AD 699 490

COLLECTED PAPERS PREPARED UNDER WORK UNIT AAA: FACTORS AFFECTING EFFICIENCY AND MORALE IN ANTIAIRCRAFT ARTILLERY BATTERIES

Francis H. Palmer, et al

Human Resources Research Organization
Alexandria, Virginia

November 1969

Professional Paper 33-69
November 1969

Collected Papers Prepared Under Work Unit AAA:

Factors Affecting Efficiency and Morale in Antiaircraft Artillery Batteries

HumRRO

HUMAN RESOURCES RESEARCH ORGANIZATION

43

The Human Resources Research Organization (HumRRO) is a nonprofit corporation established in 1969 to conduct research in the field of training and education. It is a continuation of The George Washington University Human Resources Research Office. HumRRO's general purpose is to improve human performance, particularly in organizational settings, through behavioral and social science research, development, and consultation. HumRRO's mission in work performed under contract with the Department of the Army is to conduct research in the fields of training, motivation, and leadership.

Published
November 1969
by
HUMAN RESOURCES RESEARCH ORGANIZATION
300 North Washington Street
Alexandria, Virginia 22314

Prefatory Note

In Work Unit AAA, Factors Affecting Efficiency and Morale in Antiaircraft Artillery Batteries, the Human Resources Research Organization undertook a study of a number of particularly effective and relatively ineffective on-site antiaircraft batteries. Information was sought to determine certain of the less obvious human factors that contribute most heavily to group performance. The research took place during the years 1954-1955, and was conducted at Fort Ord, California, by HumRRO Division No. 3.

The papers in this collection include two presentations at professional meetings, two professional journal articles, and a task paper summarizing the research program.

Because of the continuing relevance of the subject matter of these papers, they are being issued in a group as part of the HumRRO Professional Paper series. This series was initiated in order to provide permanent record of specialized aspects of HumRRO work, and deposit in the scientific and technical information storage and retrieval systems of the Department of Defense and the Federal Clearinghouse.

CONTENTS

BATTERY EFFECTIVENESS: ASSESSMENT OF COMPARATIVE PERFORMANCE[1]

Francis H. Palmer and Thomas I. Myers

This article describes the development of realistic measures to identify highly efficient and less efficient antiaircraft batteries and discusses the extent to which the several measures of performance are related. Under specific discussion are range of radar pickup, firing range scores, radar maintenance, artillery maintenance, defense commander's rating, and adverse personnel actions.

Every artilleryman knows that there is a difference between the operating potential of a radar gun system and the system's actual performance when men are operating it. The vital point is the size of the difference. In a sense the job of the commander, whose mission can be fulfilled only through use of such a system, is to develop the skill of the men he has available so that the machine potential and the man-machine potential of the equipment differ as little as possible.

The commander cannot shoot higher or farther or with greater accuracy than his equipment system permits. He *can* develop the efficiency of his men to levels where they approach the maximum potential of the machine.

Just what are the personal elements that contribute to the differences between potential and actual performance? This problem is being studied under the sponsorship of the Chief of Army Field Forces. Extensive observations of leadership techniques and administrative practices, along with their relation to the performance of units, are being made in this general research program.

In any effort to determine what factors influence the performance of a group of people, one of the first considerations is how to measure that performance, accurately and in meaningful ways. The problem of developing such measures exists in all sciences and is commonly referred to as the "criterion" problem—usually the toughest question in any research design. If you are interested in studying the personalities of outstanding combat riflemen, you must first identify the actions essential to outstanding combat performance and, on the basis of these actions, single out the men to be studied. If you are interested in

[1]This article appeared in *Antiaircraft Journal*, vol. 97, no. 6, November-December 1954.

studying the characteristics of a good antiaircraft battery, you must determine what performances are essential to the fulfillment of the unit's mission, and then develop techniques with which those performances can be measured.

This article describes the first step in such a program: that is, the development of realistic measures that can identify highly efficient and less efficient antiaircraft batteries. A second purpose will be to discuss the extent to which the several measures of performance are related—to see, for example, if a unit that is highly proficient in artillery maintenance is likely to be as successful in radar pickup and other essential activities.

In the early spring of 1953, the HumRRO research staff discussed the project with military staff members, representative battery commanders, and operations center personnel. The aim of the discussions was to identify the activities a battery must perform successfully in order to accomplish its mission.

The military advisers on the research reached general agreement that the essentials for satisfactory battery performance were these:

1. A battery should be able to pick up incoming targets on its radar at a range commensurate with the maximum potential of its equipment.

2. A battery should be prepared to engage an incoming target when the target comes within gun range.

3. A battery should maintain its equipment in a manner which insures its being prepared to fire on a few minutes' notice.

In addition, several commanders pointed out the importance of some measure which would reflect the extent of a unit's adverse personnel actions. The environment of the on-site battery in the zone of interior (ZI) is sufficiently complicated by such factors as relations with civilian communities that a measure of this sort seemed justified.

When these critical activities had been identified, the researchers worked out methods of evaluating units on the elements of performance implied. These measures were:

1. Range of Radar Pickup
2. Firing Range Scores
3. Radar Maintenance
4. Artillery Maintenance
5. Defense Commander's Rating
6. Adverse Personnel Actions

These six activities were measured in the following manner.

Range of Radar Pickup. A large number of tracking missions were flown under the direction of the operations officer of the defense concerned. The strikes composing these missions were flown at an altitude of 15,000 feet or higher. Units were not scored for strikes where masking significantly interfered with possible pickup. For each strike the battery concerned reported the time, coordinate, slant range, and azimuth at the point of locking-on with the track radar of the M33. These reports were checked against the track maps of each strike developed by the Antiaircraft Operations Center (AAOC), and the battery was scored for each strike. Battery performance over the many strikes was averaged out so that an individual score of average pickup was available for each unit.

Firing Range Scores. Firing range scores were decided on as the best approximation for a direct measure of battery preparedness for engagement. The last official score the unit had received during range firing was used. Such a score is of course not completely adequate as a measure of preparedness. However, the ratings necessary to ascertain specifically whether a unit was prepared to engage at maximum fuze would have been uneconomical in terms of personnel and time. For the purposes of the study, then, the range scores were accepted.

Radar Maintenance. Three methods were worked out to evaluate radar maintenance. In the first method, the ordnance team responsible for each unit's radar was asked to examine its job orders for the three previous months. Each job order was judged, in the presence of a supervisory technician, with regard to whether the repair had been made necessary by inadequate maintenance or by fair wear and tear. When a repair was attributed to inadequate maintenance, further judgments were made about the seriousness of the repair and the extent of poor maintenance. The scores for this measure, then, were made on the basis of judged relationships between poor maintenance and repair incidents.

The second method also was based upon the job orders. The orders were counted, without reference to the conditions precipitating the repair. This score therefore indicated how often ordnance had worked on each set.

The third method used the AAOC records and the unit repair reports to evaluate maintenance. Each day that the unit was considered out of action because the radar was inoperable was tallied, and a score for each unit determined.

Artillery Maintenance. Artillery maintenance was scored by a method similar to the first method listed for scoring radar maintenance. The ordnance team responsible for the unit's equipment made judgments about the precipitating causes of repair, and a battery score was determined.

Defense Commander' Rating. The defense commander instructed his staff to keep records for several activities over a three-month period. The ten activities specified were: S1 functions, S2 functions, S3 Artillery and Operations, S3 Training, S4 Food Service, S4 Supply, communications, generators, vehicles, radar.

Evaluations were based upon a score of 4 for superior, 3 for excellent, 2 for very satisfactory, 1 for satisfactory, and 0 for unsatisfactory. The unit's score was an average of its scores in the 10 inspection areas. The range of average scores for the batteries was from 3.41 to 1.00

Adverse Personnel Actions. This measure was based on episodes expressed by General Courts Martial, Special Courts Martial, Summary

Courts Martial, and entrances in the battery punishment book during the previous three months. The seriousness of the offense was weighted on a 4-3-2-1 scale. The resulting score might be called a "punishment score." A separate analysis was also made on number of AWOLs.

These six measures were applied to antiaircraft artillery units of three defenses—San Francisco, Seattle, and New York. For each defense the batteries were ranked from high to low, according to their scores on each of the measures. Relationships between the batteries' rankings on every possible pair of measures were evaluated by methods of statistical analysis known as the "correlation technique." These procedures were used to determine the presence and extent of any relationship or co-variation of the batteries on the various measures; they also permit identification of measures upon which the same battery tends to score high or low.

By using this means of analysis then, the following relationships or co-variations were identified:

When the battery rankings for *range of radar pickup* were paired with the ranking for the other measures, range of radar pickup proved to be related to the *defense commander's rating*. Range of radar pickup was also related to *radar maintenance* when maintenance was measured by either the second method (total number of job orders) or the third method (days out of action because of radar failure) but *not* when it was measured by the first method (repairs judged by ordnance personnel to have been necessitated by inadequate maintenance). Range of radar pickup was not related to firing scores, artillery maintenance, adverse personnel actions, or AWOL rate.

The *defense commander's rating*, when paired with the other measures, did relate to *range of radar pickup*, and to *radar maintenance* as measured by the second and third methods. In addition, the *defense commander's rating* was related to the AWOL situation, in that the greater the number of AWOLs, the poorer the rating given the unit by the commander. It did not relate to any of the other measures.

Artillery maintenance, personnel actions, and firing scores did not relate to radar maintenance, radar pickup, or commander's rating. AWOL did not relate to any measure other than commander's rating.

Thus three ratings—radar maintenance, when measured with certain procedures, range of radar pickup, and the defense commander's rating—are related, in that when a battery is high on one it is likely to be high on the other. With the exception of AWOL, the other measures do not show any relationship to these three, nor do they relate to each other.

The relationship between defense commander's rating and AWOL is particularly interesting when one considers that the AWOL rate did not relate to any of the operational measures. It suggests that a commander considers AWOL rate when making a unit's composite rating, but that AWOL is not one of the factors determining the unit's actual operational performance.

Analysis of the techniques of leadership and administration in the units should clarify the lack of co-variation or correlation between some of these measures of performance. Data on leadership and administration practices were collected while battery performance was being evaluated, and practices which identify highly efficient units and inefficient units were described.

Results of this phase of the study may be summarized thus:

Available measures for comparison of the performance of antiaircraft units have been shown to be reliable and accurate.

In regard to efforts to predict unit efficiency, this point has been established: The fact that a unit rates well on one performance measure does not imply that it will have a high rating on all such measures. A degree of generalization is justified from range of radar pickup to radar maintenance and the way a commander ranks his units. There is nothing, however, to suggest that a unit that ranks high on these three measures will rank high on range firing scores, personnel action indices, or maintenance other than radar.

Future evaluations of unit performance should take these findings into account. A "good" unit is a composite of many attributes; often a unit may have some of these attributes and not have others. Performance of any given type should be evaluated by observing that particular activity. For the most part, a separate measure should be used for each activity a commander considers important for assessment.

CREW DESCRIPTION DIMENSIONS AND RADAR CREW EFFECTIVENESS[1]

Thomas I. Myers and Francis H. Palmer

Performance effectiveness of 30 U.S. Army antiaircraft radar crews and their leaders is studied through the use of group dimension variables of Harmony, Intimacy, Procedural Clarity, and Stratification. The crew effectiveness criterion was a reliable (.87) measure of their ability to locate target aircraft. It was found that the leader's Stratification rating of the crew correlated highly with group effectiveness; a possible non-linear pattern was seen in regard to the crew's rating of the degree of Intimacy within the group.

This paper presents results pertaining to group dimension variables, specifically, four of the scales found on the Crew Dimensions Description Questionnaire (CDDQ), developed by Hemphill, Rush, and associates at the Personnel Research Board, Ohio State University.[2] These measures were examined in terms of several criteria of sensitivity and in terms of validity in relation to an operational criterion of radar crew performance.

The commonplace observation that groups *as groups* seem to differ from one another simply marks the behavior of groups as a bona fide subject matter of scientific scrutiny. From this single point of departure--that of a common general problem—distinguishable approaches to this problem appear, generated perhaps as much by differences in emphasis and purpose as in conceptualization. One straightforward method of approaching these group behavior phenomena is to attempt to dimensionalize some of the attributes in respect to which groups seem to vary. The fruits of this method would be measuring devices, using as subjects members or observers of the group. There are no magical means whereby the "real" or, in fact, any variables become specifiable. Thus, even here *a priori* "hunches" must precede, and, to some degree, determine the end product. Nor is such an emphasis on a descriptive device to be confused with a more ultimate causal schema. However, interest in the direct attempt to measure group dimensions certainly

---

[1]This paper was presented by Dr. Myers at the American Psychological Association convention, San Francisco, California, September 1955.

[2]John R. Hemphill, Carl F. Rush, Jr., *et al.*, research on air crew composition and performance in operations and combat, under U.S. Air Force contract.

does not preclude attempt at theoretical formulation. Such an approach *does* offer practical tools for exploratory stages of research.

The Hemphill group sought to develop measures of group dimensions by abstracting from the literature of social science some 14 dimensions of group behavior. Subsequently, items relevant to these *a priori* dimensions were created and subjected to item analysis selection procedures. The end product was a 150-item questionnaire covering 13 dimensions. In the HumRRO study, undertaken at Fort Ord, California during the planning stages of a large correlational research task, it was decided that several of the CDDQ scales might be applied to small intact Antiaircraft Artillery radar crews; Harmony, Intimacy, Procedural Clarity, and Stratification were the four dimensions chosen.

The present study, one aspect of the larger research, presents data about these dimensions in terms of a series of questions that fall under two general headings: (a) Characteristics of the CDDQ measures as independent variables and (b) Validity of the measures. The questions pertaining to the former consideration include:

Are the tests reliable?

Do the four scales differentiate between crews?

What are the interrelationships between the dimensions?

What degree of relationship exists between the ratings of leaders and their subordinates?

The questions pertaining to validity of the CDDQ scales include:

Are crew ratings correlated with a reliable group performance criterion?

Are leader scores related to group effectiveness?

Are there interaction effects between leader and follower scores which consistently relate to group productivity?

By way of procedure, the leader and subordinate members of 30 U.S. Army Antiaircraft Artillery radar crews were individually administered an untimed questionnaire consisting of modified items from the Harmony, Intimacy, Procedural Clarity, and Stratification scales of the Crew Dimensions Description Questionnaire. The *Harmony* scale items seemed to measure the degree of smoothly cooperative functioning of the group; *Intimacy* referred to the extent of personal familiarity among group members; while *Procedural Clarity* indicated how clearly defined and understood were the activities of the group. *Stratification* appeared to estimate the tendency toward hierarchy within the group in terms of ambition, influence, and prestige.

Basic data for the study were secured by deriving, for each scale and every crew separately, the average score of subordinate crew members and the individual score of the crew's leader. Thus for each of the 30 groups, there was a leader score and a mean crew score on each of the four Crew Dimension scales.

Independently of the survey testing phase, data were collected from each military unit to provide a crew performance criterion—crew ability to locate target aircraft. This measure of unit productivity was found to be reliable with the obtained odd-even coefficient of equivalence equal to .87. It was against this criterion—crew effectiveness in picking up targets on their radar sets—that the various CDDQ scores were compared in the validity analysis.

In describing the results, we shall first consider the CDDQ measures as *Independent Variables*. To estimate the reliability of individual CDDQ scores, odd-even coefficients of equivalence were obtained from a random sample of 140 crew members. The scales and their reliabilities were: Harmony .82; Intimacy .71; Procedural Clarity .84; Stratification .45. These findings indicate that the traditional reliabilities of the scores were generally high, with that for Stratification running somewhat lower (Table 1A).

Table 1

Characteristics of the CDDQ Measures as Independent Variables

(N = 140 Crew Members)

A. Reliability of Individual CDDQ Scores

Scale	Coefficients of Equivalence (odd-even)	
	Number of Items	Coefficient
Harmony	18	.82
Intimacy	6	.71
Procedural Clarity	6	.84
Stratification	4	.45

B. Differences Between Crews on the CDDQ Dimensions

Scale	Four Single Analyses of Variance of Differences Between Crew Means	
	F Ratio	p
Harmony	2.51	<.001
Intimacy	1.84	<.025
Procedural Clarity	2.16	<.005
Stratification	2.40	<.005

To evaluate the significance of difference between crew mean scores on the dimensions, four simple randomized analyses of variance were run. The findings showed that crews differed markedly from one another in their self-evaluations on these dimensions (Table 1B).

In order to examine the relationships between dimensions, product moment intercorrelations were calculated for two different sets of scores: Leader scores and crew mean scores (Table 2). The two resulting matrices were similar, showing a negative correlation between Harmony and Stratification. All other correlations were insignificant. Thus the CDDQ dimensions proved to be independent in this empirical study, with the exception of Harmony and Stratification.

Table 2

Intercorrelations Between CDDQ Dimensions

Scale	Intercorrelations for Leader Scores ($N = 30$)			Intercorrelations for Crew Scores ($N = 30$)		
	Harmony	Intimacy	Procedural Clarity	Harmony	Intimacy	Procedural Clarity
Harmony						
Intimacy	.26			.16		
Procedural Clarity	.27	.08		.23	.32	
Stratification	−.60*	.10	−.27	−.38*	−.24	−.10
	*$p<.01$			*$p<.05$		

To measure agreement between leader and subordinates in evaluating the group, the method was simply that of correlating the scores of the crew and its leader, for each of the four scales (Table 3). These correlations indicated close agreement between leader and crew in rating Procedural Clarity and Stratification. No significant agreement was observed for Harmony and Intimacy.

Turning to the validity of the measures, the analysis proceeded in two somewhat parallel ways. Pearson correlation coefficients were computed between scores and the performance criterion, crew ability to locate target aircraft. In addition, a three-group analysis of variance procedure was employed, to allow for the possibility of curvilinear relationships. In this latter method, individual leaders (or crews) were categorized into High, Medium, and Low groups with reference to their scores on the independent variable. Arbitrarily, high and low groups were defined as 27% extremes. Group effectiveness scores of the three groupings were then analyzed by simple analysis of variance (Table 4A).

Table 3

Agreement Between Leader and Subordinates in Evaluation of Group

Scale	Product-Moment Correlations Between Leader Score and Mean Crew Ratings ($N=30$)
Harmony	.27
Intimacy	.01
Procedural Clarity	.59*
Stratification	.38**
*$p<.01$, **$p<.05$.	

Considering CDDQ scores of the leaders, it was found that the leaders' Stratification rating correlated significantly with the criterion. The other leaders' ratings did not relate to group performance (Table 4B).

Table 4

Relationships Between CDDQ Scores and Crew Ability to Locate Target Aircraft

A. Correlations Between Crew Performance and Crew Mean CDDQ Scores ($N=30$)		B. Correlations Between Crew Performance and Leader CDDQ Scores ($N=30$)	
Scale	Coefficient	Scale	Coefficient
Harmony	−.11	Harmony	−.17
Intimacy	−.04	Intimacy	.14
Procedural Clarity	.12	Procedural Clarity	.08
Stratification	.24	Stratification	.42*

*$p<.05$

In respect to crew CDDQ scores, it was found that none of the dimensions correlated with effectiveness. The analysis of variance showed differences at the .05 level between criterion effectiveness of crews with high, medium, and low intimacy ratings. The pattern of this effect seemed to be non-linear. *Outstanding performance* was associated with crews who rated their group as *moderately high* on the Intimacy dimension. Crews with more extreme self-rating scored more poorly on the effectiveness measure (Table 5).

Table 5

Analyses of Variance of Crew Performance Scores

Scale	Four Analyses of Variance of Crew Performance Scores Where Crews Were Categorized Into Three Groups (27% in extremes) on the Basis of Crew Mean CDDQ Scores			Four Analyses of Variance of Crew Performance Scores Where Crews Were Categorized Into Three Groups (27% in extremes) on the Basis of Leaders' CDDQ Scores		
	F Ratio	*p*	Apparent Relationship	*F* Ratio	*p*	Apparent Relationship
Harmony	<1	*NS*		<1	*NS*	
Intimacy	3.86	<.05	Curvilinear*	3.33	*NS*	
Procedural Clarity	<1	*NS*		<1	NS	
Stratification	2.52	*NS*		5.47	<.025	Linear**

*High performance associated with medium Intimacy; low performance associated with high and low Intimacy.

**Positive relationship: low, medium and high performance associated with correspondingly low, medium, and high levels of Intimacy.

The final step in the validity analysis involved testing for any leader-follower interaction effects that might be consistently related to group effectiveness. There were two scales—Harmony and Intimacy—on which leader and follower ratings of the same group did not agree. For these two dimensions, a factorial analysis of variance was run, considering leader scores and crew scores as categorical classification variables. For example, a crew was assigned to a cell of the table by jointly considering the leader's and crew's Harmony score. The scores to be analyzed were our group effectiveness measures.

Leader-follower CDDQ score interactions were found to be insignificant in relation to productivity. The interaction F ratios for both Harmony and Intimacy analyses were small.

Assessing these validity findings is necessarily an interpretational problem, due to the correlational design of the study. It is also of interest that the leader's Stratification rating is our nearest analog to Fiedler's ASo score[1] of sociometrically defined leaders, which has proven to be an effective predictor of group performance. It, too, appears to involve the extent to which a person perceives and/or recognizes differences between group members.

Our finding that perceived Intimacy to a high degree is associated with poorer performance, fits rather nicely with some of the sociometric data from the present study. It was found that highly effective crews tended to choose off-duty pass companions from outside their immediate group. The Intimacy scale relationship tends to strengthen the interpretation that strong emotional allegiances between group members may be inimical to at least certain types of group performance.

To summarize these findings, it was shown that the four CDDQ scales were generally reliable, that they differentiated between crews, that with one exception, the dimensions were empirically independent, and that leader and follower agreement was high with regard to Procedural Clarity and Stratification, but not with regard to Harmony and Intimacy.

In terms of these various characteristics, then, the CDDQ scales were, in our experience, highly satisfactory instruments. Summarizing the validity analysis in reference to a crew performance criterion—ability to locate target aircraft—it was found that the leader's Stratification rating of the crew correlated highly with group effectiveness, and that a possible non-linear pattern was seen in regard to the crew's rating of the degree of Intimacy within the crew.

[1]Assumed Similarity of Opposites.

SOCIOMETRIC CHOICE AND GROUP PRODUCTIVITY AMONG RADAR CREWS[1]

Francis H. Palmer and Thomas I. Myers

This study is part of a research effort to identify the human factors that differentiate effective from ineffective antiaircraft batteries by studying variables that contribute to group productivity in large military units. Radar crew members were asked to make choices for off-duty association. It is suggested that congeniality and social interaction are negatively related to group productivity.

The research reported here is part of a larger HumRRO study to identify the human factors that differentiate outstandingly effective antiaircraft batteries and ineffective batteries. The purpose of this study was to identify variables in large military units that contribute to group productivity.

The locale of the study, the antiaircraft battery in the Zone of the Interior, is an ideal research climate. These units are the same in size, have approximately the same caliber personnel, the same equipment, and the same mission. Each individual unit is relatively isolated from community and military affairs. Very little social congress exists with other military installations.

The data reported will refer specifically to those parts of the study that related interpersonal relationships of the radar crew and their productivity on a single criterion of performance.

When the amount of positive interpersonal relationships among members of small groups has been related to productivity, different studies have obtained paradoxical results. In some cases productivity has been associated with congeniality and increased social interaction, and in others it has been demonstrated that such congenialities are detrimental to productivity. For example, Roethlisberger and Dickinson have shown that production on an assembly task was increased when social interaction and interpersonal friendliness were encouraged among a group of girl workers; while Fiedler, working with basketball teams and other groups, has concluded that warm interpersonal relationships in these groups are detrimental to productivity. Such studies

---

[1]This paper was presented by Dr. Palmer at the American Psychological Association convention, San Francisco, California, September 1955.

can not be compared directly, partly due to lack of precision with respect to the definition of social interaction, and partly because the tasks on which the different groups studied were not the same.

It is the purpose of the present data to take a specific situation with a specific group and to determine the relationship between degree of social interaction and the productivity of the group. The radar crews of 40 antiaircraft units in a single defense were used. Each crew, consisting of 8 to 13 individuals, is a part of a larger unit—the battery—an organization of roughly 100 persons. The crew consists of three status individuals—the range platoon sergeant, the radar mechanic, and the chief radar operator—and subordinate members whose primary role is that of operating the equipment.

Crew productivity was determined for the principal operational activity of an operating crew, that is, range of radar pickup (RRP). Before a battery can engage a target, the radar crew must be able to identify, acquire, and lock on an aerial target. This complex team process is the crucial point in battery effectiveness. The more apt the crew is at locking on targets at a great range, the more effective the entire unit will be. Therefore, the measure of productivity in this instance was the average range of pickup for each of the 40 crews over a three months' period of locking on targets. One hundred and four strikes were flown on the defense over this period. The average range for each unit was, of course, the average number of yards in thousands of yards for each unit. When averages for the odd and even pickups for each unit were obtained, and the two distributions correlated, the reliability of this measure was shown to be .87.

Interpersonal relationships were measured by the simple sociometric device of asking each member in the battery what three persons in the battery he would most like to "go on pass with." Consequently, it was possible to determine the extent to which members of the radar crew chose among each other for off-duty association. This was calculated for each unit by counting the number of choices made by members of the crew to other members of the crew, and by dividing by the N of the total crew. Two other measures consisted of taking the three status individuals and determining the extent to which they chose their off-duty associates from their crew, and secondly, taking the subordinate members of the crew and determining the extent to which they chose other members of the crew. Three scores were computed for each unit: total score, score for status individuals, and score for subordinates. When these three scores were correlated with group productivity, as determined by range of radar pickup, there were these results:

(1) The extent to which the crew as a whole chose each other as pass companions was negatively correlated with performance on Range of Radar Pickup (RRP) (-.45); that is to say, the greater amount of positive choices from crew member to crew member with respect to the "go on pass with" variable, the poorer the performance.

(2) The extent to which the status members of the crew chose their off-duty companions from the crew was negatively correlated with performance on RRP (-.51).

(3) The extent to which nonstatus or subordinate members of the crew chose pass companions from the crew was negatively correlated with RRP (-.31).

With the RRP measure of group productivity, there is no question but that social interaction, if it be defined in terms of the sociometric variable used here, is negatively related to group productivity. There is some suggestion in the data that this is more relevant for status members of the group than for nonstatus members.

I would like to suggest an hypothesis to explain the different results that have been obtained by studies of this sort. It revolves around the notion that researchers have been preoccupied with what is going on in a group and not with what the group is doing. The idea of *natural* and *imposed* goals, sometimes referred to as extrinsic and intrinsic goals, suggests that the activities of human groups may be placed on a continuum. This starts from the truly *natural* goal, in which the individual members of the group belong to the group solely for the purpose of satisfying individual needs, to the *imposed* goal, which consists of tasks imposed upon the group by some external authority, with no role as need satisfiers to group members. An example of the natural goal might be a women's bridge club where interpersonal relations must be congenial for the individual members to obtain gratification of their particular needs. The imposed goals can be thought of as the radar crew's task of RRP. Where the goal is imposed on the group, high social interaction may create competing goals, a consideration that may be detrimental to the accomplishment of the group task.

It is suggested, then, that social compatibility and congeniality are positively related to group productivity only so long as the natural goals of the individual members and the defined goals of the group are the same, and that to the extent that the defined goal of the group is at odds with the natural goals of its members, social interaction will become more and more inversely related to productivity.

LEADERSHIP AND GROUP ACHIEVEMENT[1]

Francis H. Palmer

This article discusses research on leadership and leader training as related to multiple group goals and performance in the achievement of those goals. Although relating to research in a military setting, some implications for nonmilitary contexts are included.

"My main problems are with people." The lieutenant looked out over the tubes of his 120 millimeter antiaircraft guns pencilling above the busy crews of the battery. "My job," he continued, "is to make people perform so that the battery shoots as accurately as the weapons system permits."

In these words a young Army officer puts his finger on the crux of leadership situations today in which we are trying to achieve a maximum performance of man-machine systems. Those words, "make people perform so that . . . " are on the minds, not only of artillerymen, but also of factory supervisors, farmers, industrialists, and many kinds of teachers.

Getting people to perform in one way or another is, after all, leadership in action. But when one thinks in terms of a machine-tailored goal, it is no simple matter to bring human performance to the far limits imposed by modern "machine" environment.

The artilleryman has, like most leaders, a heritage to fall back on: experience, schooling, opinion, legend, war stories. All of these serve him. But in his intricate man-machine situation, he is still asking himself many questions.

What can he do to help his team achieve its mission? What are the best relationships between himself and his team for the fostering of achievement? Are these relationships constant, or does his role as a leader vary with the activity of the group and its composition at any given time?

Group Goals

To answer some of these questions—even as partially as we are prepared to answer them—we must leave the leader and think for a moment in terms of the *goals* of the group organization. It is often assumed that a group has a single goal, easily associated with a specified activity. The truth is, however, that human groups are seldom engaged in a single activity.

[1]This article appeared in *Adult Leadership*, vol. 5, no. 2, June 1956.

Even relatively simple groups such as a railroad section crew are usually pursuing several goals. These goals may include not only laying track for the railroad, but also any of the various individual satisfactions sought by members of the group. In fact, the Human Resources Research Office has designed studies that assume that to be proficient a group must excell in *several* activities and, moreover, that those human factors that contribute to high performance in one activity may be entirely *unrelated* to performance in other activities.

These two assumptions—as well as certain implications for leader-group relationships—have evolved from a project aimed at helping the antiaircraft lieutenant answer some of his problems.

To outline this project—first of all, the research task seemed to be to identify those human factors within antiaircraft batteries that differentiated particularly good and poor organizations. It proceeded along these lines:

1. A large number of senior Army officers and battery commanders were asked to enumerate those activities essential for the performance of the antiaircraft battery's mission.

2. Measures of these activities were developed so that the battery could be rated in a reliable and valid manner as it performed its duties.

3. Then, measures of certain human characteristics of the battery were developed. These included leadership roles and techniques used in the unit, the interpersonal relationships of the men, the personalities of various key personnel, and the background, attitudes, and level of morale of the unit's personnel.

4. A large sampling of batteries was then chosen and tested from two standpoints: from that of performance of its essential activities and from that of its human makeup.

5. Finally, measures of performance were correlated with human factors, thereby indicating which human factors were consistently present or absent in high achievement of the several activities essential for the success of the group.

Such a study might be expected to produce certain "obvious" conclusions. For example: that a unit composed of well-shaven, responsible, neat, intelligent men would naturally operate more efficiently in achieving its goals than one which was not; or, if the men were well grouped from the standpoint of friendship and ability to get along well—like the Harlem Globetrotters, perhaps—they would form a better "team" and, hence, could be expected to operate more effectively: in short, that if all the admirable human traits and relationships one can think of were somehow tossed into a mixture, we would have a close-to-ideal organization. Let us look at some of the results of the study.

Surprising Results of the Study

It was found that a unit which was particularly good at one type of activity, such as identifying and tracking aircraft with its radar, was no more likely to be good at maintaining its equipment than a unit which was poor in the operation of its radar; and that a unit whose Absence Without Leave (AWOL) rate was low was no more likely to maintain its equipment well or operate its radar well than a unit with a high rate of absences.

It was found that human characteristics that related to one type of activity were seldom related to another. Indeed, on occasion, a characteristic related to desirable performance on one task was negatively related to desirable performance on another!

It was found that radar crews whose members more often chose one another as off-duty companions performed more *poorly* when operating their equipment than those crews who limited their associations to duty hours.

What, exactly, do these findings show? That it doesn't make a bit of difference whether equipment is properly maintained, or whether a unit's AWOL rate is high? Or that it really doesn't matter whether teammates are compatible? Not at all.

Although each of these factors has, of course, important implications outside the scope of this study, the findings do show that we must be careful in generalizing about group performance on several activities simply because the group has performed well on one. They also show that there *are* work situations in which congeniality among group members does not contribute to proficiency on specific tasks.

Furthermore, they suggest that we need to learn a good deal more about families of human activities—that is, types of group tasks that require the same interpersonal background. Only when this is known can we begin to generalize safely from the results of one study about those factors that will contribute to performance on tasks other than those concerned in the study.

Applications of Leadership Principles

Finally, the present results suggest hypotheses that may have bearing on the training of young leaders. Perhaps we can now say to the young leader that his behavior might well vary with the specific activity in which his organization is engaged, and that in the usual case where the organization is involved in several activities, the appropriate role should be tailored to the activity that he considers most essential.

We might also say to him that the degree to which he should maintain social distance between himself and his followers will vary according to whether the unit's task is more closely associated with the satisfaction of *individual* needs among unit members or with more *impersonal*—or externally imposed—goals. In other words, if the primary goal, at a particular time, is to erect a unit social center, he should be more familiarly involved with his men; if the primary goal is a specific military mission, he should maintain relatively greater social distance.

We may say to him, too, that in certain activities his leadership will not be as effective or as important as the leadership of his subordinate officers. For example, the performance of radar crews seemed to be most closely related with the characteristics of the crew leaders, while attitude and morale measures were more closely associated with the actions of the battery commander. He must realize, then, that at all times he does not *directly* play the most important leadership role for all segments of his organization.

These findings are at once general and specific. They are derived from a military situation, but it is likely that they apply as well to many public, commercial, vocational, and educational situations.

It is fundamental tenet, of course, that production, working hours, personnel qualifications, and plant operation in general are all geared to the goals of the organization. Now we should begin advising our young leaders to consider carefully the specific goals of their organizations before they initiate leadership activity. It boils down, then, to a generality with specific applications. In time, the researcher hopes to whittle down the generality itself. In the meantime, it is a direction in which we can confidently travel.

HUMAN FACTORS AFFECTING THE PERFORMANCE OF ANTIAIRCRAFT BATTERIES[1]

Francis Palmer, Thomas Myers,
Bertram Gold, and Paul Metzger

Operational performance in Range of Radar Pickup and Radar Maintenance measures by antiaircraft batteries and the Defense Commander's Rating practices in 40-on-site defense organizations were studied. Leadership techniques, battery practices, and interpersonal relationships were studied to determine the extent to which human factors served to discriminate between high and low efficiency units. Implications for personnel assignment and for training are given.

INTRODUCTION

Getting the most out of a weapons system depends not only upon the machine, but also upon the men operating the machine. The knowledge, training, and organizational structures most appropriate for the optimal operation of a new weapons system frequently lag behind its technological development. This means that the potential performance of a weapons system can be described on several different levels.

One such level of performance is the potential of the weapons system, as limited by its own specifications (machine potential). A second level is the performance attained by the weapon when it is operated by personnel (man-machine potential). Dissimilarity between machine potential and man-machine potential can be regarded as being largely due to human factors.

This paper is concerned with the relationship of certain human factors to performance of antiaircraft (AAA) batteries. Specifically, suggestions are made for ways of measuring and improving performance; that is, ways of narrowing the gap between units that perform well and those that perform poorly. However, the main emphasis is on comparing existing antiaircraft batteries and isolating those factors associated with variations in unit performance.

It is obvious that the commander of an antiaircraft battery cannot obtain better performance—that is, shoot higher, more accurately, or more rapidly—than the machine potential of his equipment permits; although there are many forces beyond his control that limit his ability

[1]Summary task paper prepared in March 1956.

to effect change, he *can* organize and develop his personnel so that the difference between *actual* and *potential* performance is minimized.

Population Studies

There were several reasons for selecting on-site antiaircraft batteries in the Zone of the Interior for this research. First, these batteries are among the many military organizations whose performance is dependent on the interrelationships of man and machine. Second, there are a large number of batteries in the several antiaircraft defenses, each having the same mission, the same equipment, and similar personnel. Third, and perhaps most important, the commanders of several of the defenses indicated an awareness of the effect of human factors on unit efficiency and expressed interest in the possible results of this study. The research would have been impossible without their assistance.

The number of battalions that compose the several defenses varies considerably, but the organization of the fire unit itself, the gun battery, is basically consistent from defense to defense. The geographical location of a battery is, of course, a function of the terrain where the defense is organized. Some defenses have units on mountain tops in highly remote situations; others have some of their batteries in the middle of city parks.

The gun battery Table of Organization states that a captain will command the unit, but only about one of every four units sampled had an officer of that rank, and fully 50% were commanded by second lieutenants. The commander may or may not have had one or two other officers in his unit. In short, most batteries were understrength in officer personnel and many of the officers lacked experience.

The noncommissioned officer situation was similar. Seldom did a unit have its authorized complement of first three graders. The remaining positions were filled by corporals who were draftees and did not have sufficient in-grade qualifications for promotion. The battery commander who could fill six or seven of his top 12 positions with regular Army NCOs was rare.

Personnel turnover during the year previous to the collection of these data was high. Most of the officers felt that this was particularly harmful to crew-type performances, and most severe when these performances were highly technical as, for example, in the radar crew.

For the most part, the physical environs of the batteries were comfortable. With few exceptions all units had permanent mess halls and prefabricated barracks and latrines. Site improvements by unit personnel varied greatly and may or may not have been related to an organization's state of operational readiness.

In every case the unit was equipped with the M33 radar, either the 90mm or the 120mm gun, and the appropriate power supply. Although all of the radar equipment was relatively new, the number of operational hours varied from battery to battery.

Approach to the Problem

The design of the study provided for (a) the identification of those battery activities essential to performance of the unit's operational mission, (b) the development of the reliable measures of these activities, and (c) the administration of tests and evaluative devices to unit personnel. The last activity permitted the determination of details concerning leadership techniques, battery practices, and interpersonal relationships. When both the measures of performance and the measures of human factors were available, the extent to which the human factors served to discriminate between high and low efficiency units was evaluated.

This isolation of the human factors that co-varied with unit efficiency was intended, primarily, as a basis for suggesting hypotheses about the central determiners of group effectiveness, whose causal relation to performance might then be subjected to appropriate rigorous experimental test. In addition, it was clear that the descriptive information in regard to discriminating human factors might well serve as an interim aid to the military. This paper has this latter function as its goal.

THE PERFORMANCE OF BATTERIES

Like most work groups, the members of the antiaircraft artillery battery perform several activities in the course of their duties. These include locating and locking on targets with the M-33 radar, maintaining radar and artillery equipment, meeting the demands of superior headquarters, and having adequate administrative-personnel procedures. These activities are integral to the life of any unit, and a problem arises respecting the extent to which success in one activity is related to success in another. If the battery is good at picking up targets with its radar, will it be good at maintenance and its other functions?

Large organizations, military and civilian, have a continuous problem of assessing their own proficiency. Commanders characteristically acquire and act upon cues that they believe are indicators of the performance level of the group. If the group's mission is complex and the performance difficult to measure, supervisors are gradually compelled to use cues as short-cut methods for evaluating performance. This is particularly evident when the product of a group cannot be measured in terms of dollar profit, number of units manufactured, or some other direct and simple index.

As an organization grows older, it is more likely that certain cues will be passed down by senior members and will, as a consequence, become traditional. Because of such factors as changes in organizational mission or technological or social development, these traditional cues may become invalid. Consequently, traditional means of estimating performance must be periodically examined to determine whether they are, in fact, accurate indicators of the performance mission. If the cue is not an appropriate index for assessment, its use yields an unrealistic evaluation of unit worth and is probably more damaging to subsequent proficiency than no evaluation at all.

An example of such a cue becoming traditional is Absence Without Leave. Many commanders use this measure as an evaluation of a unit, and often generalize about the overall unit performance from it. The data below show that this cue is not related to any of the other performances measured in this study. Figure 1 shows the relationships between AWOL and other measures of performance taken for 40 antiaircraft batteries.

Relationships Between AWOL and Performance Measures

(N = 40 Antiaircraft Batteries)

Performance Measure	RRP	RM	AM	DCR	AWOL
Range of Radar Pickup (RRP)					
Radar Maintenance (RM)	None				
Artillery Maintenance (AM)	None	None			
Defense Commander's Rating (DCR)	Positive	None	None		
Absence Without Leave Rate (Low)	None	None	None	Positive	

Figure 1

It was found that while AWOL rate was unrelated to how well a unit performed on Range of Radar Pickup (RRP), Radar Maintenance, or Artillery Maintenance, still a low rate was positively associated with the Defense Commander's Ratings of his unit. Also, when the unit was good at Range of Radar Pickup, the DCR tended to be high. But RRP was not related to either maintenance measure, nor was DCR; Radar Maintenance and Artillery Maintenance were not related to each other.

The most important result shown by these data seems to be that batteries are involved in several activities, all of which are important to the final military goal; and that there is no basis beyond the limits of the relationships shown for assuming that a unit will be good on one performance because it is good on another. If commanders wish to assess maintenance, then maintenance measures should be developed and used; if they wish to assess the operational ability of the radar crew, then measures specifically designed for that purpose should be developed.

RANGE OF RADAR PICKUP

The operational effectiveness of an antiaircraft gun or Nike battery is greatly influenced by the range at which members of the radar crew are able to locate and lock on incoming aerial targets. This chapter deals with the problems of (a) developing a measure to evaluate the comparative performance of radar crews in respect to Range of Radar Pickup (RR)), (b) describing the difference between the mechanical potential of the M33 radar and the actual man-machine potential of a group of such weapons systems under field conditions, and (c) explaining the human factors in radar crews that relate to variations in their performance. Finally, the implications of these data, with respect to training, selection, motivation, and morale, were considered.

Developing a Performance Measure

The antiaircraft defense commander usually deploys his batteries in a circular or elliptical manner around a designated vital area. Consequently, the several radars in the defense are situated on varying terrain and any simulated raid on the defense by real targets affects each unit differentially as a function of its position. One of the problems in measuring crew performance, then, is that of controlling or allowing for the effect of the battery's geographical location on the Range of Radar Pickup. Other problems involve the control of such factors as *clutter* (electronic interference) and *masking*.

It should also be stated that maintenance of equipment—in the sense of day-to-day adjustment of the radar set—is a variable in the measurement of Range of Radar Pickup; that is, it is assumed that some crews are better than others. However, maintenance (so defined) is such an integral part of crew operation that the most valid measure of pickup performance should include its effects. Consequently, an effort was made to measure or control the frequency and caliber of such adjustment.

Measurement of crew performance was accomplished as follows: over a three-month period 104 scheduled aircraft approaches were made on the defenses. Three or four such approaches ("strikes") made on a particular day are referred to as a "mission". The radar crew of each battery was instructed to participate in all missions except when it was involved in field problems or at the firing range.

A "track map," showing the actual course of the targets over the defense, was maintained by the Antiaircraft Operations Center for each strike. Also, for each strike, a *battery report* filled out by an officer in the firing unit was forwarded to the Brigade Radar Officer. This report included the date, the strike number, the location of site, and the coordinates, slant range, altitude, and azimuth describing the position of the target at the time the locking-on process was completed by the radar crew.

Consequently, information was available indicating the distance at which the *unit* reported it had locked on the target at a particular time, and the distance at which the *operations center* fixed the target at that time. When these plots were correlated within certain tolerances, the distance from the individual site to the target at the lock-on point was measured, and the unit given a score in thousands of yards for that strike.

An average range of radar pickup was derived for each battery over a three-month period. The coefficient of reliability for the measure was a very satisfactory .87. Further analysis indicated that such factors as clutter, masking, and the number of strikes in which a unit participated over the three-month period were not related to the final score on RRP; that is, units were not penalized or rewarded by the influence of these factors.

Machine Potential Versus Man-Machine Potential

Although data on the mechanical potential of the M33 Fire Control System and on the performance anticipated for planning purposes are not relevant to this paper, the magnitude of performance differences between high-scoring and low-scoring batteries is important. The average performance of the top one-third of the batteries was 30,000 yards better than the average performance of the bottom one-third (Figure 2). The range of performance between the highest and the lowest scoring batteries was 50,000 yards. *Thus, human factors, differences in the effectiveness of human beings in different batteries, appear to be responsible for significant differences in RRP and, hence, in the operational success or failure of the antiaircraft defense.*

Differences in Actual Range of Radar Pickup Performance for 40 Antiaircraft Batteries Within a Single Defense

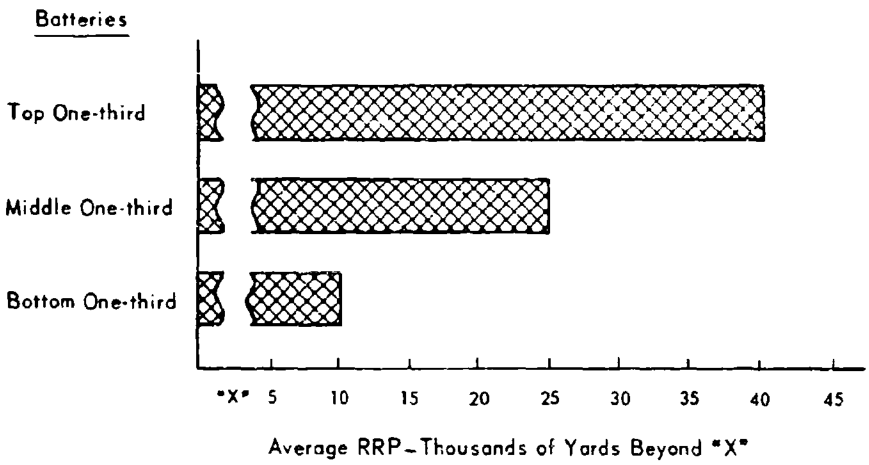

Figure 2

The Measurement of Human Factors

Specific human factors thought to be associated with variations in battery performance were identified by measures yielding data on the intelligence and education, interpersonal relationships, morale and motivation, leadership techniques and group structure, personality, and background characteristics of the battery as a whole and the radar crew in particular. Tests, rating scales, and other instruments were used to measure characteristics of battery members. A comparison of the results of these scales and tests, and the measures of Range of Radar Pickup, indicates the relationship between specific human factors and performance.

Scores were determined for each of the individuals tested. Then, radar crew scores were derived by determining the average score of each member. These averages, as well as the individual scores of key personnel, were correlated with crew performance on RRP. The following section summarizes characteristics that vary consistently with crew performance.

Characteristics of Crew Leaders

Radar crew leaders—the Range Platoon Sergeant, Radar Mechanic, and Chief Radar Operator—in crews that performed well on RRP were found to differ from leaders in poorer crews in several ways.

Personality Traits. When crew leaders were comparatively more assertive in social situations and placed higher value on perseverance and efficiency as goals, crew performance on RRP was better. Other personality dimensions, such as enthusiasm, preference for dealing with people rather than objects, emotional maturity, and inclinations toward authoritarianism, did not discriminate between leaders of good and poor crews.

Intelligence. Leaders of good crews received higher scores on intelligence tests than did leaders of poor crews. The average intelligence of subordinate members of good crews was no higher than that of poor crews.

Interpersonal Relationships. Leaders of good crews showed a strong tendency to choose "pass companions"—men with whom they spent their off-duty time—from battery members other than the radar crew. Those in poor crews more often chose off-duty companions from their own crew.

Morale. A tendency existed for leaders of good crews to rank their batteries higher when they were compared to other batteries in the defense. Attitudes about the caliber of their officers, as well as attitudes toward the unit's food or pass policies, did not discriminate between the leaders of good and poor crews.

Generally, the leaders of good crews were different from leaders of poor crews in that they were more intelligent, were inclined to value perseverance and efficiency more highly, and were more assertive in interpersonal relations. They maintained a greater social distance (as indicated by their choice of off-duty companions) from the subordinates than did leaders of poorer crews.

The leader whose intelligence scores related most highly to crew performance was the Chief Radar Operator. Other results suggest that men in this position influence crew performance more than any other.

Furthermore, it should be noted that the behavior of the Range Platoon Sergeant, as perceived by crew members, on such variables as his Initiating Structure (how well he defined and organized procedures and work for the crew) and Consideration (the degree of personal consideration he showed crew members) did not relate to crew performance on RRP.

Characteristics of Crew Subordinates

The characteristics of the non-status members of radar crews—all but the three leaders mentioned above—were, in several instances, related to proficiency in RRP.

Personality Traits. The tendency for non-status members of good crews to value perseverance and efficiency as goals in themselves differentiated them from their equals in poor crews even more effectively than it differentiated their leaders. They were also inclined to show greater will control (as characterized by independence of action) than subordinates in poor crews. Other personality measures did not differentiate between the crews of good and poor units.

Intelligence. While the average intelligence of crew leaders related to performance, the average intelligence of non-status crew members did not. Crews in which subordinates had, on the average, higher General Technical (GT) Aptitude Area I scores performed no better on RRP than crews with lower average intelligence. Similarly, no relationship existed between the average educational level of subordinate crew members and their performance.

Interpersonal Relationships. As was true of their leaders, the non-status members of good crews were more inclined to choose off-duty companions from battery members outside the crew, while poor crews tended to choose companions from within the crew. Also, members of good crews were more often chosen as companions by non-crew battery members than were members of poor crews, but were less often chosen as *most valuable* to the battery. (Battery member choices for most valuable crew member, in the case of the good crews, were directed more toward the status members than the non-status members.)

To some extent, then, the characteristics that differentiate the subordinate members of good and poor crews are the same as those that differentiate their leaders. Good crews are endowed with the tendency to value perserverance and efficiency—perhaps related to a "sense of duty"—and they tend to choose non-crew personnel as off-duty associates.

Although the intelligence of leaders related to performance, the intelligence of subordinates did not. When considering this finding, one must remember that radar crews as a whole were found to be a more intelligent group than other battery crews. Within the limits of this selected sample, then, intelligence does not relate to performance. The responses of individual members in describing their crew showed that such factors as the degree of harmony, the clarity with which procedures were defined, the degree of group identification, and relative stratification—all within the crew as measured—did not relate to this performance. In other words, if there are such factors that demonstrably influence performance, they are either different from those measured or more subtle than the instruments and consequently were not detected.

Members of good radar crews were more often chosen as off-duty companions by other battery personnel than were members of poor crews. This result, coupled with the previous findings of greater extra-crew choice among the good crews, may bear on the general problem of the level at which cohesiveness, in the sense of friendliness, is important to a military unit. It would appear that the batteries whose crews perform best are cohesive at the battery level rather than at the crew level. One explanatory hypothesis would be that when crews associate more with members of the larger unit, a greater understanding may be gained of the roles played by the several sections in the battery. That is, better interpersonal communication between sections, resulting from off-duty association, may increase the likelihood of efficient performance.

A second hypothesis that may apply to this finding—it is not exclusive of the first—would be that those crews whose members associate with one another most off-duty, may develop goals as a group associated with off-duty experiences. These goals compete with the operational goals of the crew on duty. Consequently, crews with comparatively less off-duty association may be more likely to expend their energies toward the military goals of the crew.

When the senior NCOs in the battery, the section leaders, were considered in terms of the extent to which their longevity in the battery was related to performance, it was found that length of time in battery for this group was positively related to improved RRP. Time in battery for crew leaders and crew subordinates did not relate to RRP.

The individual characteristics of the battery commander showed no relationship to crew performance on RRP. The attitudes of senior NCOs toward the commander's methods of operation were unrelated to performance. This fact suggests that the leadership most important to the radar crew and its proficiency on RRP is *within* the crew itself, and that either the actions of the battery commander are somewhat remote from this operation, or there is not sufficient difference in the methods used by various commanders to affect performance markedly.

Implications for Planning and Evaluating Operational Performance

The measure of RRP performance developed in this study is statistically reliable and, by the judgment of relevant authorities, is valid. Therefore, a technique is available that will enable commanders to evaluate the relative and absolute performance of their units at a given time and to determine such specific factors relevant to defense readiness as:

(1) The average pickup for each unit, when the average is based on *all* opportunities with zero scores allowed for no pickup.
(2) The average range of pickup for each unit in those instances where pickups were made.
(3) The incidence of locking on and reporting wrong targets.
(4) Sectors within the defense where responsible units as a group are particularly effective or ineffective.
(5) The probability of pickup at a given range and azimuth for those units that could effectively engage the target.
(6) The percentage of strikes when one, two, or three units could engage a target.

Selection of Personnel

The average Aptitude Area I scores of the three senior radar people were related to performance on RRP. This was especially evident for the Chief Radar Operator. Efforts to select personnel for these positions, particularly the Chief Radio Operator, from those with high intelligence might contribute to improved performance.

However, while among subordinate crew members no relationship existed between intelligence and performance, it does not follow that intelligence is unimportant to performance. Radar crews were demonstrably more intelligent than other battery personnel, but the average intelligence of the subordinate members of the several crews did not vary from crew to crew enough to warrant action on the findings.

In the selection of radar personnel, the results suggest that individuals who are particularly endowed with a "sense of duty"—who value perseverance and efficiency as goals in themselves—are most productive at RRP. Lacking a test score, a unit commander would find this particular selection difficult unless he knows his men very well. However, as commanders become more familiar with the characteristics of their men—particularly if they are alert to specific characteristics—selection of radar personnel in these terms may become possible under field conditions.

Implications for Personnel Assignment

Findings which showed that crews performed better when the senior NCOs had been in the battery longer have implications for personnel assignment. With the introduction of Operation Gyroscope, the Army has made substantial progress toward stabilizing personnel, but at the present time the benefits of this program have not been extended to antiaircraft batteries in the Zone of the Interior. Even if a policy for the stabilization of all personnel were impossible, emphasis might very profitably be placed on the stabilization of the Senior NCOs.

Implications for Training

On-site training of radar personnel is usually emphasized by unit commanders and their superiors. Present findings show that, beyond a certain point, no relationship exists between performance and the time spent by the average crew member, or his leaders, on the radar set. It is suggested that the process by which existing crews reach their best performance under present conditions is not so complex as generally thought, and evidently is one in which a comparatively high peak performance is reached quickly.

Two interpretations of this fact may be offered: the first, that crews may reach their best performance quickly; the second, that obstacles to learning may arise fairly rapidly and deter the crew from reaching proficiency at levels above those shown in this study. That is, a crew can become as proficient as possible in a few months, under present conditions.

One of the most striking results of this research concerns the *actual* RRP average of crews compared to how well they *think* they are doing—crew estimates are almost always unrealistic. It is estimated that personnel (at most levels) tend to overestimate RRP performance by 20,000 to 40,000 yards. Their estimates appear to be based upon recall of outstanding pickups; they tend to forget incidents when no pickup was made or when the target was acquired too late for action. The figure reported, then, is often a reflection of "best effort" rather than "average effort."

One of the most important potential influences on performance is adequate proficiency measurement. Accurate information about unit capabilities can be made available to operators. Knowledge of results is a powerful incentive for learning. This fact has been repeatedly demonstrated in experiments with the human learning process. Yet, during the period of this study, commanders and crews were not accurately informed of their absolute or relative level of performance. Quite possibly, higher commanders were not familiar with actual levels of performance. If a reliable and *appropriate* measure of the performance is used, and if the results are fed back to the participants, interbattery competition may well become a strong motive for high performance.

The Special Devices Center has contracted for equipment capable of simulating most of the relevant factors associated with the pickup process of the M33 radar. If this equipment meets the military characteristics of the contract specifications, an outstanding training device will be available to on-site commanders, for obtaining information in the operational, training, and research areas:

(1) Operational: (a) The maximum man-machine potential of the M33 radar under existing field conditions (such data appears essential for planning purposes); (b) the effects of a faster, more realistic raid pattern on a crew trained largely in acquiring and tracking relatively slow propellor-driven aircraft; (c) saturation points existing with respect to changing targets, and so forth.

(2) Training: (a) Training practices and simulated conditions that will help to expose and overcome primary obstacles to crew training; (b) how quickly a crew can be expected to acquire proficiency, and ways a commander can find out how well his crews compare with such expectations.

(3) Research: (a) The personality characteristics, skills, and capabilities that provide a crew with the greatest potential for performance; (b) the lower limits that may be set on intelligence, and other factors before performance is seriously depreciated; (c) the organizational procedures and leadership practices that are most appropriate for maximum crew performance; (d) the successful role of the battery commander in a unit whose equipment is highly technical may differ from ordinary leadership concepts.

A large amount of information associated with the study is not reported here because it did not meet the requirements of statistical significance. The study was designed largely as a source of hypotheses about the performance of batteries.

RADAR MAINTENANCE

Description of the Measure

Unit rankings on efficiency of Radar Maintenance were determined by inspections of Ordnance files over a three-month period, and through subsequent judgments by Ordnance personnel about the cause of the repairs. For each job order describing a repair a judgment of either "fair wear and tear" or "inadequate maintenance" was made. Incidents judged to be the results of inadequate maintenance were then categorized into three groups and weighted according to their importance; the sums of these weighted scores represented the battery score on maintenance.

Illustrations of those incidents that contributed to the scores are: (a) damaged threads on a leveling jack, brought about by forcing the jack with a heavy and inappropriate object, (b) damage to van or antennae resulting from faulty or incomplete march order preparation, and (c) burned-out servo motors, resulting from too great an initial surge of current from the generator.

The nature of these incidents suggests that the computed score was a measure of the maintenance activities of the entire crew rather than those of the Radar Mechanic alone. Other indications support this notion.

For example, a previous HumRRO report[1] has shown that 47% of the malfunctions of M33 radars in the field were originally detected by operators. Furthermore, the nature of the mechanic's job in the on-site unit limits his activities; only a few mechanics perform maintenance more

[1] *The AAFCS M-33 Mechanic: Analysis of Field Activities and Problems With Implications for Training*, HumRRO Information Report by Staff, RADAR, March 1954.

complex than regulations prescribe—replace chassis, switch tubes, and so forth.

The nature of the incidents referred to Ordnance, the restricted aspects of the mechanic's job, and the role of operators in identifying malfunctions, all lend support to the notion that the measure used in this study was one indicating the efficiency of the entire crew rather than the performance of an individual.

Although the radar maintenance measure is more an index of crew maintenance effectiveness than a measure of mechanic proficiency, the mechanic might be expected to exert an important influence on group performance by virtue of his role in the group as technical expert.

Characteristics of the Radar Mechanic and Maintenance Performance

The Radar Mechanic, a Fort Bliss trained technician, qualified for training with a minimum Aptitude Area IX score of no less than 100. The mechanics who were sampled in this study ranged between 102 and 136. When the AA-IX score of the mechanic was correlated with his unit's maintenance score, no relationship was demonstrated. These results indicate that while this particular measure may be appropriate as a selector for school success, it may not be related to how well radar maintenance is performed in the field within the restricted score range of those completing school.

When mechanics' scores on Aptitude Areas I through IV were related to the maintenance measure, a significant relationship was demonstrated in the negative direction; that is, the higher the mechanic's score on these measures of general intelligence, the poorer his unit ranked on the Radar Maintenance measure. Few individuals in this sample had scores below 100 on any of the Aptitude Areas used; therefore, the negative relationship may be said to have been shown only for that intelligence range at the high end of the scale represented by the sample. Quite possibly, a cutoff point exists, below which mechanics' lower intelligence scores begin to relate to poorer maintenance--that is, the relationship is perhaps curvilinear. But the findings stand within the limits of the intelligence area sampled. A negative relationship between intelligence and performance in a technical field is unusual; its presence might well be studied to determine what further factors are operating. Concrete actions implication would follow if, for example, part of the explanation for these findings was that the less intelligent mechanics—within the limits of the present sample—were more inclined to be satisfied with the restricted nature of their duties, and consequently maintain a better level of motivation in their work.

Other attributes of radar mechanics are related to maintenance performance. When the battery had good Radar Maintenance, the radar mechanic had been in the battery longer and was *less* dominant in interpersonal relations.

With respect to the Radar Mechanic and performance on Radar Maintenance the data are for the most part unexpected. Although mechanics in good crews are more experienced in maintenance, they are also less intelligent and less dominant. Not all of these results were expected; their unusual nature suggests that more should be known about the Radar Mechanic operating under field conditions.

Crew Characteristics Related to Maintenance Performance

When the non-status members of radar crews were analyzed for those characteristics that related to maintenance performance, certain relationships were found. When the maintenance was good, the group—with the exception of the Mechanic, Platoon Sergeant, and Chief Operator—was shown to be younger. Furthermore, the shorter time they had spent in the Army, and in the battery, the better the maintenance. Also, the level of general morale was positively related to maintenance.

Maintenance can be a boring business. The routine polishing, cleaning, and oiling activities prescribed by the on-site operating procedures tend to become irksome to the soldier after a number of repetitions, and his motivation to perform well probably drops. The average time in the Army for the crews who had the best maintenance record was 10.4 months, and the average time in the battery, 4.4 months. These figures, when compared with the experiential factors associated with the poorest crews—whose average time in the Army was 22.4 months and time in the battery 9.0 months—would indicate that boredom may be related to maintenance performance. Subordinate members in good crews may not have reached the saturation point—the point at which their activities become so routine that a decrement in performance results. An alternate interpretation is that those crews whose members were in service longer were "short timers," had little service remaining, and were less motivated to perform.

A most interesting result of this study involves the unexpectedness of the findings. Those relating to the mechanic, added to measures of crew experience, suggest that characteristics influencing maintenance may not be those usually expected, and that much more research should be done on maintenance per se if these somewhat paradoxical results are to be understood.

Utilization of the Findings

The maintenance problem is, of course, not limited to AAA units—it applies to every Army organization with responsibility for the upkeep of materiel. The information gained from this study, regarding those variables that proved to be significantly related to the maintenance criterion, suggests that motivation may be an essential factor in good maintenance.

While the present results have some immediate use in selection and training, the most important implications, then, rest in this hypothesis.

Research on motivational factors should at least be considered equally with the checklist development and job analyses.

DEFENSE COMMANDER'S RATING

A third criterion of unit performance was the composite rating of each battery made by the Defense Commander. The Defense Commander's Rating (DCR) is based on observations made by the Commander and his staff over a period of three months. The inspection system includes ratings of battery administration (S-1), intelligence (S-2), artillery knowledge and training (S-3), food service and supply (S-4), and communications, generators, motors, and radar activities.

It should be emphasized that the DCRs of the various units were based on an inspection system developed by a particular Commander and his staff. The findings can be extended or generalized to units in other defenses only in so far as we can assume that the inspection systems of other Commanders are like that of the defense studied.

Battery Commander Characteristics That Varied With DCR

The image created by a military organization is closely associated with its Commander. His status marks him as the primary source of communications between higher echelons and battery personnel. When superiors desire changes in the battery's activities, they must be brought about through him. Because he is so closely identified with the unit, many come to perceive Commander and unit as one. This association has contributed to the belief that a unit assumes the characteristics of a Commander, and that, conversely, the Commander's personal characteristics will be expressed in unit performance.

This study permits an examination of the concept that a unit and its Commander may be subsumed into a single image, and that the Commander's traits are reflected in unit performance. If this suggestion is valid, it would be expected that variables showing differences between Commanders would be related to the unit ratings given by the Defense Commander. If, under these circumstances, relatively few of the Commander's characteristics are related to unit performance, or if, in comparison to other individuals and groups within the unit, few of his characteristics are related to performance, then the hypothesis would appear to have little practical value.

Unit Commanders Characteristics That Varied With DCR

Information describing the Commander's mode of leadership was available from the Leader's Behavior Test. This instrument permitted key personnel to describe the manner of the Commander's actions with respect to the extent of his consideration for his men and how he initiated structure when a particular assignment was passed down.

Commander's scores on these two scales did not vary consistently with the Defense Commander's unit ratings.

Another leadership evaluation test depicted a variety of hypothetical (yet typical) decision situations, and elicited from battery members an index of the extent of their agreement with the Commander's decisions. Agreement scores of the various individuals and groups in the battery did not consistently relate to the unit's Defense Commander's Rating. A related leadership evaluation test—designed to measure the predictability of the Commander's decisions—failed to relate to unit rating on DCR.

In addition, 11 scales were administered to the 40 unit Commanders to determine their relative inclinations on certain personality dimensions. One tendency stood out as a characteristic that varied consistently with unit DCR: the greater the Commander's tendency to worry and be frustrated by problems associated with his duties, the higher his unit's DCR. Such a Commander showed more concern for elements of command that were especially sensitive to his superior.

That the Commander's personality is relatively unimportant to unit DCR, except in the above sense, is emphasized by a listing of some of those dimensions often mentioned as qualities of the good unit leaders, but which, in fact, showed no relationship to DCR performance. Such personality characteristics as the extent to which the Commander was more dominant in interpersonal relationships, his tendencies toward authoritarian attitudes, his degree of will control, his independence of action, and his evaluation of perseverance and efficiency as goals—all commonly prescribed prerequisites for the good Commander—did not vary consistently with DCR.

Certain relationships existed between unit DCR and subordinates' perceptions of their Commander. These were global perceptions rather than perceptions about specific details of organizational ability or leader behavior. When the Radar Crew gave their Commander a high rating as an officer, the unit's Defense Commander Rating was high. But when the Radar Crew approved of the specific details of their Commander's method of operation, units had no better DCR performance than when the Radar Crew disagreed with their Commander's specific actions.

The absence of relationship between unit DCR and the Commander's personal characteristics, both actual and perceived, would imply that the influences exerted by the Commander are not so great as is often supposed, or that they are too subtle to be measured by the instruments used in this study. Compared to other individuals and groups in the battery, the Battery Commander's characteristics appear to have little relation to criterion scores.

First Sergeant Characteristics That Varied With DCR

While the Commanders of high DCR units reflected tendencies to worry and be concerned with the dispositions of their superiors, the

First Sergeants responded in the *opposite* manner. That is to say, the unit high on DCR tended to have an easily worried Commander and a First Sergeant who took obstacles as they came, unexcited by routine obstructions or inconsistency of command criticisms. His role in the good battery was one of maintaining calm and selecting from the Commander's worries the essentials upon which to act.

The First Sergeant in the unit with a high DCR was more dominant in interpersonal situations, was more mature, and possessed more will control with respect to his independence of action than did his counterpart in the low DCR unit.

All things considered, the First Sergeant in the good unit fit his stereotype better than did the Commander. His characteristics more often varied consistently with unit rating on DCR than did the characteristics of the Commander.

Battery Characteristics That Varied With DCR

As was true in the case of the First Sergeant, the Senior NCOs in good batteries were no more intelligent than their equals in poorly rated units, nor were they better educated. As a group, they had been in the battery longer. This longevity factor, like RRP, was the one that stood out most importantly in describing the Senior NCOs of a good unit. Because length of time in the battery related to both RRP and DCR, it was a strong argument for the stabilization of such key personnel.

Characteristics of Junior NCOs related to Defense Commander's Rating in many instances where characteristics of the Senior group did not. This fact was also true when RRP was considered. In fact, one of the strongest relationships of the entire study was between the intelligence scores of Junior NCOs and their unit's Defense Commander's Rating—the higher their average intelligence, the better their unit's DCR. This relationship was not found for the Senior group; that is, the average intelligence of Senior NCOs was not related to DCR performance.

Subordinates' perceptions of their unit's NCOs were extremely important. When the battery as a whole rated their NCOs high, DCR was high. This relationship was also very pronounced for the Radar Crew in that the higher it rated the unit's NCOs the higher the unit's DCR.

Conversely, Commanders of good units rated their NCOs no better than did Commanders of poor units. This fact has significant bearing on the present use of efficiency ratings for these men. If the NCOs of good units, perceived as better by their subordinates (and actually better in terms of unit DCR), receive no better efficiency ratings from Commanders than do their counterparts in poor units, then it might be advisable to question the value of these ratings as a basis for subsequent promotion and placement of NCOs.

As a group, Junior NCOs of high DCR units had been in the Army considerably longer, were more apt to be Regular Army personnel and were more often planning to reenlist, than NCOs of low DCR units.

When the entire enlisted population of units was considered, it was found that the average enlisted man in the good unit had been in the Army longer and expressed higher general morale. He tended to hold his battery in higher esteem than did the average EM in the unit with low DCR, thus demonstrating that he could realistically evaluate his unit, even though he had, perhaps, little information about the rest of the defense.

Radar Crew Characteristics That Vary With DCR

The Radar Crew proved to be the group with characteristics most sensitive to the DCR. What these men *were* on the average, and how they felt about conditions in the battery varied consistently with DCR more often than was true of any other individual or group. When they tended to value perseverance and efficiency as goals in themselves, when they rated their officers and NCOs high, or when they ranked their battery high as compared to other units in the defense, their unit's DCR was high.

As with RRP, neither the average intelligence, nor length of Army service, nor time in the battery of the Radar Crew was related to DCR. More battery experience for the key radar NCOs, however, was associated with a higher DCR rating, as was longer on-the-job tenure of the chief operator.

Utilization of the Findings

Some of the factors identified above are controllable and some are not. Some have implications for personnel policy at the higher level, and some for Commanders in the field. A summary follows of those characteristics related to Defense Commander's Rating that are most relevant to the Commander and to lower level headquarters:

(1) The primary attribute of the Commander as shown in this study is that he be the type of individual who is concerned about details. He must be eager to please authority and sensitive to its demands.

(2) The First Sergeant should, as much as possible, fit the Army "stereotype" for his job. He should be stable and possess a quality of independence.

(3) Senior NCOs need not be exceptionally intelligent or well-educated, but the degree to which they are stabilized in their battery appears to be an important factor.

(4) Junior NCOs should have longer Army experience, plus comparatively high intelligence and education.

(5) Crucial characteristics of the Radar Crew correlated with a high DCR are not that its members be stabilized in the battery, intelligent, and well-educated, but that they share certain personality characteristics.

With the exception of Senior NCOs, time in battery, the experience factors so frequently related to DCR were not related to Range of Radar Pickup. These relationships between DCR and longevity may well be explained by the fact that experienced soldiers know better how to satisfy an inspecting officer. That such experience is not related to the objective operational measure, RRP, suggests that although Commanders emphasize such factors, they should not mistakenly believe that they are also related to operational performance. While variables associated with experience may contribute to the "good soldier" qualities expected by a Commander, they are not after a certain point, related to unit performance on Range of Radar Pickup.

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Human Resources Research Organization (HumRRO) 300 North Washington Street Alexandria, Virginia 22314	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

COLLECTED PAPERS PREPARED UNDER WORK UNIT AAA: FACTORS AFFECTING PROFICIENCY AND MORALE IN ANTIAIRCRAFT ARTILLERY BATTERIES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Professional Papers

5. AUTHOR(S) (First name, middle initial, last name)

Staff members of Work Unit AAA

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
November 1969	41	3

8a. CONTRACT OR GRANT NO. DAHC 19-70-C-0012 (current) b. PROJECT NO. 2Q062107A712 (current) c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Professional Paper 33-69 9b. OTHER REPORT NO.(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES Presentations at professional meetings, journal articles, and a task paper.	12. SPONSORING MILITARY ACTIVITY Office, Chief of Research and Development Department of the Army Washington, D.C. 20310

13. ABSTRACT

Research conducted in 1954-55 to determine the contribution of certain human factors to the effective performance of crew members of antiaircraft artillery batteries is described in these five papers: "Battery Effectiveness: Assessment of Comparative Performance," by Francis H. Palmer and Thomas I. Myers; "Crew Description Dimensions and Radar Crew Effectiveness," by Thomas I. Myers and Francis H. Palmer; "Sociometric Choices and Group Productivity Among Radar Crews," by Francis H. Palmer and Thomas I. Myers; "Leadership and Group Achievement," by Francis H. Palmer; "Human Factors Affecting the Performance of Antiaircraft Batteries," by Francis H. Palmer *et al.*

DD FORM 1 NOV 65 1473

Unclassified
Security Classification

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Group Behavior						
Human Relations						
Individual Behavior						
Individual Differences and Training						
Training Concepts and Theory						

Unclassified

Security Classification