety of reasons:

- Social psychologists in academic settings deal with extremely small groups (compared to military units), and tend to see only gifted individuals.
- Those working in military settings have contributed important work, already known to the COTR and incorporated in his work.
- Those working in industrial settings tend to deal with individuals who are relatively competent. Those who are not competent do not remain in the positions studied, given the play of the labor market.
- Training programs for the disadvantaged are almost exclusively concerned with individual skills.

Thus the investigator established that useful research in the area mapped out for his work had been well covered by the COTR, and that other, seemingly useful work, would not in fact be rewarding. In a consultation between the investigator, the COTR, and the HERO project coordinator, the investigator was asked to direct the focus of his investigation for the remainder of the project to literature on the relationship between the tested aptitude of the individual and his performance in operating complex machinery, that is, to the relationship between aptitude and man/machine systems.

As the best summary of the research on this question, and one specifically focussed on a military environment and the question of large numbers of low-aptitude personnel, the investigator has submitted the attached paper by Dr. Earl Hunt and Dr. Marcy Lansman of the University of Washington. Dr. Hunt has long been one of the foremost contributors to studies in the field of intelligence, and Dr. Lansman has done much work in the human factors field.

COMMENTS ON THE CATEGORY IV PROBLEM

Earl Hunt and Marcy Lansman
The University of Washington

Upon the initiation of the all-volunteer force, concern was expressed about the ability of the services to attract quality personnel. A target of 20% Category IV enlistees was established for the Army. Early analyses suggested that this was not a problem, and that in fact only 11% of the Army enlistees were in Category IV (Cooper, 1977). It was subsequently found that the test norms were in error, and that 45% of current enlistees fall into this category (Holden, 1980). The implications of this fact have been hotly debated. The extreme points of view seem to be that (a) the Category IV classification is only a statement about reading, which is not relevant to many Army tasks, or (b) that the Army contains many decidedly subnormal individuals, who might well be on public assistance, or even institutionalized, were they not in the Army. We believe that neither view is justified. In this report we analyze the capabilities and weaknesses that may be expected in Category IV enlistees. The report focuses upon behavior in using equipment. In preparing the report we have drawn primarily upon the general psychological literature on the relation between intelligence and social and economic performance. While a complete scholarly review has not been possible in the time available, we believe that there is considerable consistency in this literature, and that it is applicable to the military situation.

The report begins by discussing some of the general findings concerning "intelligence". The Armed Forces Qualification Test (AFQT) will be related to the major dimensions of intelligence, as defined in the psychological literature. The next section deals specifically with Category IV individuals. Here we will consider the characteristic reactions of similar individuals in civilian educational and occupational settings. The third section describes the sorts of problems that Category IV individuals are likely to encounter when they deal with military equipment. Some remarks will be made about their general reaction to work situations. The final section of the report suggests possible responses of the Army to the presence of large numbers of Category IV enlistees. This section also presents some topics for further research, and a brief list of further references.

The Interpretation of Aptitude Test Scores

The definition of a Category IV person is that the enlistee has obtained a score on the Armed Forces Qualification Test (AFQT) that falls within the 10th and 30th percentiles with respect to World War II enlistees. The AFQT is a paper-and-pencil test composed of subtests from the larger Armed Services Vocational Aptitude Battery (ASVAB). The AFQT score is obtained currently by summing scores on a test of vocabulary, paragraph comprehension, arithmetic reasoning (with the problems presented in words), and speed of simple arithmetic computation. Statistical studies have shown that the AFQT correlates highly with reading grade level. Indeed, these correlations are so high that scores on the AFQT could be substituted for scores on standard reading tests (Mathews, Valentine, and Sellman, 1978). To say that the AFQT "just measures reading" does not denigrate the test. Reading itself is important in a number of adult settings. Military personnel must read such things as directions, orders of the day, and equipment operating instructions in order to function in the services (Sticht, et al., 1972). Perhaps more importantly, reading ability is a generally accurate predictor of the ability to deal with the spoken language. Although there is such a thing as "specific reading disability," for most individuals difficulty in comprehending written material is indicative of difficulty in comprehending speech. The importance of this should be emphasized. Although "learning to read" is undoubtedly important, responding to low AFQT scores by sending people to reading schools may miss the point. Given that a person has been exposed to from eight to twelve years of public education, lack of reading ability is generally indicative of a more pervasive problem in linguistic comprehension.

Can one extrapolate from the AFQT score to something other than "verbal comprehension"? To what extent can one say that a person in Category IV has low intelligence? The answer to this question is more difficult. We will refer to "verbal ability" to indicate the ability to deal with both the written and spoken language. This is part of intelligence, but only a part. In our opinion, intelligence itself is not a single characteristic of an individual, but is rather an abstract term that refers to the tendency of a number of mental competencies to be positively correlated. One should think of intelligence as analogically similar to athletic ability - which may be expressed in several ways, and is obviously composed of more basic factors - instead of thinking of

intelligence as analagous to height or blood pressure. The question "Is a person intelligent enough for this job?" is almost always too simplistic. Instead the question should be "Does the person have, or can the person acquire, the mental competencies necessary for a particular job?" Given the AFQT score, one has a reasonable measure of verbal ability. What are the other competencies that might have been measured, and what information does the AFQT carry about them?

Although the exact nature of the underlying dimensions of intelligence is a matter of considerable controversy amongst psychologists, there is broad agreement on certain facts. Two broad domains of cognitive skill can be identified. One is the ability to deal with verbally presented material, either in speech or writing. The second is the ability to deal with visual-spatial problems, such as might be called upon in recognizing that one piece of a machine fitted against another in a particular way. Within the domain of verbal ability a further distinction is usually drawn between verbal comprehension - the ability to understand the material presented - and reasoning - the ability to draw conclusions from that information and to generalize conclusions from one situation to another. Reasoning ability is not entirely restricted to the verbal domain, but it is highly correlated with verbal comprehension, especially in young adults. Thus if one knows an enlistee's verbal comprehension ability, a quite accurate estimate of reasoning ability can be made. In particular, low comprehension ability is a good indicator of low reasoning ability.

Prediction is much less accurate across the verbal and spatial domains. The correlation between word knowledge (a good test of verbal comprehension) and a spatial reasoning test in the ASVAB is only .32 (Fruchter and Lee, 1977). Thus while one could guess from a very low verbal score that spatial reasoning was "below average", because the two are positively correlated, one's guess might not be accurate. A guess of "as low as the verbal score" would definitely not be warranted. On statistical grounds alone, one would expect a person with a low verbal score to have a somewhat higher spatial score.

A third domain of cognitive skills can be called, loosely, "psychomotor abilities." These are the abilities involved in real time co-ordination of perception and motor output. It is difficult to summarize this field succinctly, because it covers a great many tasks, ranging from the simple co-ordination required in, say, pointing a rifle, to the complex combination of memory, attention, perception, and motor skills needed to pilot a helicopter. Individual differences in psychomotor abilities are large, but appear to be highly specific to the particular acts of co-ordination involved. Thus performance on tasks

involving gross motor movements of arm or leg may not predict performance on tasks involving finger dexterity. Furthermore, the correlations between psychomotor performance and performance on tests of either verbal or spatial ability are at best modest. Correlations between verbal ability measures and psychomotor tasks will range from 0 to .4, depending on the particular task involved. Correlations can be expected to be slightly higher between psychomotor performance and spatial ability tasks. These facts pose a considerable problem for personnel selection. On logical grounds alone, one can assume that psychomotor skills are involved in machinery operation. The specific nature of these skills, however, makes it hard to talk about predicting who will be a "generally good machinery operator". One has to specify just what machinery will be used, and often it will not be possible to estimate the person's skill in machinery use until after the person has been trained to use the machinery.

Anticipated Behavioral Characteristics of Category IV Enlistees

The definition of a Category IV individual is somewhat variable, depending upon the lower cutoff limits that were in use at the time of the enlistment. In order to relate the Category IV classification to classifications used more generally outside the military, we shall assume that we are dealing with individuals who would score between 80 and 90 on a measure of verbal intelligence. Such individuals can accurately be characterized as having "low verbal aptitude," but it is a mistake to regard them as high grade mental retardates. (Borderline retardates would have scores in the 65-75 range, and would undoubtedly present much greater training and supervision problems than do the Category IV enlistees.)

What is the general behavior of such individuals? A low AFQT score suggests that the enlistee has, for the most part, been unsuccessful in school, even though a high school certificate may have been achieved. The failure to learn to read suggests strongly that the enlistee has not in the past, and will not in the future, absorb material well in a formal classroom setting. Our best guess is that the same enlistee will be relatively stronger in spatial and visualization than in language skills, although still somewhat below average. We have very little information about the person's psychomotor skills.

Individuals such as this have a great deal of trouble in formal educational settings. This is hardly surprising. Having already had great difficulty with one setting, the public schools, the person is not a good bet to do well in formal Army schooling. Perhaps because of their low reasoning skills, Category IV individuals are likely to have great difficulty in grasping general principles,

Table 1

Median Scores on the Army General Classification Test for Enlistees from Several Occupations (Data Gathered from the World War II Mobilization Sample)

85 - 89	Teamster
	Miner
	Farm Worker
	Lumberjack
91 - 94	Marine Fireman
	Laundry Machine Operator
	Laborer
	Barber
	Shoe Repairman
	Jackhammer Operator
	Groundman, Telephone, Telegraph, or Power
	Section Hand, Railway
95 - 104	Truck Driver
	Cook
	Construction Machine Operator
	Orderly
	Longshoreman
	Stationary Fireman
	General Painter

either when taught by expicit discussion or by example. Unfortunately, it appears that one of the most efficient ways to teach general principles is to rely on formal, verbal argument - precisely the sort of instructional technique that is hard for a person with low verbal aptitude.

The civilian experience also suggests that the low-aptitude person takes longer to teach. This is at least in part because the economical "verbal" mode of instruction will not work. The knowledge that the low-aptitude person takes away from training is highly specialized to the precise content of that training. Thus training must be specific to the job the person is going to do. Furthermore, it must be one-on-one, hands-on training in which the person is shown what to do, not told what to do. Given these qualifications, though, the low-aptitude individual is capable of operating in a variety of economic roles. People who do very poorly in school are often capable of achieving useful economic roles, provided that they find the right situation. The ideal niche for a low-aptitude person is one in which problems are predictable (and hence the individuals can be trained in advance to react to them) and in which the general environment is stable. One of the striking characteristics of a low-aptitude person is that behaviors that have been learned in one situation do not readily generalize to other situations that seem only superficially dissimilar.

A less abstract idea of what the Category IV is capable of can be obtained by considering the sorts of jobs that Category IV individuals hold in civilian life. Table 1 shows a selected list of jobs that were held, prior to enlistment, by "Category IV" individuals in the original World War II reference population. With only slight updating we believe that it would be accurate today. For comparison, the Table also shows selected jobs held by individuals who fell into Category III. A comparison of the jobs shows that in a competitive society, Category IV individuals hold unskilled or slightly skilled jobs. They generally do not operate complex machinery, nor do they hold jobs in situations in which one must make decisions about what to do. Note, though, that the jobs held do require operation of fairly simple machinery. On the other hand, almost none of these jobs require extensive verbal communication.

An analysis based on "jobs held" understates the ability of low-aptitude persons. In an economically competitive society, these individuals may not be able to compete for positions that they could occupy if they were employed. Table 2 shows the Department of Labor's estimates of the minimum qualifying scores for a variety of jobs. (The scores in this table are based on a General Aptitude Test Battery composite similar to the AFQT.) There are interesting

Table 2

Minimum General Intelligence Score Required for Success in Several Occupations According to the General Aptitude Test Battery Norms

Heavy Equipment Operator	75
Parking Enforcement Officer	75
Truck Driver	75
Mail Carrier	80
Orderly	80
Electronics Assembler	85
General Clerk	85
Key Punch Operator	85
Licensed Practical Nurse	85
Milkman	85
Telephone Operator	85
Bank Teller	90
Dental Assistant	90
Fire Fighter	90
Refrigeration Mechanic	90
Typist	95
Grocery Clerk	100

apparent inconsistencies between Table 1 and Table 2. In several cases the median scores obtained by persons in an occupation are considerably above the minimum scores required to perform in the same occupation.

The truth probably lies somewhere between these extremes. Because of competitive factors in the labor force, the civilian economy almost certainly underutilizes low-aptitude individuals. This is important because it suggests that our "common sense" observations about the sort of jobs people do will lead us to underestimate the capabilities of people similar to Category IV enlistees. On the other hand, the GATB test scores indicate what a single individual might do in a normal work setting. From the employer's point of view, a work force consisting of, say, 40% minimally qualified individuals is quite different from a work force that contains a few minimally qualified people. Unfortunately, the tradeoff between these factors is not known, and is unlikely to be obtained from studies of the civilian economy.

The Use of Equipment

Is the equipment found in the Army likely to pose a problem for the Category IV individual, and if so where will this problem lie and what can be done about it? In answering these questions it is important to distinguish between training, operation, and maintenance. The three phases of equipment use appear to require considerably different cognitive skills.

Training. Training is a form of education, and education is hard on a Category IV enlistee. It will take longer to train Category IV soldiers to operate equipment. When they are trained, they will not generalize readily from training to field situations. Thus even when a skill is mastered in training, the Category IV soldier may not realize that it is applicable in a superficially quite different situation.

The highly efficient lecture mode of instruction will be difficult for the Category IV soldier, even when lectures make heavy use of audio-visual procedures and demonstrations. Audio-visual procedures may require just as much imagination as dealing with text, especially if they are schematized. Even training films are idealizations that require the viewer to imagine himself being in the situation. Equally important, Category IV individuals have not learned the skills necessary to learn from lectures, film strips, or audio-visual

presentations. There is no particular reason to believe that the Army can can succeed in teaching these skills where the public school system failed. We are similarly pessimistic about proposals to "teach the Category IV's to read." The probability of significantly improving reading skills in the time available is small. (This observation, of course, should be tempered for anyone who has had little exposure to English.)

The answer to the training problem appears to be simple, direct, and expensive. Category IV individuals learn best by doing. They need closely supervised experience with the specific tasks that they are going to do in the field. If possible, they should be trained by the same people who supervise them in the field. The training can be expected to go slowly, and the trainer will have to be extremely alert to the actual progress of the student.

We shall return to this basic point several times. The available psychological literature, although seldom directly relevant to the Category IV problem, contains a great deal of tangentially relevant information. All of this information is consistent in indicating that low-aptitude individuals do not generalize easily from training situations. Thus training must be as much like the field situation as possible.

Operation. The typical finding is that Category IV individuals can operate equipment providing that it is not "too complex." But what does complexity mean? For our purposes, one can distinguish between complexity due to the fact that many things must be done in a predictable sequence, and complexity because there are many decision points. Changing a tire is complex in the former sense, while driving in traffic is complex in the latter.

Low-aptitude individuals can be trained to handle detailed but predictable tasks. Indeed, mental retardates (who are considerably below the "Category IV" level in general mental aptitude) have been trained to operate reasonably detailed but predictable equipment. Naturally, any equipment operation involves dealing with some contingencies. The Category IV person can handle these, provided that the contingencies that occur can be specified in advance. To illustrate, people with low general aptitude obtain jobs as janitors. These jobs require numerous decisions about what equipment to use, but the distinctions are usually clear cut.

The Category IV soldier is likely to fail in equipment operation when an unexpected contingency occurs, or when situations change in such a way that

previous specific training is no longer appropriate. The low-aptitude person can tolerate little change in the superficial aspects of a work situation. For example, changes in supervisory personnel would be likely to be more disruptive to Category IV than to cognitively more competent individuals, even though such changes should make little, if any, difference in equipment operation per se. For similar reasons, a change in assignment, even though not accompanied by a change in duties (e.g. a transfer from one infantry company to another) might be more disruptive to the lower level person. Changes in equipment will also be disruptive. When new equipment is introduced into a unit there is an inevitable retraining period. Normally, knowledge of the principles used governing the operation of the old equipment will help in understanding the new equipment. This will be less the case with Category IV personnel.

This section is well summarized by the remarks of an Army Armor Corps captain whom we interviewed while preparing this report. He pointed out that preparing the power pack of a tank for removal is a complex but predictable task that requires the execution of a number of steps in a particular sequence. Based on his experience, he felt Category IV soldiers could learn to do this task. Repairing a sight mount "in the field" is a simpler but less predictable task. The soldier must find whatever is at hand that can hold the sight in place, install it, and recalibrate the weapon. Such improvisation requires insights into the purpose of equipment, rather than into the way particular pieces of equipment fit together. These insights are likely to be hard to develop in the Category IV soldier.

Clearly the key to utilizing the Category IV soldier is to make the situation as predictable as possible. Military operations are, by their nature, often unpredictable. Thus in a changeable situation it will be extremely important that the Category IV soldier be closely supervised, and supervised by someone with whom he is familiar, so that the supervisor can act as decision maker, filter out the unfamiliar, and present the task to the soldier as being one on which he is already trained. Such a requirement poses a considerable problem for the combat arms, because casualties among noncommissioned officers (NCOs) can be predicted to be extremely disruptive. If the private soldiers are of low aptitude, one can neither count on finding leaders to replace the lost sergeants from within the unit, nor can one easily transfer new sergeants into leadership positions.

Maintenance. Maintenance and repair lie somewhere between training and equipment operation. Some aspects of maintenance are routine and highly predictable, such as changing the oil in an engine. Regularly scheduled routine maintenance, even of a detailed nature, can probably be performed by Category IV soldiers. Maintenance that requires fault diagnosis is likely to prove difficult, because this sort of task requires logical reasoning, an ability that is likely to be weak in a person falling into Category IV. Some maintenance and repair tasks involve the following of relatively abstract instructions, presented either in writing or by schematized diagrams. Both ways of presenting repair instructions are likely to be hard for a person with low verbal aptitude.

Category IV soldiers involved in maintenance and repair will require a large amount of supervision. Our feeling is that the optimal repair shop containing a large number of Category IV individuals should be organized around highly specialized work stations. A supervisor would determine what needed to be done for a specific piece of equipment, and would move that equipment, step by step, from one work station to the next. The shop atmosphere should be kept as stable as possible.

A General Comment. Although our charge is to discuss the use of equipment, we would like to raise a "sociological" point. Most of the present information concerning low-aptitude workers is based on the performance of isolated individuals working with cognitively more competent colleagues. Such people undoubtedly benefit from informal, supportive social networks. Indeed, the low-aptitude person probably needs these networks more than his or her more talented colleagues. Very little is known about the effects of having a work force composed predominantly of Category IVs. (Rather more is known about the problems encountered in "sheltered workshops," but these institutions generally deal with people very considerably below the Category IV level.) The following speculative remarks are offered.

1) Close supervision will be essential. The supervisor must act as a buffer between the individual and a potentially confusing world. In military terms, the burden is once again on the noncommissioned officer.

2) For this reason, it is important that first-line supervision be stable. Frequent rotation of NCOs would be a serious problem. In combat situations NCO casualties will not be easy to replace.

3) If half the work force are Category IVs, then half are not. Higher category personnel will preempt informal leadership roles. Thus the morale and motivation of these individuals becomes important. Their example will be followed. Will it be a good or bad example?

4) Unfortunately, the pervasive supervision appropriate for Category IV soldiers may hurt the morale of more competent individuals in the same unit. To illustrate, the "assembly line" operation suggested for a repair shop staffed largely by Category IV personnel is known to be distasteful to more competent individuals; witness the current move toward craftwork and personalized completion of jobs in many industries. This could present a still further problem for supervisors.

5) If an appropriately stable situation can be maintained, Category IV soldiers will find it agreeable and will not be motivated to leave the service either for change or for potential career advancement. Neither will they expect career advancement within the service. Thus longer enlistments may be possible at the lower ranks in order to capitalize on a considerable training investment.

Recommendations

Policies in Dealing with Category IV Individuals

1. If possible, training should be accomplished within the operating units. Training should be on the equipment used in the field, and should be "hands on" training. Allowance should be made for the length of time required. This could be as much as 50% greater than the normal length of training.

2. The noncommissioned officers who deal with Category IV personnel are crucial. They will be required to give individualized training, to check frequently to make sure that apparently simple tasks really have been learned, and to tailor training to the capabilities of the individual soldier. Probably the most important single thing that the Army can do to utilize Category IV personnel is to provide adequately trained and motivated NCOs. These NCOs will not have to have a high degree of technical skill, but will have to deal with difficult teaching and leadership situations.

3. The operating environment should be stable. Once trained, the Category IV soldier should be kept on the same job, and in the same unit. Supervisory

personnel should be stable. As retraining will be long, serious consideration should be given to the utility of introducing new equipment into a unit. When new equipment is introduced, careful preparation for the changeover will be essential.

4. Repair procedures should rely on exhange of parts (which assumes an adequate logistical system) and routine operations. Little reliance can be placed on the ability of field personnel to improvise.

5. The Category IV soldier should be trained to do one highly specific job. Multiple personnel assignments, such as are routinely used in some military situations (e.g. the submarine service) will not be feasible.

6. In evaluating the readiness of a combat unit consisting of large numbers of Category IV personnel, consideration should be given to unit performance when NCO and junior officer casualties are simulated.

7. Studies of job performance should not be based on a narrow definition of the work required. The total adjustment of the soldier to the Army must be considered. This is particularly important in the case of a Category IV person, who may be less able to divorce work and general social adjustment. Even when the Category IV soldier is assigned to a nonverbal job, the soldier still has to deal with a verbal Army.

Research Questions. There is a great deal that we do not know. Indeed, most psychological research focuses either on individuals with superior aptitude (typically college students) or upon mental retardates, whose performance is much below that of low-aptitude individuals. As was pointed out, studies of how low-aptitude persons function in equipment-oriented jobs are not likely to be conducted in the civilian work force, for the simple reason that civilian employers have more economic labor forces available to them. If the Army were to plunge ahead with an intensive training program for Category IV soldiers, however, it could be an extremely expensive one. It is important that such training be evaluated carefully. It would also be desirable to extend the initial categorization of soldiers beyond the realtively crude categorization system based on the AFQT. The following suggestions are offered.

1. Is it possible to develop tests in the psychomotor field that would predict performance, although they might not predict training? This has been tried before, without notable success. The new microcomputer technology may make the construction of appropriate psychomotor tests feasible, but this contention has not been proven.

2. Most of our current data concerning the relation between tests, and between test and work performance, is based upon studies of the entire range of capabilities in the population. Studies of the relationship between test and work performance within the Category IV population itself should be conducted. Much of this data must already be available to the Army, through records of ASVAB scores and work records.

3. Studies of the social aspects of the work situation should be instituted. Of particular concern are such questions as (a) the trade-off between unit performance and the percentage of Category IV individuals in the unit, (b) the appropriateness of special training for officers and NCOs who deal with such individuals, and (c) the effect of structuring job situations in various ways both upon the Category IV soldier and upon other soldiers in the same unit.

Suggested Further Readings

There is really very little that we have found that is directly relevant. Most of the psychological literature deals with individuals at a higher or lower range of ability. This is certainly true of review or textbook articles on "intelligence," for these are heavily weighted toward studies in educational settings, and usually of college students, retardates, or the elderly - all the people who do not enlist in the Army. The following three general references are most helpful.

A good overview of intelligence testing in general can be found in the first section of the book

Sternberg, R. Intelligence, Information Processing, and Analogical Reasoning. Hillsdale, N.J.: Erlbaum Associates, 1977.

The first part of this book contains a good description of various approaches to the definition of intelligence. The remainder of the work is concerned with a specific research project and is of almost no relevance to the question at hand.

A more general textbook approach is found in

Willerman, Lee. The Psychology of Individual and Group Differences. San Francisco: Freeman, 1978. (See in particular chapters 1,2,4,6, and 7.)

A discussion of the use of tests in industrial research is found in

Dunnet, M.D. "Aptitudes, Abilities, and Skills." In M.D. Dunnette, editor. Handbook of Industrial and Organizational Psychology. Chicago: Rand-McNally, 1976.

References

Cooper, R.V.L. Military Manpower and the All-Volunteer Force. Technical Report R-1450-ARPA RAND Corporation. September 1977.

Fruchter, D.A., and M.J. Ree. Development of the Armed Services Vocational Aptitude Battery: Forms 8, 9 and 10. AS AFHRL Technical Report 77-19. March 1979.

Holden, C. "Doubts Mounting about All-Volunteer Force." Science 209 (1980): 1095-98.

Mathews, J.J., L.D. Valentine, and W.S. Sellman. Prediction of Reading Grade Levels of Service Applicants from Armed Services Vocational Aptitude Battery. US AFHRL Technical Report 78-82, December 1978.

Sticht, T.G., J.S. Caylor, R.P. Kern, and L.C. Fox. "Project REALISTIC: Determination of Adult Functional Literary Skill Levels." Reading Research Quarterly 7 (1972): 424-65.

US Department of Labor. Manpower Administration. Manual for the USTES General Aptitude Test Battery. Washington, 1970.

APPENDIX B

COMBAT EFFECTIVENESS AND CHARACTERISTICS OF SOCIETY*

Introduction

A survey of modern military history suggests that there is a general consensus that the following forces had high combat effectiveness and that, as a corollary, their troops had a high level of capability:

- Spanish infantry (16th Century)
- Turkish janissaries (16th Century)
- Swedish troops (17th Century)
- Prussian infantry (18th Century)
- French Imperial Guard (Napoleonic wars)
- British infantry (Napoleonic wars)
- Zulu warriors (1870s)
- US cavalry (late 19th Century)
- Prussian/German infantry (c. 1864-1945)

There is an equally strong consensus as to which forces have performed poorly in combat, in large part because their soldiers had a low level of capability. Among these were:

- Indian Moghul troops (17th Century)
- US militia (1812)
- Italian forces (World War II)
- Iraqi forces (present)

In considering these cases the reader should keep in mind the distinction between troop capability and combat effectiveness of units. The military historian, in assessing troop performance, can see only the effectiveness of troops in units. Troop capability, on the other hand, describes what troops contribute to combat as individuals. The troop capability concept does not include the impact of leadership, unit training, and logistics on combat. Of these additional characteristics, leadership appears to have exerted an especially clear and persistent effect.

*This appendix is based on the work of Robert McQuie.

In the analysis that follows, the effect of leadership has influenced both the performances of the forces in combat and the judgment of the military historians about that performance. While one may assume for analysis purposes that combat effectiveness, including the influence of leadership, is a good approximation of troop capability, the two terms do not refer to quite the same characteristic, and readers should bear in mind that in this appendix the assumption has been made that for each of the different forces analyzed, combat effectiveness and troop capability are approximately the same.

Objective

The observations about "best" and "worst" forces, noted above, lead to the following conclusion: Since we are able to agree on which forces were the best in combat and which were the worst, the process can be carried a step further to gain agreement about the relative effectiveness of the forces falling between the two extremes. If the level of effectiveness of a force (or, roughly, the capability of its troops) can be placed on a scale, then this characteristic of force effectiveness can be related to various characteristics of the nation or society that employed the force.

The problem of scaling intangible qualities is one that has been addressed by many academic disciplines, ranging from psychology to marketing. In this appendix, the determinants of troop capability were examined by the same overall approach employed in many of the social sciences. Expert judgment was employed to develop a scale of combat effectiveness for troops engaged in combat since the beginning of World War II. Then the degree of association between combat effectiveness and national characteristics which might be expected to influence troop capablility, including demography, economics, politics, education, military attitudes, and climate, was measured. Based on the results of this measurement, it was possible to isolate some characteristics strongly associated with combat effectiveness and troop capability. This should contribute to development of appropriate factors for war games.

Procedure

The steps involved in carrying out the analysis are listed below. Each is then discussed separately in some detail.

1. Seventeen forces that were engaged in combat during the period from the start of World War II to the present were selected.
2. The judgments of a group of military historians about the comparative effectiveness of these forces were obtained.
3. The judgments were converted into a scale on which the effectiveness of each force was quantified.
4. Historical data about the demographic, economic, political, military, educational, and climatic characteristics of the nation to which each force belonged was assembled.
5. The degree of association, or correlation, between each characteristic and the quantified estimates of force effectiveness from Step 3 was calculated.
6. Conclusions that appeared appropriate were drawn:
    a. National characteristics strongly associated with combat effectiveness were identified.
    b. Judgment was made as to the feasibility of constructing a formula relating national statistical characteristics to force effectiveness and troop capability, for war gaming purposes.

1. Forces Selected

The evaluations were based on forces in combat about which the evaluators had detailed knowledge. These forces also had to be ones that were involved in a number of combat engagements, so that the evaluation would reflect a true combat effectiveness rather than a nonrepresentative performance in one battle. The forces selected were:

Force	Dates
Chinese	1950-52
Egyptian	1967-73
German	1943-44
Iraqi	1967-73
Israeli	1967-73
Japanese	1943-44
Jordanian	1967-73
North Korean	1950-52

Force (Continued)	Dates (Continued)
North Vietnamese	1966-70
South Korean	1950-52
South Vietnamese	1966-70
Soviet Union	1943-44
Syrian	1967-73
United Kingdom	1943-44
United States	1943-44
United States	1950-52
United States	1966-70

2. Judgments of Military Historians

The second step involved asking three military historians to compare the combat effectiveness of all the forces, two at a time. The historians were asked to compare "troop quality," because it was the investigator's judgment that this would be the most familiar designation for combat effectiveness, as distinguished from the effects of numbers and weapons. It was recognized, however, that it would be impossible for these historians, or anyone, to judge troop quality at the force level, and that what was actually being judged was the combat effectiveness of the forces.

Each historian was asked, "Which force had the higher troop quality, the Germans or the British in 1943-44?" and then, "Which force had the higher troop quality, the Germans in 1943-44 or the Chinese in 1950-52?" The procedure was carried out with all possible pairs of the forces listed above. In general, the number of comparisons is given by the formula $n(n-1)/2$, in this case 136 comparisons. Figure 1 shows the form on which these comparisons were made.

It should be noted that, although the size of the panel may seem small, in establishing a consensus of informed judgment the qualifications of the panel members were judged to be much more important in this preliminary effort than the size of the sample. This panel was the best informed one that could be questioned in the time available. The form has been circulated to six additional highly qualified military historians, and their responses will be used to adjust the scale obtained, if necessary. Since the agreement among the three queried was close, it is not expected that there will be significant changes when all the responses are received.

HISTORICAL EVALUATION &
RESEARCH ORGANIZATION
2301 Gallows Road
Dunn Loring, VA 22027
(703) 560-6427

Figure 1

TROOP QUALITY WORKSHEET

Prepared by_____

DIRECTIONS
For each pair of countrie select the one with the higher quality troops. Place its symbol in the appropriate cell. In case of a tie, use an equal sign (=).

Nation	Date	Symbol	US4	GB	G	JA	SU	US5	C	NK	SK	US6	NV	SV	E	IR	I	JO	S
United States	1943-44	US4																	
Great Britain	1943-44	GB																	
Germany	1943-44	G																	
Japan	1943-44	JA																	
Soviet Union	1943-44	SU																	
United States	1950-51	US5																	
China	1950-51	C																	
N. Korea	1950-51	NK																	
S. Korea	1950-51	SK																	
United States	1966-70	US6																	
N. Vietnam	1966-70	NV																	
S. Vietnam	1966-70	SV																	
Egypt	1967-73	E																	
Iraq	1967-73	IR																	
Israel	1967-73	I																	
Jordan	1967-73	JO																	
Syria	1967-73	S																	

3. Scales of Troop Capability

Each historian's judgments were converted into a scale, ranging from a value of zero for the force with the lowest effectiveness or capability, to a value of 100 for the force with the greatest effectiveness or capability. This conversion was carried out by the Method of Paired Comparions.

This method was devised in the early 1930s by Thurstone, one of the fathers of psychometrics. It is based on a mathematical fact that may not be readily apparent, unless one has explored the mathematical implications of choosing between two alternatives. The fact is that when a person makes a series of such choices among two alternatives, he has carried out the mathematical equivalent of assigning all of the alternatives to locations on a scale. This scale can be normalized so that its maximum and minimum values are any preselected values, often zero and one hundred. While the individual has difficulty in assigning items being evaluated directly to the scale, he has no difficulty in evaluating which of any two of them is preferable to the other.

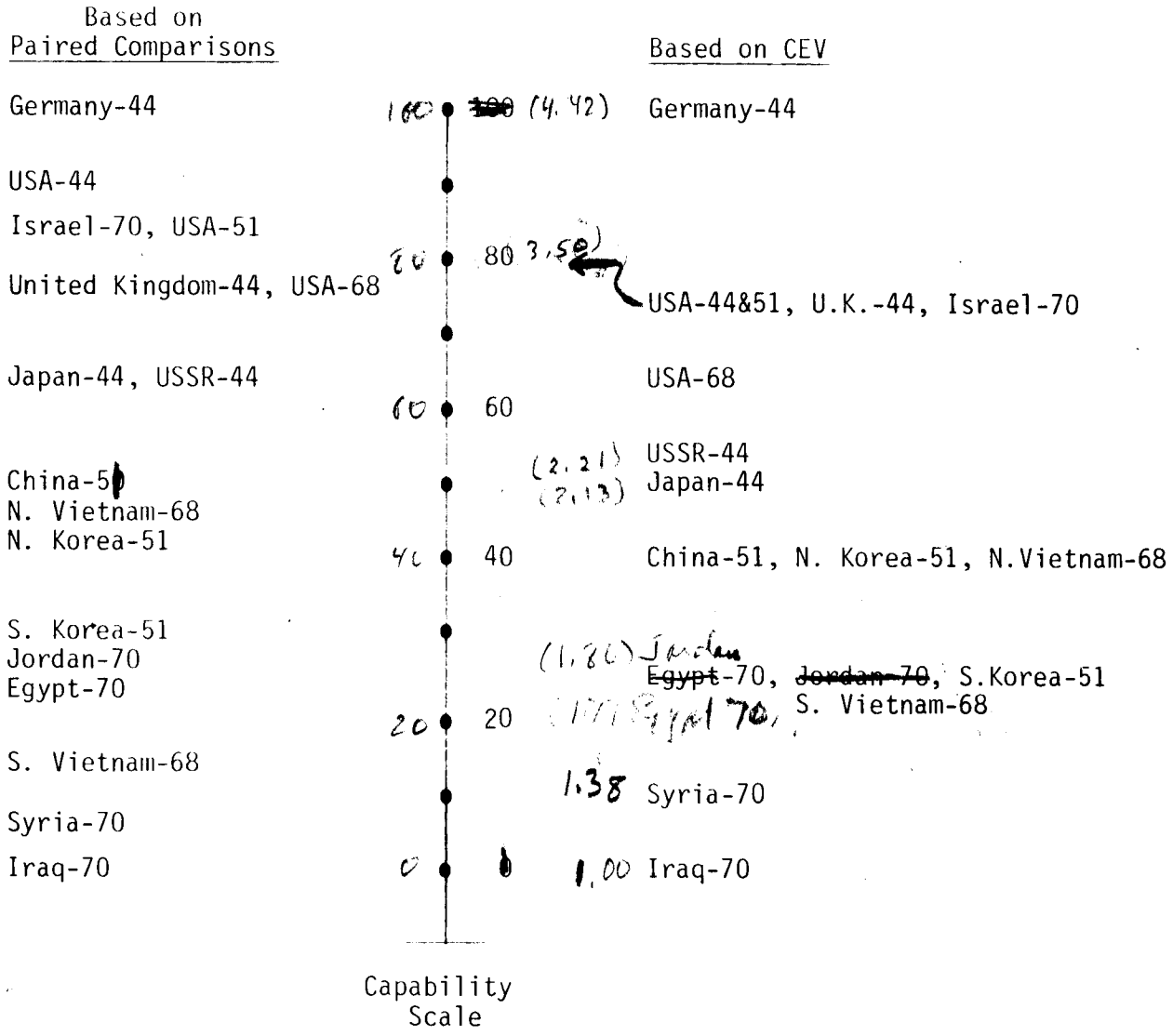
Since Thurstone's day, his original idea has grown into a field of mathematics with a large literature in its own right, nonmetric scaling. One of the more recent texts on the subject of scaling, listed in the bibliography, is that of van der Ven.

Using this mathematical methodology, the choices of each of the three historians were converted into points on a scale running from zero to 100. The results were listed statistically to determine whether the three scales might be significantly different from one another. This was done by applying the Friedman Two-Way Analysis of Variance test. A null hypothesis of no significant difference among the three could not be rejected at either a 1% or 5% level of confidence. Accordingly, the judgments were pooled, and the scale value assigned to each force was taken as the average of the values resulting from the judgments of the three military historians.

In a large number of judgments, a rater might make logical lapses. That is, he might state that Force A was superior to Force B and that Force B was superior to Force C, and then state that Force C was superior to Force A. To test the possibility of this having occurred, Kendall's Coefficient of Consistency was calculated for the evaluations of each of the three raters. The results were greater than .9 for each of them, indicating a high degree of internal logic in each of the three sets of judgments.

Figure 2

COMPARISON OF SCALES OF TROOP CAPABILITY



The resulting scale is shown on the left hand side of Figure 2. Also plotted on this table are the troop capability estimates based on the Relative Combat Effectiveness Value (CEV) for each of these forces, as obtained from HERO's Quantified Judgment Model. As may be observed, there is very little difference between the scales obtained by the two methods.

4. Data about the Nations

Data was next accumulated about each of the nations whose forces was being evaluated. The sources of the data were the cross-national data publications of the United Nations, the World Bank, and related sources. (See bibliography.) These general collections of data were supplemented in a number of cases by single items of data from other publications (also listed in the bibliography).

Data was accumulated on the following characteristics:

Area	Characteristic
Demography	Birth rate
	Household size
	Male life expectancy
	Population density
	Protein consumption per person
Education	Male literacy (%)
	Primary schooling (%)
	Secondary schooling (%)
Economics	Energy consumption per person
	GNP per person
	Physical quality of life index
Politics	Civil disorder index
	Civil liberties index
	Political power index
	Political rights index

Area (Continued)	Characteristic (Continued)
Military	Battle deaths per month per million inhabitants
	Civilian/military ratio at start of war
	Military power index
	Military spending per person
Climatic	Rainfall
	Temperature in coldest month
	Temperature in hottest month

These characteristics were selected according to two guidelines. First, the Book of World Rankings was used as a rough screen of what might be worthy of further investigation. This volume ranks the nations of the world according to approximately 450 characteristics. In cases where a large number of the nations being evaluated were at either the high or the low end of a ranking, that characteristic was tentatively selected for evaluation. Second, an attempt was made to get at least two characteristics that were logically related to each broad area, and for which a plausible argument could be made that they might be related to troop capability. Thus, Average Household Size was selected because it was felt that the type of family structure in which a soldier grew up might be related to his capability in combat.

Several of the items of data are index numbers, well supported by discussion in the scholarly literature, that measure characteristics too complex to be reflected in a single item of data. Among these are the Physical Quality of Life index, devised by the Overseas Development Council, and the Political Power index devised by the Georgetown University Center for Strategic and International Studies. An attempt was made to pick indexes that are supported by explanations published in the open literature of the particular discipline involved.

The data was gathered for a period that was approximately contemporary with the war in which the troops were being evaluated. Most items were for the nation within two or three years of the time of the war and none were more than 10 years before or after it. Where data this current could not be located or calculated from other data, the item was left blank. Seven of the 24 characteristics had 12 or more items of data, while 5 characteristics had 7 or fewer items of data.

5. Characteristics Associated with Combat Effectiveness

The product-moment correlation coefficient (r) was calculated between the scaled value representing force effectiveness (or troop capability) and each of the characteristics of the nations listed previously. These coefficients were then interpreted to answer the following question: Is the association between effectiveness and a particular characteristic strong enough to reflect an actual association between the set of characteristics for all countries and all wars during the period or could it have been due to the fact that only a small sample of data was employed?

After applying a standard statistical procedure, it is possible to make an inference about whether the association or correlation in the particular sample is "significant," that is, whether it arose from an underlying relation or from the particular sample of data employed.

A high value of the correlation coefficient does not imply a cause and effect relation between the two variables. It only means that a positive change in one is accompanied by a change (positive or negative) in the other. The correlation coefficients were evaluated at two levels of significance, the 5% level implying one chance in 20 of the association's resulting from chance, and the 1% level, implying one chance in 100. These are the levels of testing commonly employed in most statistical analysis. The results are shown in Figure 3.

6. Conclusions

An examination of Figures 1 and 2 permits the following conclusions to be made:

1. The German Army in World War II had the highest level of combat effectiveness of those examined.
2. The Iraqi Army from 1967 through 1973 had the lowest level of combat effectiveness of those examined.
3. US troops in the Korean War had slightly lower combat effectiveness than those in World War II, while those in Vietnam had a slightly lower effectiveness than those in Korea.
4. The Japanese and Soviet armies in World War II had a much lower level of combat effectiveness than the United States troops in the Vietnam War.

Figure 3

CHARACTERISTICS ASSOCIATED WITH TROOP CAPABILITY IN COMBAT

Characteristic	Probable Degree of Association		
	Strong[a]	Likely[b]	Uncertain[c]
Demographic	Household size (-) Birth rate (-)	Male life expectancy Population density Protein Consumption	Rural population density
Economic		GNP per capita Physical quality of life index Energy consumption	
Educational	Male Literacy	Secondary schooling	Primary schooling
Political		Civilian disorder index Political rights index (-)	Political power index Civil liberties index
Military			Battle deaths rate Civilian-military ratio Military spending Military power index
Climatic	Temperature-Hottest month (-)		Temperature-Coldest month Rainfall

a. < 1 chance in 100 of being due to chance (1% confidence level)
b. < 5 chances in 100 of being due to chance (5% confidence level)
c. > 5 chances in 100 of being due to chance (>5% confidence level)

5. Of the national characteristics that were strongly associated with effectiveness (1% confidence level), three had a negative correlation, that is, increases in the characteristic were associated with decreases in combat effectiveness. These were:

   Birth rate
   Household size
   Temperature in the hottest month

6. One strongly associated characteristic (1% confidence level) had a positive correlation. This was:

   Male literacy

7. None of the characteristics reflecting the political or military characteristics of the society were strongly associated with combat effectiveness.
8. The high level of correlation noted above indicates the probability that a formula could be derived to estimate combat effectiveness for the forces of any two forces that are being studied in a war game.
9. No significant differences are observable between a scale of combat effectiveness or troop capability derived from informed military judgment and one based on the CEV employed in the Quantified Judgment Method.

BIBLIOGRAPHY

General Data Sources

Bureau of the Census. Historical Statistics of the United States. 2 vols. Washington, D.C.: Government Printing Office, 1975.

Demographic Yearbook 1948. Lake Success, N.Y.: United Nations, 1949.

Kurian, George Thomas. Book of World Rankings. New York: Facts on File Publications, 1975.

Mitchell, B.R. European Historical Statistics, 1750-1950. New York: Columbia University Press, 1975.

Mitchell, B.R., and H.G. Jones. Second Abstract of British Historical Statistics. Cambridge University Press, 1971.

Showers, Victor. World Facts & Figures. New York: Wiley, 1979.

_____. World of Figures. New York: Wiley, 1973.

Singer, J. David, and Melvin Small. The Wages of War 1816-1965, A Statistical Handbook. New York: Wiley, 1972.

Statistical Yearbook 1948. Lake Success, N.Y.: United Nations, 1949.

World Almanac and Book of Facts 1974. Ed. by G.E. Delury. New York: Newspaper Enterprise Assn., 1973.

The World Bank. World Tables 1976. Baltimore, MD: Johns Hopkins University Press, 1976.

Sources of Specific Items of Data

Cline, Ray. World Power Assessment, A Calculus of Strategic Drift. Washington, D.C.: Georgetown University Center for Strategic and International Studies, 1975

Encyclopedia of World War II. Ed. by Thomas Parish. New York: Simon & Schuster, 1978.

Freedom at Issue. New York: Freedom House. January-February 1977.

Information USSR. Ed. by Robert Morwell. Oxford: Peogamba, 1962.

International Encyclopedia of Social Sciences. New York: Macmillan, 1968.

Military Balance 1973-74. London: International Institute of Strategic Studies, 1973.

Morris, Morris David. Measuring the Conditions of the World's Poor. New York: Pergamon, 1979.

Sellers, Robert C. Reference Handbook of the Armed Forces of the World. Garden City, New York: Robert C. Sellers & Assoc./Praeger, 1966, 68, 77.

Special Operations Research Office of American Univ. Area Handbook for Israel. DA Pam. 550-25. Washington, D.C.: Hq. Dept. of the Army, September 1970.

_____. Area Handbook for Korea. Washington, D.C.: The American University, November 1964.

UNESCO. World Illiteracy at Mid Century. Westport, Connecticut: Greenwood Press, 1957.

van der Ven, A.H.G.S. Introduction to Scaling. New York: Wiley, 1980.