

AD-A079 174

ARMY PERSONNEL RESEARCH OFFICE WASHINGTON DC
AUTOMATED PROCEDURES FOR OBTAINING AND SCORING PEER RATINGS, (U)
AUG 64 F F MEDLAND , J OLANS
APRO-RESEARCH STUDY-64-2

F/G 5/9

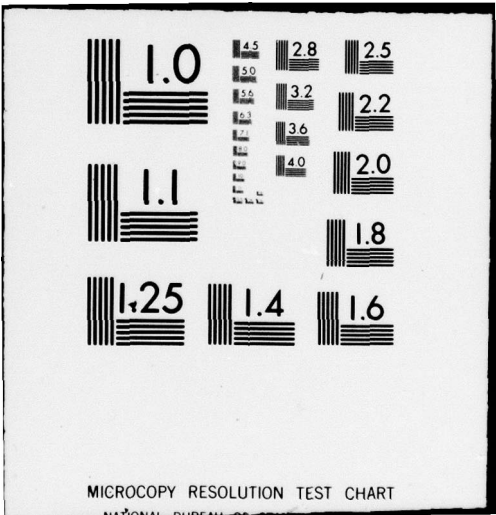
UNCLASSIFIED

NL

| OF |
AD A
079174



END
DATE
FILMED
2 - 80
DDC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

~~LAST COPY FOR REFERENCE USE ONLY~~
~~Return to Publications Br., USAPRO~~

LEVEL

HP (1)

ADA 079174

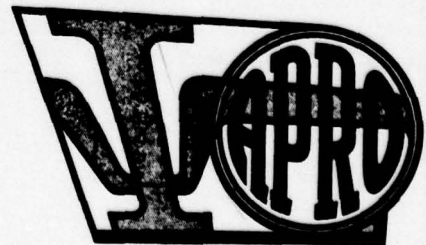
Research Study 64-2

**AUTOMATED PROCEDURES FOR
OBTAINING AND SCORING PEER RATINGS**

D D C
RECEIVED
DEC 26 1979
RECEIVED

A

DDC FILE COPY



U. S. ARMY PERSONNEL RESEARCH OFFICE

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

79 12 19 149

14 APRO-Research Study-64-2

6 AUTOMATED PROCEDURES FOR
OBTAINING AND SCORING PEER RATINGS,

10 Francis F. Medland and Jerome Olans

Submitted by

Leonard V. Gordon
Chief, Behavioral Evaluation Research Laboratory

Approved by

J. E. Uhlauer
J. E. Uhlauer
Director, Research Laboratories

Charles S. Gersoni
Charles S. Gersoni
Colonel, GS
Commanding

12 15

Research Studies are special reports to military management. They are usually prepared to meet requests for research results bearing on specific management problems. A limited distribution is made--primarily to the operating agencies directly involved.

11 August 1964

040 650

JRC

PREFACE

Enlisted men who are capable of becoming effective combat NCOs should be identified early in their Army service so that they can be given appropriate training. Research on the selection of NCO leaders aims at developing a two-stage screening process: the measurement of leadership potential during basic training and the measurement of leadership skills acquired through an extended period of military service.

One objective of the NCO LEADERS Task is to evaluate peer ratings as predictors of NCO potential and to assess the relative effectiveness of various methods of obtaining peer ratings. The present publication reports on a portion of subtask a, "Development of a preliminary potential NCO leadership screening technique," FY 1964 Work Program.

The entire research task is responsive to special requirements of the Deputy Chief of Staff for Personnel and of the U. S. Continental Army Command, as well as to requirements to contribute to achievement of the objectives of Department of the Army R and D Project No. 2J024701A722, Selection and Behavioral Evaluation.

Accession for	
NTIS GRA&I	<input checked="" type="checkbox"/>
DDC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist.	Avail and/or special
A	

040 650

70B

BRIEF

Requirement:

To explore the feasibility of applying computer techniques in obtaining and scoring peer ratings in Basic Combat Training companies.

Procedure:

An experimental prototype of the Army Personnel Rating Machine and a rating procedure utilizing the Army Rating Punch Card were assessed for operational feasibility in field tryout on basic training companies at Fort Ord, Calif., in late 1963.

Findings:

The present study indicated the feasibility of adapting computer techniques to the rating process and pointed up the differing operational characteristics of the two methods.

With either procedure, both speed and accuracy of scoring ratings and computing average ratings would be increased.

If adequate equipment to accommodate a full company of raters were made available, the computerized rating machine could reduce time spent in rating sessions from about one hour (current procedure) to 20 minutes. Results are immediately available. With the current prototype equipment, additional processing of results would be necessary for a computerized assignment procedure, a drawback which could be met by addition of print-out equipment. Initial cost of equipment and maintenance costs would be high.

With the punched card method, more time would be required for rating sessions than under present procedures. Results would be directly utilizable in a computerized assignment procedure. Cost would be low, since only regular tabulating equipment is required.

Utilization of Findings:

Development of satisfactory automated rating procedures would make practicable the large-scale use of peer ratings in classification and assignment as well as for special purpose selection.

AUTOMATED PROCEDURES FOR OBTAINING AND SCORING PEER RATINGS

In research investigations in which ratings have been studied, peer ratings have consistently been among the best predictors of subsequent Army performance. On the basis of results of a number of studies, peer ratings were introduced operationally in the recruit training programs of the Army late in 1961. The ratings are used as part of the selection procedures for a USCONARC-initiated program of NCO leadership training. The routine of the program is now established so that peer ratings are obtained on recruits in every basic combat training company formed in Army training centers.

More widespread use of ratings in personnel assignment is anticipated for the future. For example, ratings may be used with increasing frequency in the selection of personnel for unique assignments such as Special Warfare. Further, study is being made of the applicability of peer ratings as a part of initial classification procedures, possibly as a component of aptitude area composites leading to personnel assignments to infantry, armor, artillery, or other combat MOS.

A practical limitation to extensive use of peer ratings in the past has been the administrative complexity of obtaining ratings on large numbers of personnel within a restricted time sequence. Because of the developing importance of ratings in Army selection programs, the U. S. Army Personnel Research Office undertook an investigation of means by which certain administrative procedures in the rating program could be simplified and automated. Ultimate goals of the study are:

1. Improve the timeliness with which scores are obtained, so that results could be available for use in assignment to initial MOS training.
2. Reduce clerical cost and labor required in conducting peer rating sessions and processing the results so as to improve the operational utility of the program.

The sections which follow outline current operational requirements and procedures and expanded requirements which would be generated by more extensive use of ratings. Two experimental procedures for automating phases of the rating program, the Army Personnel Rating Machine and the Army Rating Punch Card, are described. Preliminary field evaluations resulted in delineating the major features to be considered in the case of each method.

CURRENT OPERATIONAL PROGRAM

Peer rating sessions are conducted late in the fourth or fifth week of the Basic Combat Training cycle. The rating session, requiring

between one and two hours, is scheduled for an entire training company at one time. The peer rating group is the squad, ranging in size from 10 to 18 trainees. The normal training company has 16 squads. However, an occasional company may include as many as 24 squads. Typically, the individual training center conducts rating sessions on 8 to 10 BCT companies each week, additional sessions being required during periods of peak input.

At a rating session, each squad member is given a squad roster and a rating form. He first arranges all other members of his squad in order of merit based on estimated leadership potential. He then assigns a numerical scale value--integral values from 1 (poorest) to 7 (best)--to each of the individuals he has rank ordered. He enters these values on the rating form. The individual's leadership potential score is the mean of the ratings assigned to him by all other members of his squad.

Control of the operational program at the training center is generally maintained by the adjutant general. Four to six enlisted men are responsible for the preparation, conduct, and processing of results of the rating program, including the following administrative steps:

Preparation. Arrangements for physical facilities; scheduling the rating session; coordination with the training company in establishing the personnel complement of each squad; mimeographing rosters of the squad (the peer group).

Rating session. Conduct of the rating session by test administrator and four to five proctors, in accordance with standard test administration procedure and specific instructions for obtaining peer ratings.

Processing results. Manual transcription of individual ratings into a score matrix; computation of average scores.

Currently, eligibility for the NCO leadership training is restricted to trainees who have been assigned to combat MOS and who have a standard score of 100 or higher on the appropriate combat aptitude area. Rating scores are computed only for eligible trainees (perhaps 15 to 20 percent of the total number). Further, since Advanced Individual Training assignees are not selected for acting leader positions until late in the BCT cycle, results of the ratings are not needed immediately. In the case of eligible trainees transferred to another training center for Advanced Individual Training, scores are completed and forwarded to the new center with the individual's transfer orders. There is no requirement for entry of the leadership potential rating in the individual's permanent records, since scores serve only local and one-time assignment purposes. Computation is therefore completed at any convenient time consistent with local requirements.

REQUIREMENTS OF AN EXPANDED PROGRAM

If a rating program should be made part of the differential classification system, certain requirements will have to be added to those imposed by the current operational system:

Timeliness. Rating results for trainees will be required by the Department of the Army (Office of Personnel Operations, Washington, D. C.) in time for collation with other classification data used in making initial assignments to school or on-job training. Current estimate is that rating scores would have to be available at least two weeks prior to completion of the Basic Combat Training cycle. This requirement will entail very close timing, since judgments by peers must have opportunity to mature before ratings can be taken, four to five weeks as a minimum. Ratings may either be scored in the field and results transmitted to the Department of the Army, or raw data may be transmitted to the Department of the Army for centralized scoring and processing. If the latter course is followed, provision must be made for transmitting results to the Basic Training Center prior to completion of the BCT cycle.

Number processed. Rating scores will be required for a more extensive segment of the basic trainee population rather than for a limited number as at present.

Recording. Provision will have to be made for entry of the rating score on the trainee's Enlisted Qualification Record (DA Form 20).

EXPERIMENTAL AUTOMATED PROCEDURES

THE ARMY PERSONNEL RATING MACHINE

The Army Personnel Rating Machine is a digital computer with a capability for recording and computing both rank order and graphic ratings by multiple raters. The machine was designed by engineers of the National Bureau of Standards following specifications outlined by APRO staff members. Each rater is given a rating panel equipped with plugs (up to 20) numbered to correspond to the individual members of the rating group, and sockets into which the plugs are inserted. Rank or rating is established by placement of the ratee plugs in appropriate sockets. One arrangement of sockets provides for the serial or rank-ordering of the ratees. A second arrangement of sockets provides for a columnar arrangement (seven columns corresponding to seven points of a graphic rating scale) for the rating of the ratees (Figure 1). The rater considers each of the other individuals of his group in turn, makes his rating, and completes his task by inserting the plugs representing the ratees in the appropriate sockets.



Figure 1. Rating panel in use. The rater is placing ratee plugs to rate his peers on a 7-point scale.

Each of the individual rating panels is wired to a central computer, which contains means of connecting the rating panels, the computing equipment, the information displays, controls, and power supplies. In order to maintain second decimal place accuracy in scoring, digital logic circuitry is used in the computing equipment. Figure 2 shows the central console.

Results on the group can be determined immediately. When the entire group has completed the rating procedure, the console panel shows, for each member of the group, his man number, the number of individuals who rated him, and the average of the ratings he received. Although not included in the prototype rating machine, provision can be made for a print-out of the results and/or punched tape record for use in subsequent statistical processing.

FIELD TEST OF THE RATING MACHINE

Tests to assess the feasibility of the rating machine were conducted in the operational rating program at Fort Ord, California, in the second half of 1963. The prototype machine developed for this field test provided for no more than 20 raters at one time. For operational use, a machine would have to have facilities for handling a full company at a time.

The field tryout involved 70 squads from 29 different training companies. The squads were split into random halves; one half rated the entire squad on the rating machine, the other half rated the entire squad by the current paper-pencil operational procedure. At the end of each session with the rating machine, personnel who conducted the session provided what was essentially a critique of the session by completing an evaluation form consisting mainly of open-end questions. These evaluations were concerned with the general efficiency of the rating machine as compared with current procedures--timing and control of the rating procedures, assistance needed by raters, etc.--rather than with effectiveness of the ratings obtained. Results are summarized below:

Administration. The actual rating session required 15 to 20 minutes per squad, a reduction from about one hour for the current procedure.

Certain difficulties were encountered in preparation for the ratings and between-session arrangements. Ratee plugs in the rating panels had to be maintained in position until after the scores had been recorded. In order to return the rating panels to neutral position for the next session, trainees had to be kept in their seats until scoring was completed so that they could place the ratee plugs in storage position. The alternative was for administrator and proctors to clear all panels between sessions.

THE RATING PANEL

- 20 NUMBERED RATEE PLUGS

Plugs are labeled with names of the men to be rated.

- AN ARRAY OF SOCKETS FOR RANKING

The rater places the labeled plugs in the top row of sockets, ordering his men from best to poorest.

- AN ARRAY OF SOCKETS FOR GRAPHIC RATINGS

The rater places each labeled plug in one of seven columns which represent the seven steps of a scale.

- A LOCKING DEVICE

The scores are registered when the cover of the rating panel is closed.

THE RATING PANEL

● **20 NUMBERED RATEE PLUGS**

Plugs are labeled with names of the men to be rated.

● **AN ARRAY OF SOCKETS FOR RANKING**

The rater places the labeled plugs in the top row of sockets, ordering his men from best to poorest.

● **AN ARRAY OF SOCKETS FOR GRAPHIC RATINGS**

The rater places each labeled plug in one of seven columns which represent the seven steps of a scale.

● **A LOCKING DEVICE**

The scores are registered when the cover of the rating panel is closed.

Each of the individual rating panels is wired to a central computer, which contains means of connecting the rating panels, the computing equipment, the information displays, controls, and power supplies. In order to maintain second decimal place accuracy in scoring, digital logic circuitry is used in the computing equipment. Figure 2 shows the central console.

Results on the group can be determined immediately. When the entire group has completed the rating procedure, the console panel shows, for each member of the group, his man number, the number of individuals who rated him, and the average of the ratings he received. Although not included in the prototype rating machine, provision can be made for a print-out of the results and/or punched tape record for use in subsequent statistical processing.

FIELD TEST OF THE RATING MACHINE

Tests to assess the feasibility of the rating machine were conducted in the operational rating program at Fort Ord, California, in the second half of 1963. The prototype machine developed for this field test provided for no more than 20 raters at one time. For operational use, a machine would have to have facilities for handling a full company at a time.

The field tryout involved 70 squads from 29 different training companies. The squads were split into random halves; one half rated the entire squad on the rating machine, the other half rated the entire squad by the current paper-pencil operational procedure. At the end of each session with the rating machine, personnel who conducted the session provided what was essentially a critique of the session by completing an evaluation form consisting mainly of open-end questions. These evaluations were concerned with the general efficiency of the rating machine as compared with current procedures--timing and control of the rating procedures, assistance needed by raters, etc.--rather than with effectiveness of the ratings obtained. Results are summarized below:

Administration. The actual rating session required 15 to 20 minutes per squad, a reduction from about one hour for the current procedure.

Certain difficulties were encountered in preparation for the ratings and between-session arrangements. Ratee plugs in the rating panels had to be maintained in position until after the scores had been recorded. In order to return the rating panels to neutral position for the next session, trainees had to be kept in their seats until scoring was completed so that they could place the ratee plugs in storage position. The alternative was for administrator and proctors to clear all panels between sessions.

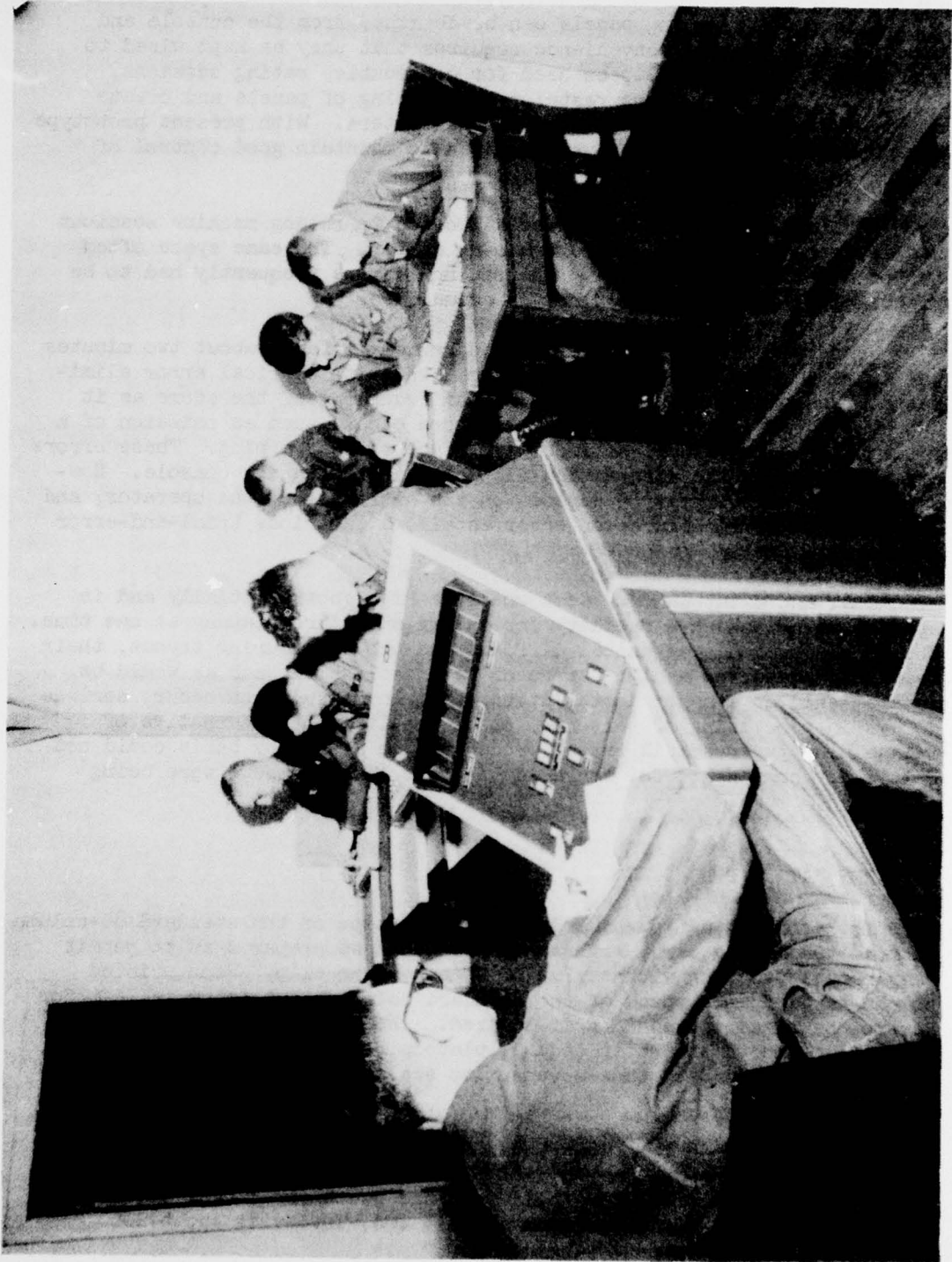


Figure 2. The rating console sums all scores for each ratee, counts the raters per ratee, averages the ratings, reports the final score, and displays ratee number, number of raters, and score.

Although the rating panels can be detached from the console and used in any location, convenience requires that they be kept wired to the console if they are to be used for consecutive rating sessions. Present wiring limitations restrict the spacing of panels and create difficulty in maintaining privacy between raters. With present prototype equipment, close proctoring was required to maintain good control of rating procedures.

Physical space which could be allocated to rating machine sessions was severely restricted at the training center. The same space often serves a variety of purposes. The rating machine frequently had to be disconnected and stored when not in actual use.

Scoring. Scores can be processed extremely fast--about two minutes per squad is required. Scoring is accurate, with clerical error eliminated. The only manual operation is the recording of the score as it appears on the console. Some rater errors occur, such as omission of a rating for an individual or incomplete insertion of a plug. These errors are reflected in the information displayed in the rating console. However, their identification depends on the alertness of the operator, and localization of the source of error entails a laborious trial-and-error search.

Cost and maintenance. Cost would be high, both initially and in upkeep, to establish a facility for rating an entire company at one time. While only nominal mechanical difficulties occurred during tryout, their resolution required a high degree of technical skill such as would be assured only at a major installation. Should malfunction occur, serious delays in processing of personnel could result. The alternative of maintaining paper-pencil rating procedures on a standby basis could not completely compensate for the delay in case large numbers were being processed.

THE ARMY RATING PUNCH CARD

The Army Rating Punch Card is an adaptation of the standard 80-column machine tabulating card, with punch positions so prepared as to permit easy manual punching with stylus or opened paper clip. Each printed column takes up the space of two standard columns, but space is left for key-punching additional data if desired. The card contains columns for identification of the rater (squad, platoon, man number) and for rating up to 20 squad members on a seven-point scale.

FIELD TEST OF THE ARMY RATING PUNCH CARD

Tests to assess the feasibility of the punch cards were conducted in the operational rating program at Fort Ord, involving about 2000 trainees in 10 BCT companies.

The punch card used in the field tryout (Figure 3) was one designed for a study of peer ratings of ROTC students in Summer Camp. Procedures were adapted for trial in the Army Training Center. The card could easily be redesigned so as to apply specifically to Basic Training companies. Each trainee was provided with a rating card in which identifying information--installation and company--had been printed. He was also given a roster of the rating unit (the squad) and a rating form. The rater punched additional identifying information on the rating card--platoon and squad number and his own man number as it appeared on the roster. He accomplished ratings of his fellow squad members on the rating forms as in current operational rating sessions. He then transferred his ratings to the punch card, identifying each fellow squad member by man number and indicating the rating by punching a number from 1 to 7 in the appropriate man number column of the rating card.

Evaluations of the sessions were obtained from administrators and proctors by procedures similar to those used to evaluate the rating machine.

Administration. With the punch card system, ratings for an entire company can be obtained at one time. Time required for the rating session increased by 15 to 20 minutes beyond the time required for current operational sessions. Preparation of squad rosters and scheduling was much the same, the preparation of the cards being the only additional requirement. Extremely alert proctoring was required to prevent rater error. There is no easy way to check completeness of rating at the time the rater turns in his card. However, if error was detected, or if a rater found he had mispunched, the error usually had to be corrected by completion of a new card.

Scoring. Clerical error was eliminated in the scoring and recording of results. However, rater errors which had gone undetected at the rating session were found. By checking with the rating form, errors of a chain-reaction type were discovered: if a punch was entered in a wrong column, succeeding columns would also be mispunched. Effect of inconsistent ratings detected during computer processing could be eliminated at the cost of losing one rater for one rating score. It may be possible to include more effective self-checking procedures in future manuals.

The punched cards are processed by tabulating equipment, and results can be available in a short time. Results can be used directly in subsequent computer processing, should such processing be desired.

Cost and maintenance. Cost is nominal. There is no requirement for equipment beyond tabulating equipment now available, and no additional expense except for cost of the cards. No maintenance skill is required beyond that currently available for maintenance to tabulating machines.

RATINGS

1		2		3		4			5		RATINGS																					
FILE NO. A B		YOUR SQUAD		YOUR MAN NO. A B		YOUR COLLEGE CODE A B C			SQUAD YOU ARE RATING		RATEE MAN NO. →		01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20
0	0	0	0	0	0	0	0	0	0	0	0	0	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	
6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	
7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	
8	8	8	8	8	8	8	8	8	8	8	8	8	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
9	9	9	9	9	9	9	9	9	9	9	9	9	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	
1	1	2	2	3	3	4	4	5	5	6	6	7	7	8	8	9	9	10	10	11	11	12	12	13	13	14	14	15	15	16	16	

SCALE VALUE OF RATINGS

PT4573, ARMY RATING PUNCH CARD

CODE BELOW SHOULD BE USED BY RATER

Figure 3. Facsimile of Army Rating Punch Card used in field tryout.

GENERAL OBSERVATIONS

The field tests were aimed at assessing the feasibility of procedures rather than at investigating the psychometric equivalence of results. However, spread of ratings obtained with either type of experimental equipment was approximately the same as that attained with current operational procedures: 25 percent of the ratings were in category 4 (average), another 25 percent in the combined categories 5, 6, and 7 (best), and 50 percent in combined categories 1, 2, and 3 (poorest).

While not conclusive as to the overall superiority of one method over the other, the present study indicated the feasibility of adapting computer techniques to the rating process. Both the rating machine and the punch card system provide score data without clerical tabulation or calculation, except for the manual recording of the scores displayed on the console of the rating machine. This requirement would be eliminated if print-out equipment were added. With the rating machine, scores are instantly available. Using the current prototype machine, further processing would be required before the scores could be used in computerized assignment procedures. Again, use of print-out equipment to reproduce the ratings on tape or punch card would meet this objection. The punch card procedure adapts readily to current operational procedure in which ratings obtained on roster forms are transferred to punch cards. There is some delay in availability of results, since the cards are shipped to a central location where they are processed by computer. Results are subsequently mailed to the Army Training Center where they are required as guides in making leadership assignments during training. In the event ratings should enter more widely into classification and assignment, changes in the current practice of central processing would in all probability follow. For example, transceiving of information (both to and from the computer center) rather than mailing is under discussion as a means of making scores immediately available to the training center.

Additional studies of these methods, as well as future investigations stemming from these studies, will be conducted and reported by the Statistical Research and Analysis Laboratory of the U. S. Army Personnel Research Office.