PRL-TR-67-13

December 1967

PROCEEDINGS
TWENTY-FIFTH ANNIVERSARY SYMPOSIUM

PERSONNEL RESEARCH AND SYSTEMS ADVANCEMENT

Sponsored by

THE PERSONNEL RESEARCH LABORATORY
AEROSPACE MEDICAL DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
Lackland Air Force Base, Texas

Edited by

Angelo L. Fortuna, Major, USAF

This symposium was held in San Antonio, Texas, November 1 - 3, 1966, under USAF Contract AF 41(609)-3173.

Published December 1967

Notices

FOREWORD

Rapid advances in technology have been paralleled by the growth in complexity of personnel subsystems. While technological advancement finds ready expression in news media, the techniques developed for efficient manipulation of personnel systems suffer the quiet dignity of esoteric journals, technicals reports, operating manuals, and the like. Although publicity for the sake of publicity should not be a goal of the personnel researcher, effective systematic methods for communication of the discipline certainly warrant promotion. It was with this intent that the Symposium on Personnel Research and Systems Advancement was sponsored by the Personnel Research Laboratory of the United States Air Force. The Symposium climaxed the twenty-fifth anniversary year of that Laboratory following an uninterrupted but undoubtedly cyclothymic era – an era started in 1941 with the Army Air Force and transitioned to the United States Air Force with its "birth" in 1947.

Historically, personnel research has assumed a sophisticated pose on the threshold of its own computer age. Because the great number of people which constitute the military personnel systems, military planners have been particularly responsive to automated techniques starting with the punch and "shuffle" to the point where computerized models are being implemented as a matter of course. The great breakthroughs in personnel systems technology are on the verge of appearance. Then delay is a function of the need for interactive relationships among many research disciplines. Each study requires the attention of psychologists, economists, mathematicians, operations analysts, and a host of support personnel. Principled narratives, often ingenius idiosyncratic products, rarely deserve the attention of the effective manager. For efficient, effective operations, what man *knows* is of much greater value than what man *thinks*. New systems permit the personnel manager to record his memories laterally, longitudinally, and with great accuracy. Automated systems are not prone to forgetting.

The program for the Anniversary Symposium was evolved within the constraints of the several purposes of the event. All presentations were to be made by eminent persons who had been members or were members of or were in close relationship with the Personnel Research Laboratory. Since so many prominent persons had formed this population, there was, in effect, no constraint. In fact, the problem was one of deciding which of the many should be invited to speak. Topics for the program were generated prior to the determination of appropriate speakers, but the individual was then granted complete latitude in his presentation. Most chose to remain relatively close to the topic suggested.

Another idea used in selection of speakers was the attempt to have representation from military, governmental, business, and academic sources. As might be expected, the military systems are disproportionately represented. Since the lay personnel subsystems of the military services are lucrative targets for research and development, the representation is not an inequity in a real sense. The work being done with the military subsystems has almost direct translation to the large corporation. There is no doubt that many major concerns are initiating programs for computer-based personnel systems dependent upon inputs in real time. The problem appears to be one of locating managers trained in the many facets of such programs. The widespread distribution of ideas, information, models, and suggestions presented during the Symposium should serve an important function in developing personnel researchers, planners, and managers in both academic and operational settings.

The papers presented have been edited only with the intent of providing some common form for the reader. Even this design was modified for one paper to keep it in the style appropriate to its discipline. The order of presentation of the papers at the Symposium has been altered for the purpose of providing some common chapter characteristics. Admittedly, the commonness of the chapter breakdowns may not always be too apparent.

In addition to excellent reviews of activities underway in the field of personnel research, there are some relatively novel expressions for systems advancement. Dr. Leiman's suggestions with regard to a new form of job engineering are highly creative and socially attuned. Dr. Dailey presents a new approach to mental testing and does so with considerable clarity. Those topics which generally deal with the future of personnel research are for the most part in the form of tributes to the USAF Personnel Research Laboratory.

The evidence for candor and freedom of expression is quite obvious in the papers by Dr. Carlson and Dr. Carter. In all, the papers foliated the Symposium proceedings most adequately. Their continued use should be equally valuable to all those who are fortunate enough to read them.

Angelo L. Fortuna
Symposium Coordinator

ACKNOWLEDGMENTS

The remarkable success enjoyed by the Twenty-fifth Anniversary Symposium is a testimonial to the dedication of the many people involved in its production. Recognition of each of them is not possible, but certain individuals must be noted. Major Angelo L. Fortuna served resourcefully and tirelessly as Symposium Coordinator; Mr. Leonard N. Jaimes did yeoman's service in the handling of all financial matters and as the Laboratory troubleshooter; Mr. Jack Harmon, of Southwest Research Institute, polished all arrangements for the Symposium into coordination so excellent that it was not apparent. Captain George H. Walther should be especially recognized for his arrangements for securing and operating the remote inquiry station for the Air Force Military Personnel Center computer system which operated during the course of the Symposium. Miss Pat Lyster of the Personnel Research Laboratory directed composition of the proceedings manuscript, which was prepared for the composer by Mrs. Helen Widner and Mrs. Joann Archer. Final composition of this manuscript was made possible through the assistance of Mr. Rees Walker, Mr. Howard Kochwelp, Mr. William Bernard, and Mr. Dan Maher, technical consultants from the IBM Corporation. Our appreciation of the services provided by those named above and many others who served is gratefully acknowledged.

The Symposium was made possible by the support of Colonel James H. Ritter, Commander of the Laboratory, who lent confidence and enthusiasm for us to see it through.

J. W. Bowles, Jr.
Symposium Director

TABLE OF CONTENTS

RECENT PERSONNEL RESEARCH IN THE ACADEMIC WORLD

John T. Cowles
The University of Pittsburgh

Introduction

This review was prepared with the intent of limiting it to inclusion of work done by persons identifying themselves with a university and mainly to publications dated only during the past year or two. University projects done under contract with the Armed Services or Governmental agencies were also excluded, but a fairly large portion of the research to be commented upon was, of course, done often with National Institute of Health (NIH) or other grants in industrial or business settings, the common laboratory of personnel psychologists based in the university. Published studies of university personnel problems involving university employees as subjects are rare indeed, although college students are sometimes used as convenient respondents in published studies of methodological problems.

It will be noted, too, that most of the studies cited were published in the two principal civilian personnel research journals, the *Journal of Applied Psychology* and *Personnel Psychology*, or in selected recent books. The *Annual Review of Psychology* is an excellent secondary source, but in keeping with a biennial topical rotation scheme, each volume contains reviews only of selected topics, such as Personnel Selection one year and Personnel Management the next. A few recent textbooks and handbooks have been included, since they represent a characteristic contribution of university scholars who usually have been involved in research of their own. There is no well-defined field of personnel psychology today, no division of the American Psychological Association with that title, no standard textbook with that title, it is a composite or blend of various specialties investigating problems in such areas of work and employment as selection and recruitment, classification and placement, job analysis, job evaluation, organization, motivation and morale, proficiency measurement, training, human engineering, accidents and safety, pay and satisfaction, human relations, training and development, separation and retirement, and so on. The areas of human relations and organization and motivation have greatly expanded in recent years, largely through the increased participation of the rapidly developing social sciences in studies of these particular personnel areas. Computer technology, simulation, and mathematical models are also invading the general field. The systems approach, too, with man-machine systems, communication systems, organizational dynamics, and the complexities of automated learning have all changed the face of personnel research. It is utterly fascinating to a present-day reviewer of the field, perhaps even a bit bewildering to one who entered the field of personnel research before these new developments had begun to flourish. In fact, we find that some of these developments have become new specialties of their own, with a language of their own, journals of their own, and congenial workers who sometimes appear preoccupied with self-fascination if not mutual mystification. So, it is with some diffidence that I peer softly into some of these ganglia of the great nerve net of personnel psychology of today.

Recruitment and Selection

Within the scope of this survey of recent research, only one study dealt with recruitment as such. This study by Carroll (5) concerned the relationship of various college graduate characteristics to recruiting decisions. Earlier studies had usually questionnaired employers regarding their preferential ratings of the general traits of potential employees, such as intelligence, college grades, extracurricular activities, and so on. This particular study took a more direct tack and assessed the characteristics of 211 male graduates as shown more objectively in their records, plus a rating of their appearance in the official photograph on file. It was found, incidentally, that this latter rating could be reliably done on the general scale dubbed "appearance." These assessments included age, grades, height, major field, marital status, fraternity membership or not, military draft status, leadership positions in college, membership in college organizations, work experience, work preference, weight, and weight-height ratio. Criteria of recruitment success included such items as numbers of visit offers, job offers, visit-interview ratio, job-visit ratio, and a combined criterion. Only three of the assessed traits related significantly to any of these five criteria: marital status, appearance, and office experience. Appearance alone correlated significantly with as many as three of these criteria, with consistently low positive correlations. In short, what employers have traditionally reported they value as to college graduates' characteristics, namely, intelligence and grades rather than employment experience or test scores, is hardly in accord with their actual recruitment practices.

Selection studies seem to have become somewhat more sophisticated, but their results are hardly more successful or promising than the fuzzy studies of a decade ago. But there are distinctly encouraging signs. Not often do we find today the former one-man style of "criterion development," but rather multidiscipline criterion analysis, especially among motivational or social variables, rather than among the well-worn cognitive and skill variables of the past. Global criteria, with a single measure of "job success," are quite suspect today. But I must admit numerous studies still resort to easy criteria, such as supervisors' ratings of managers' effectiveness, and take for granted that these criteria have some kind of ultimate or intrinsic construct validity without further empirical verification. Who is to argue with the boss? Nash (42), for example, developed an empirical Strong Vocational Interest Blank (SVIB) key for predicting executives' overall rankings of managerial effectiveness, using a sample of 461 managers from 13 varied Minnesota companies. The 57 items showed a correlation of .33 on a hold-out sample. Of note was the description of interest patterns shown by the valid items: effective managers seemed to be characterized by interests in independent and intensive thought, tangible outcomes, physical and social activities, some risk, less regimentation of time, interest in non-technical and non-agricultural pursuits, and less interest in activities requiring long periods of attention and concentration on details, as well as less interest in service-oriented, humanitarian-oriented, and esthetic pursuits. The author concludes that such a study of interest patterns of effective managers should be of help in vocational guidance or possibly in selection. We might add that an analytic view of the criterion "effective" would be essential before we could put to work these results on interest patterns.

Another recent study by Schletzer (48) explored the relationships between standard (and appropriate) SVIB keys and three job satisfaction inventories for 185 male graduates from various professional schools, using SVIBs that had been filled out while the subjects were in the 12th grade, as well as current SVIB retakes. Of the 56 relationships investigated, only one correlation was significantly different from zero, namely the SVIB engineering key and the Job Dimension Inventory for the engineering subsample. Although the later SVIBs showed greater congruity with the appropriate occupational interest keys, they were no closer to present job satisfaction than the earlier SVIBs. These results are hardly promising but are in accord with other studies of the relation of standard keys of interest inventories and ordinary measures of job success. Somewhat more successful was a study by Megargee and others (40) relating the California Psychological Inventory (CPI) dominance (DO) scale to a person's assumption of a leadership role, or his not assuming that role, in a paired-person problem situation requiring one of the pair

to serve as leader. When the problem instructions emphasized the task itself, only a chance proportion of high "dominance" types assumed the leader role, whereas, when the instructions emphasized the leadership role, the persons with high DO scores assumed leadership in 90 per cent of the pairs. We may conclude that this CPI scale was predictive in this situation. Now, why oh why, will not this investigator or someone else try out this apparently promising personality scale in other situations? We can only hope.

Guion and Gottier (24) have, in fact, made a recent summary of validity studies of personality measures in personnel selection. They found that only about one-third of recently reported studies bothered to check *predictive* validity, and only in seven of the 30 studies in that one-third had the researchers troubled to perform a cross-validation. Of the studies published in the *Journal of Applied Psychology* and *Personnel Psychology*, in the entire 12-year period (1952-1963), something less than one-eighth of the validity studies resulted in significant coefficients. There was no clear pattern of suitable criteria, and the few significant validities were associated with custom-built instruments for particular situations. The authors of this review concluded that raw empiricism is still a better tool, that is, research-based selection test development, rather than the application of existing "standard" instruments. In short, we can only agree with them that "some personality instruments may work for some purposes in some situations," but don't count on it! A very recent general review of occupational aptitude tests and their validity has been published by Ghiselli (18).

Some of the newer techniques, like the Frederiksen in-basket technique, have not proved as fruitful as we might have earlier anticipated. Two other general reviews tend to concur with these discouraging general conclusions, namely the 1963 review by Dudek (11) and the 1965 review by Biesheuvel (2), both in the *Annual Review of Psychology*. For example, the latter cites the evaluation by Locke and Hulin (36) of all available validity studies of the well-publicized Activity Vector Analysis, in which it was found that in only two out of the 18 studies were the procedures adequate for validation of a presumed industrial selection technique. In neither of those two studies was the presumptive device of sufficient validity to be of any practical use for its intended selection purpose, namely selecting insurance salesmen. He also mentioned the similar negative evaluations by Campbell and co-workers of various devices for selecting high-level personnel, including the unhappy finding that even objective cognitive tests had no useful relation to supervisors' ratings of intellectual functions.

It is apparent, even granting that we might find adequate criteria and develop likely predictors, that our classical validation models, based on simple correlation or ordinary multiple correlation, are outmoded. We should be seeing in the near future some fair trials, in practical personnel selection or classification problems, of the more powerful computerized multivariate techniques such as canonical correlation, which can handle the numerical complexities of multiple prediction of multiple measures of the criterion. Horst's (27) new treatise on factor analysis of data matrices is a major academic contribution in this area and should have wide applicability. However, it appears that we have not yet developed equally advanced, systematic approaches to identifying and measuring the more significant criterion behaviors. The present reviewer shudders at some of the recently reported applications of high-powered computer techniques to trivial data, such as a principal-components factor analysis of items in a rating scale, based on only a handful of respondents! We shall have to endure a coming rash of ill-advised studies using these magnificent data processing and computing techniques, solaced by the hope that a nugget may turn up among these studies. Some interesting relationships have been discovered in the recently popular search for moderator variables, but we still need more systematic approaches, perhaps insights, to their discovery and use.

Studies of Techniques and Methodology

There continue to be published new variations of old measurement techniques, for example in studying the cues used by interviewers. Maier (38) had a person act the role of an honest or dishonest interviewee, and he discovered that untrained interviewers could detect the dishonest interviewee but could

not identify the relevant cues. This suggests that we might seek out and study such cues more thoroughly to determine the bases of interview judgments, and perhaps thereby increase the validity of the interview as a predictive device. In another study by Blumberg and others (3), various rating scale formats were compared and evaluated. It was found by analysis of variance that there were no appreciable differences, in an actual measurement situation, among formats of different direction (up-down, left-right), graphic or numeric, and whether the traits were rated one at a time, or the ratees one at a time for all traits, or whether the rater was permitted a free choice of order. We do wonder a bit if the results were much affected by the kind of rater (college students) or choice of persons to be rated (Lincoln, De Gaulle, Liz Taylor, Khrushchev, etc.). However, a study by Peters and McCormick (44) compared the reliability of numerically anchored and job-task anchored rating scales and found the latter type significantly more reliable. This latter study, by the way, was part of a larger and rather impressive study of various types of scales for the rating of job activities, which has been carried out at Purdue by these men. Findley (14) has contributed a new index of mutual acceptance of individuals in a group, which may have useful application in personnel research. Under his assumption that certain best groups are those in which there are no isolates, the index puts a premium on mutual ratings where all individuals are chosen relatively as frequently, rather than where a few distinctive ones get all the votes. It is a clever index, similar to his earlier Index of Participation and is called the Index of Assimilation. He illustrates its use with live data and provides a tentative normative scale for interpretation.

It was gratifying to find at least one study on research methodology, namely, by Rosen and Sales (46) on the effects of behavioral field research on the work performance of factory employees. This study first noted no effects of observational or survey research operations on overall worker productivity; but when certain hypothesized moderator variables such as age, urban-rural, union activity, and F-scale (authoritarianism) scores were investigated, it was found that at certain levels on these variables workers were significantly affected and the productivity results thereby contaminated. In general the workers who were either younger, or rural, or more active in the union were the ones more affected; however, no differences were associated with F-scale scores. This study, like some of the earlier ones by Ghiselli, not only illustrates the usefulness of the moderator-variable approach plus insights, but also is a lesson to other research workers who might claim no significant effects of a given variable just because it produces no overall effect.

Before leaving this section on techniques and methodology, we should note that academic personnel psychologists commonly contribute, in the form of textbooks, scholarly work of considerable value in the training of new workers in fields of personnel psychology, and in summarizing aspects of those fields. An example would be the excellent new text on personnel testing by Guion (22), or the recently revised general texts on industrial psychology by Gilmer (19), by Maier (37), and by Tiffin and McCormick (51). There are also numerous excellent texts in the more specialized sub-fields, such as Ernest McCormick's second edition of *Human Factors Engineering* which appeared in 1964 (McGraw-Hill). A book of quite a different type but worthy of special note is the newly enlarged second edition of that stimulating application by Cronbach and Gleser (10) of decision theory in the use of psychological tests for personnel decisions. This book has practical and theoretical values for anyone involved in appraising manpower costs or utility in selection or placement. We shall later mention other specific works, often combining scholarship and original research, dealing with theory of organization, group process, motivation, and so on. There is also a quantity of less notable but nonetheless necessary books, written by academic psychologists, at simpler levels of discourse for explaining the technical field to the busy executive or other laymen.

Training

No attempt was made to cover the enormous field of training in this brief review. Excellent summaries are found recurrently in the *Annual Review of Psychology* and the *Review of Educational Research*, although these deal mainly, on the one hand, with the more fundamental levels of

experimentation and theory, and on the other hand with applications to the classroom situation. Programmed learning, automated learning, sensitivity training and other new techniques of learning have been greatly stimulated by academic, laboratory research, even animal studies on operant conditioning, for example, while further stimulation has come from other sources and certainly from current needs to make teaching and learning more effective in the face of teacher shortages. A recent conference summary by Glaser and participants (20) on training research is a notable collection of reports and viewpoints on the present status of training research in a wide variety of personnel fields. Academic research is well represented. The monumental handbook edited by Gage (16), summarizing research on teaching, should not be overlooked. It has obvious applicability to personnel training programs and to the proper study and evaluation of methods of instruction; it is more pertinent to training research than its title may suggest. Other useful ideas for new training methods have come from such unlikely sources as the laboratory method, or the practice of psychotherapy in the form of psychodrama or, better, "role playing." Schein and Bennis (47) have developed a "laboratory" training method of broad application, with fascinating possibilities for observing and evaluating personnel development, and personal and organizational change. Corsini, Shaw, and Blake (8) have written up role playing in business and industry, with practical illustrations of a how-to-do-it kind that suggest applicability to personnel program needs from diagnosis and information concerning individual or group functioning to executive leadership training. However, we hasten to add, we would now like to see some solid research on the actual validity of this often entertaining technique. Are its promises fulfilled? Admittedly there is a constant need for a bridging effort on the part of psychologists to determine training requirements in applied situations, to seek out those research findings that may be useful, and then to translate and apply them with experimental cautions. But how often is the user's faith substantiated by more than enthusiasms? It is my opinion that the academic world has set a rather bad example, in the past, of perpetuating old or introducing new methods of education without adequate evaluation. Rather glaring examples can be seen in medical education, where large foundation subsidies have been made for new curricular ventures without providing for fair evaluation. The retrospective looks have resembled rationalization of the changed methods. On a more optimistic note, the new educational research and development centers on several of our university campuses have attracted brilliant young investigators with a cool eye to evaluating new methods before applying them in the educational situation. These centers will bear watching by business, industry, and the military establishment. Another significant development in the academic world, as educational costs frighteningly increase, is the new look at dollars-and-cents educational input and output. Faculty and administration are teaming to find ways to make university educational programs less of a losing venture.

Motivation, Success, and Satisfaction

The complex relationships among motivation to work, job performance, and job satisfaction variables have become the focus of a greatly augmented bulk of personnel research in recent years. We have found out, through research, that there is no simple relation between pay and productivity, pay and satisfaction, and so on. With a growing realization of the complexity of human factors in the work situation, new approaches and various newly interested disciplines are now investigating these relationships. Psychologists are studying needs and satisfactions, response styles, expectations, frustration, conflict, and the like; sociologists are studying social and community and cultural factors; human engineers are studying the fits and misfits, and the why of these, in man-machine systems; economists are studying pay plans, personnel benefits, and other perquisites not alone with regard to profit and loss but also with regard to the context of industrial aspirations and development. The waters are indeed murky but well agitated.

Representative studies during the past year or so reflect the amazing variety of current problems and approaches. Ghiselli (17), for example, investigated the maturity of self-perception of managers and found a curvilinear relationship; it is more important for success for the manager to have a pattern of self-perception that corresponds with others of his own age than characteristic of younger or older managers. By an

ingenious checklist of paired adjectives of equal social desirability or undesirability, he was able to specify the significant differences between the self-percepts of older and younger managers. Hulin and Smith (30) found that women tend to be less satisfied than men with their jobs, but that it is probably a function of associated variables, such as salary, place of work, hope for promotion, and qualification level, rather than any sex difference as such. Hulin (28) also studied the relationships of such community characteristics as prosperity, slums, unemployment, productive farming, and community decrepitude to job satisfaction and productivity. He found that average satisfaction scores were unrelated to group productivity but were negatively related to the prosperity of the community. Pay satisfaction scores tended to be more negatively related to the prosperity of the community than did the other aspects of job satisfaction. It seems that in a poor community the job alternatives are so few and so unattractive that the workers remain more satisfied than their counterparts in a more prosperous community, with its attractive competitive opportunities for employment. In short, community characteristics may be moderator variables in the relation between pay and satisfaction. In similar vein, Centers and Bugental (6) found that the perceived relative values of extrinsic and intrinsic job components are related to job level; the higher the job level, the higher the worker values the intrinsic job components; the lower the level, the more the worker values the extrinsic job components. Sex differences in these relationships were found to be negligible; however, women tended to place a higher value on "good co-workers," while men placed a higher value on "opportunity to use talent or skill."

Hulin (29) found that job satisfaction scores of women clerical workers might predict turnover, since those that left the job had had lower satisfaction scores, and these were based on attitudes toward work, pay, promotion, co-workers, and supervision. Lawler (34) found that ability is a moderator variable in the relationship between managers' perception of the relation between performance and pay and his actual performance as rated by supervisors. Among high-ability managers, there is a clear relation between their perceptions of job factors important in pay determination and their job performance, whereas among low-ability managers, there is no apparent relation of this kind He concludes that the multiplicative concept seems to be confirmed, namely, that job performance is a function of ability times motivation. Lawler (35) further studies the relation of managers' attitudes toward how their pay is determined and their satisfaction. He found that high dissatisfaction was associated with a lack of congruence between the manager's attitude toward how his pay should be determined and how it actually was determined. In general he confirmed that there may be little relation between an incentive pay system and the workers' perception of the incentive values of such a plan.

Eran (13) studied the relationship between self-perceived personality traits and job attitudes among middle-level managers. He found that managers scoring higher on personality traits believe they are fulfilling their inner psychological needs (Maslow variety) to a greater degree, and perceive more job satisfaction in terms of those same traits. It appears that they are more satisfied because their abilities are more compatible with job requirements as self-perceived. Paine and others (43) studied the need satisfactions of managerial level personnel in a Government agency. First, he compared field managers with office managers of similar level; he found that the former obtained greater satisfaction with certain higher-level needs, such as self-esteem, development, and independence of action. He also compared these results with the need satisfactions reported by managers of similar level in non-governmental organizations. The satisfaction of the Government managers was less across all need items, including social needs and need for security, than the satisfaction of private managers. However, an impending reduction in force (RIF) in the Government agency may have been influential in the latter results. At any rate, information of this kind would be useful to those who are either counseling or selecting potential managers.

Weissenberg and Gruenfeld (54) have been investigating the complex relationships among leadership dimensions and cognitive style, in a civil service setting, using various paper-and-pencil instruments. One measure, Witkin's Embedded Figures Test (EFT) seemed quite promising, since it reflects an objectively measurable personality dimension called "field dependency" and can hardly be faked; it shows a substantial

curvilinear relation (eta = .87) with Fiedler's rating form "Esteem for the Least Preferred Co-worker" (LPC). It was found that Fleishman's Leadership Opinion Questionnaire (LOQ) is highly susceptible to faking. The same investigators have also been studying such variables as supervisory characteristics and attitudes toward performance appraisals (22). Supervisors high in accepted supervisory qualities are more favorably disposed toward the appraisal and development of their subordinates, even when the appraisal scores of those subordinates has little to do with pay and promotion (as seems to happen in a civil service situation). The best predictors of attitude toward the appraisal system and corrective use of it were certain scales of Ghiselli's Self-Description Inventory and Fleishman's Leadership Opinion Questionnaire, rather than such old standbys as age, intelligence, or F-scale scores. In fairness to these active research workers, it should be mentioned that they do appear to be interested in the basic behavioral processes as well as in the correlational data. Gruenfeld (21) also recently investigated the bases of favorable attitudes toward a management development program and found that when it was contributory on the part of the employee, the perceived benefits were significantly increased. A related study is the one by French, Kay, and Meyer (15) who found that when employees really participate in planning their own development, not just think they are participating, then the various indices of satisfaction and improved human relations with supervisors show gains. In concluding this section on empirical studies of factors in job success or job satisfaction, we should cite some outstanding reviews of the research literature in this area. Nash (41) recently reviewed 47 published articles on the vocational interests of effective managers. He noted that the standard scales of the Kuder Preference Record and the Allport-Vernon-Lindzey Study of Values when used for predicting managerial success are superior to the standard scale of the Strong Vocational Interest Blank. Additionally, he pointed out the familiar faults in the great bulk of such validation studies: lack of cross-validation, low Ns, and flimsy documentation, even in the more widely publicized studies. Sells (49) and Porter (45) have prepared the latest reviews on Personnel Management in the *Annual Review of Psychology*.

Theories of Motivation in the Work Situation

There have been several recent contributions to theories of motivational factors in the work situation, but none has been either consistently supported by the experimental literature or has had sufficient generality to be of truly landmark quality. The theoretical formulations of Herzberg and associates, in various published articles since 1959, have now been given more formal presentation together with selected experimental evidence by Herzberg in his recent book of ambitious title, *Work and the Nature of Man* (25). His theoretical constructs have been controversial for some time, and within a few months of this publication at least two new studies have been reported which seem to refute Herzberg's two-factor theory of "motivators" (intrinsic influences) and "hygienes" (extrinsic influences) as differential sources of satisfactions and dissatisfactions. Burke (4) had college students rank their perceived importance of five motivators and five hygienes, and Coombs' unfolding technique was used to scale the data. There was no evidence of two independent factors underlying job satisfaction and job dissatisfaction. He also reviewed the experimental literature, and among 14 relevant studies failed to find evidence in support of Herzberg's theory. The second study was by Wernimont (55), who obtained responses to forced-choice and free-choice items presented to 50 accountants and to 82 engineers, where the items concerned past satisfying or dissatisfying job situations. Both groups endorsed more intrinsic than extrinsic items when describing both kinds of situations, and there was no evidence that either category of variables was unidirectional, as posited by Herzberg. Expectations have a strong influence on the extent of satisfaction with job factors. Another more comprehensive survey of work and motivation theory will be found in the recent book by Vroom (53). Other contributions to motivation theory are numerous in the context of other subfields of personnel psychology, for example in discussions of organizational behavior cited below.

Organizational Behavior and Theory

Under this caption has fallen the greatest proportion of recent studies by academic research workers that have come to our attention. It reflects the markedly increased interest in social psychology aspects of the employment situation by the various social sciences, along with computer technologists and applied mathematicians, who have also turned considerable attention to the development and exploitation of mathematical models of group and management behavior. It has been an astoundingly blossoming field and, now, almost a specialty in itself. A few very recent empirical studies will at least illustrate the diversity of interests. Jerdee (31) has neatly studied work-group versus individual differences in attitude to help answer the question, "Is attitude variability a function of group or individual variability?" From a sample of 190 employees representing 38 work groups in three manufacturing plants, he obtained responses to a "Triple Audit Employee Attitude Scale," developed at the University of Minnesota. An analysis of variance showed that group differences in attitude were insignificant compared to individual differences; the author concluded that the appropriate unit for study of such attitudes is the individual rather than the group. The attitudinal statements, by the way, did include group-related job factors such as supervision, as well as items concerning the work itself.

Another study, this one by Dunteman (12), explored the interrelations among 84 variables pertaining to company organizational characteristics, management attributes, incentive conditions, worker characteristics, personnel performance, and organizational functions. A multiple-choice questionnaire of 84 items was the source of data, drawn from 234 responding personnel directors at as many different manufacturing firms. Interestingly, this was only an 8 per cent return, the best that could be retrieved from almost 2,000 companies solicited! A factor analysis, with rotation to simple structure, indicated that there might be 14 factors, such as size of organization, economic growth, tardiness *vs.* family responsibility (a bi-polar factor), pay-skill level, personnel tenure, and so on. Curiously, productivity, job aversion, and theft were mutually independent. A third study by Shore (50) focussed on the role of the arbitrator, again using questionnaire returns based on 25 statements, from a sample of 28 accredited labor arbitrators, 33 management representatives, and 40 union officials. Major dimensions of the arbitrator's role as perceived by each of these three groups were defined, after a cluster analysis of items which resulted in these 5 clusters: adherence to precedent (favored by the arbitrators themselves), prophylactic purposes (favored by management), liberality (favored by management), elicitation of facts (also favored by management), and procedural formality (favored by the arbitrators). In each instance except the last, the other two respondent groups were significantly lower in preference, although they might not have differed significantly between themselves.

Many of the relevant studies in this organizational area have been reported in books which present organizational theory together with empirical evidence or illustrations. Theory in these books has ranged widely, and perhaps a bit wildly, with such facets as the following: organizational stress (33), ingratiation (32), small group process (26), sociological factors (52), optimization in the work situation of the criterion values of our society (1), and group processes in decision making (7), plus some of our earlier references to theories of worker motivation. Perhaps the time has come, in view of the mass of well-selected (!) supporting data in each instance, to attempt an integration of these partial theories into something more embracing, and look for the gaps or inconsistencies in more general theory. Certain mathematical models are being developed, of varying promise or practicality, especially in regard to small group behavior, as notably reported by the Stanford group and persons of related interest (9).

In closing we shall dare to cite a new publication which we have not yet seen, but which the publishers indicate will be very useful to workers in this last hugh area of personnel psychology: *Handbook of Organizations* under the editorship of J. G. March (39). This is promised as an extensive general survey of theories and methodologies relating to organizational and interpersonal behavior. It will have to be, I might say, the review to end my own review.

REFERENCES

1. **Bass, B.M.** *Organizational psychology.* Boston: Allyn & Bacon, 1965.
2. **Bienheuvel, S.** Personnel selection. *Annu. Rev. Psychol.*, 1965, 295-324.
3. **Blumberg, H.H., DeSoto, C.B., & Kuethe, J.L.** Evaluation of rating scale formats. *Personnel Psychol.*, 1966, 19, 243-259.
4. **Burke, R.J.** Are Herzberg's motivators and hygienes unidimensional? *J. appl. Psychol.*, 1966, 50, 317-321.
5. **Carroll, S.J., Jr.** Relationship of various college graduate characteristics to recruiting decisions. *J. appl. Psychol.*, 1966, 50, 421-423.
6. **Centers, R. & Bugental, Daphne E.** Intrinsic and extrinsic job motivation among different segments of the working population. *J. appl. Psychol.*, 1966, 50, 193-197.
7. **Collins, B.E. & Guetzkow, H.** *A social psychology of group process for decision making.* New York: John Wiley & Sons, 1964.
8. **Corsini, R.J., Shaw, M.E., & Blake, R.R.** *Roleplaying in business and industry.* Glencoe, Ill.: Free Press, 1961
9. **Criswell, Joan H., Solomon, H., & Suppes, P. (Ed.)** *Mathematical methods in small group processes.* Stanford, Calif.: Stanford Univ. Press, 1962.
10. **Cronbach, L.J. & Gleser, Goldine C.** *Psychological tests and personnel decisions.* (2nd Ed.) Urbana, Ill.: Univ. of Illinois Press, 1965.
11. **Dudek, E.E.** Personnel selection. *Annu. Rev. Psychol.*, 1963, 261-284.
12. **Dunteman, G.H.** Organizational conditions and behavior in 234 industrial manufacturing organizations. *J. appl. Psychol.*, 1966, 50, 300-305.
13. **Eran, M.** Relationship between self-perceived personality traits and job attitudes in middle management. *J. appl. Psychol.*, 1966, 50, 424-430.
14. **Findley, W.G.** Group *vs.* individual sociometric relations. *Intl. J. of Sociometry and Society*, 1966, 5 (1-2), 60-66.
15. **French, J.R.P., Jr., Kay, E., & Meyer, H.H.** Participation and the appraisal system. *Human Relations*, 1966, 19, 3-20.
16. **Gage, N.L. (Ed.)** *Handbook of research on teaching.* Chicago: Rand McNally, 1963.
17. **Ghiselli, E.E.** Maturity of self-perception in relation to managerial success. *Personnel Psychol.*, 1964, 17, 41-48.
18. **Ghiselli, E.E.** *The validity of occupational aptitude tests.* New York: Wiley. 1966.
19. **Gilmer, B. von H. (Ed.)** *Industrial psychology.* (2nd Ed.) New York: McGraw-Hill, 1966.
20. **Glaser, R. (Ed.)** *Training research and education.* Pittsburgh: Univ. of Pittsburgh Press, 1962.
21. **Gruenfeld, L.W.** Effects of tuition payment and involvement on benefit from a management development program. *J. appl. Psychol.*, 1966, 50, 396-399.
22. **Gruenfeld, L.W. & Weissenberg, P.** Supervisory characteristics and attitudes toward performance appraisals. *Personnel Psychol.*, 1966, 19, 143-151.

23. Guion, R.M. *Personnel testing.* New York: McGraw-Hill, 1965.

24. Guion, R.M. & Gottier, R.F. Validity of personality measures in personnel selection. *Personnel Psychol.*, 1965, 18, 135-164.

25. Herzberg, F. *Work and the nature of man.* Cleveland, O.: World Publ., 1966.

26. Hopkins, T.K. *The exercise of influence in small groups.* Totowa, N.J.: Bedminster Press, 1964.

27. Horst, P. *Factor analysis of data matrices.* New York: Holt,Rinehart & Winston, 1965.

28. Hulin, C.L. Effects of community characteristics on measures of job satisfaction. *J. appl. Psychol.*, 1966, 50, 185-192.

29. Hulin, C.L. Job satisfaction and turnover in a female clerical population. *J. appl. Psychol.*, 1966, 50, 280-285.

30. Hulin, C.L. & Smith, P.C. Sex differences in job satisfaction. *J. appl. Psychol.*, 1964, 48, 88-92.

31. Jerdee, T.H. Work-group versus individual differences in attitude. *J. appl. Psychol.*, 1966, 50, 431-433.

32. Jones, E.E. *Ingratiation: a social and psychological analysis.* New York: Appleton-Century-Crofts, 1964.

33. Kahn, R.L., Wolfe, D.M., Quinn, R.P., & Snoek, J.D. *Organizational stress: Studies in role conflict and ambiguity.* New York: Wiley, 1964.

34. Lawler, E.E., III. Ability as a moderator of the relationship between job attitudes and job performance. *Personnel Psychol.*, 1966, 19, 153-164.

35. Lawler, E.E., III. Managers' attitudes toward how their pay is and should be determined. *J. appl. Psychol.*, 1966, 50, 273-279.

36. Locke, E.A. & Hulin, C.L. A review and evaluation of the validity studies of Activity Vector Analysis. *Personnel Psychol.*, 1962, 15, 25-42.

37. Maier, N.R.F. *Psychology in industry.* (3rd Ed.) Boston: Houghton-Mifflin, 1965.

38. Maier, N.R.F. Sensitivity to attempts at deception in an interview situation. *Personnel Psychol.*, 1966, 19, 55-66.

39. March, J.G. (Ed.) *Handbook of organizations.* Chicago: Rand McNally, (No date).

40. Megargee, E.I., Bogart, Patricia, & Anderson, Betty J. Prediction of leadership in a simulated industrial task. *J. appl. Psychol.*, 1966, 50, 292-295.

41. Nash, A.N. Vocational interests of effective managers: A review of the literature. *Personnel Psychol.*, 1965, 18, 21-37.

42. Nash, A.N. Development of an SVIB key for selecting managers. *J. appl. Psychol.*, 1966, 50, 250-254.

43. Paine, F.T., Carroll, S.J. Jr., & Leete, B.A. Need satisfactions of managerial level personnel in a government agency. *J. appl. Psychol.*, 1966, 50, 247-249.

44. Peters, D.L. & McCormick, E.J. Comparative reliability of numerically anchored versus job-task anchored rating scales. *J. appl. Psychol.*, 1966, 50, 92-96.

References (*Continued*)

45. Porter, L.W. Personnel management. *Annu. Rev. Psychol.*, 1966, 395-422.
46. Rosen, N.A. & Salen, S.M. The effects of behavioral field research on the work performance of factory employees. *J. appl. Psychol.*, 1966, 50, 165-171.
47. Schein, E.H. & Bennis, W.G. *Personal and organizational change through group methods: The laboratory approach.* New York: Wiley, 1965.
48. Schletzer, Vera M. SVIB as a predictor of job satisfaction. *J. appl. Psychol.*, 1966, 50, 5-8.
49. Sells, S.B. Personnel management. *Annu. Rev. Psychol.*, 1964, 399-420.
50. Shore, R.P. Conceptions of the arbitrator's role. *J. appl. Psychol.*, 1966, 50, 172-178.
51. Tiffin, J. & McCormick, E.J. *Industrial psychology.* (5th Ed.) Englewood Cliffs, N.J.: Prentice-Hall, 1965.
52. Turner, A.N. & Lawrence, P.R. *Industrial jobs and the worker.* Boston: Harvard Univ. Press, 1965.
53. Vroom, V.H. *Work and motivation.* New York: Wiley, 1964
54. Weissenberg, P. & Gruenfeld, L.W. Relationships among leadership dimensions and cognitive style. *J. appl. Psychol.*, 1966, 50, 392-395.
55. Wernimont, P.F. Intrinsic and extrinsic factors in job satisfaction. *J. appl. Psychol.*, 1966, 50, 41-50.

II

RECENT ADVANCES IN PERSONNEL PSYCHOLOGY IN INDUSTRY

Harold A. Edgerton
Performance Research, Inc.

The purpose of this paper is to take a look at advances, changes and trends in Personnel Psychology in industry. It is not an attempt to cover a specific period of years, but rather to touch on those things that seem to be reflecting significant developments and pointing to needed researches that may be showing just over the horizon. It may be that we could hasten the exploitation of some of these.

In the area of personnel selection, there has been little or no change in the validity of aptitude tests over a long period of years, certainly over the last two or three decades. We have a great many journal articles reporting the validity coefficients for various tests in various situations. There is need for such studies and such publication to continue; but there is also need for more adequate analysis of the criterion, of the job situation, of the employment market, and of the value system used by management in order to generalize from this kind of information. Otherwise, we have only a list of applicabilities of tests in specific situations, and these situations are seldom described adequately. This is a "fact of life" rather than a professional failure.

With regard to the validity of tests: correlation coefficients up to .65 or .70 are sometimes obtained, but we become suspicious of those coefficients above .70, wondering in what ways there may have been error or bias in the measurement process to produce such coefficients; on the other hand we act as if ideally there should be a correlation of unity between aptitude tests and job performance. As yet, we have not really faced the question as to what extent one expects to find all of the factors affecting quality of on-the-job performance within the bounds of the skin of the individual worker and, hence, to be available as predictors.

Personnel testing in industry is undergoing considerable re-examination. There are issues regarding the applicability of aptitude tests to the culturally deprived, to the disadvantaged, to minority groups, to language-handicapped groups, and so on. The approaches to this problem have not as yet clarified into trends, unless it is the recognition that measurement must be suited to the situation. Aptitude test scores, by their nature, are affected by the prior experiences and learning of those taking the test. We see no good way of removing those portions of the variance in test scores which are due to economic or social disadvantage, or to different experiences, or to a lack of appropriate experiences. This points to the need for improvement in the interpretation of scores, the interpretation adapted or adjusted by level of test score and by the kind of individuals for whom the scores are being applied. A second type of effort is to develop or obtain tests appropriate to the kind and group. For example, tests with high verbal content may not be

the best measures of the ability of those who dropped out of school before becoming substantially literate. Use of such tests must be based on proper validation. Both the proponents and opponents of the use of aptitude tests in the hiring and placement of disadvantaged groups seem to agree that empirical validation is necessary.

There is a problem related to the fact that with appropriate training, there is a substantial improvement in aptitude test scores. With these changes, do we get corresponding change in the job potential or job performance potential of the individuals? I would like to believe so. But we need, and I am sure we will be seeing, research aimed at obtaining answers in this area. There have been many studies showing that, with practice, there is an increase in the scores on aptitude tests. We need to compare the validity of these "augmented" aptitude test scores with the validity of scores not deliberately augmented by practice on materials similar to the test content. For example: Two years ago, in a six-week training program in basic education and communication skills, a group of disadvantaged trainees improved their performance on the language score of the California Test of Mental Maturity an average of 14 mental-age months. While I feel sure that this increase in aptitude test score improved their technical trainability and their probable job performance to a corresponding degree, we do not have the data to test the belief.

Many personnel directors feel that "personality" factors are important in selection and in quality of job performance; therefore, they wish to have measures of personality included in the selection battery. The lack of predictive potency of personality tests was called to our attention years ago, both by Ellis and by Kurtz, and most recently by Guion and Gottier (5). We are not willing to believe that "personality" qualities are unimportant or have no bearing on performance either on-the-job or in training. We may not be investigating the right habit-attitude clusters, we may be using our measurement in the wrong place. Could it be that we need to include personality measures of the individual relative to similar estimates of the characteristics of the job climate or organizational climate complex? It is also possible that we are hobbled by our blind use of linear aggregative statistics.

Then there is the "invasion of privacy" issue, which has stirred considerable controversy. I hope the heat generated by this question will produce some light as well. This appears to have developed from the use of personality tests constructed and validated for use in the diagnosis or identification of serious personality and adjustment problems in our industrial realm, and in the use of such tests in personnel actions threatening to the individual rather than in the "mental clinic" treatment climate. Another factor in the situation is that of asking an individual, essentially, "to unwittingly testify against himself." This issue was well put by Dr. Allin Westin at the 1966 meeting of the American Psychological Association: "A vital element of privacy is that the individual be able to choose the time and place for revealing his innermost secrets. Otherwise, he is seared by the hot light of selective forced-exposure under conditions, or at a time, or by individuals, where such exposure could be threatening to the individual."

This does not force abandonment of the use of knowledge of individual differences. It does point to the need of careful validation of the measures used. This does not appear to be of any great issue in rank-and-file employment, in which the tests are more likely to be ability and knowledge rather than "personal indirect." The issue does come more into focus in the selection, transfer, or upgrading of executive personnel, particularly in those situations where paper-and-pencil techniques for measurement of personality and motivation are employed.

Another area in which there is some stirring is in the selection of aptitude or selection tests. Primoff (10), U.S. Civil Service Commission, has developed improved techniques for choosing selection tests in terms of their relationship to carefully identified job elements, duties, or functions. I believe that test selection becomes too stereotyped or perhaps too "off-hand." We frequently fail to do the kind of study needed for effective test selection; i.e., a careful job analysis to find out what the job content really is and from this to infer the aptitudes and even the specific aptitude tests and levels of scores appropriate for good

selection. Only once in my career have I selected a battery of aptitude tests all of which survived field validation. These tests were selected after several days of interviewing first-line supervisors. The jobs involved sales representatives, and the company was not willing to accept the amount of sales as a measure of individual performance, since they believed that a sale was the product of the efforts of many individuals in many parts of the organization. I asked each District Manager to name his best salesman, and then to describe this man, telling me what the man did or did not do, that made him the best salesman: "How did he know that this man was his best salesman?" I tried to keep the discussion on the track of objective, observable behavior, avoiding the traits, cliches, and stereotypes which could not be operationally defined. The same procedure was repeated for a salesman he considered one of his least effective. The interview was continued until the performance of at least four salesmen had been given and no new information was being presented. Out of this I developed a much better understanding of the value system of the company and had a picture of what kinds of observable behavior produced higher performance ratings and what job behaviors characterized those receiving lower performance ratings. As a result I was able to select tests each of which showed significant correlation with performance on the job as valued by the first-line supervisors and by the Vice President for Sales.

Whenever one discusses personnel research, one is bound to run into the question of "the criterion" scores. Dunnette (3) suggests prediction in limited rather than global situations. Kipnis and Glickman (6) and others have found merit in using different predictors for high- and for low-aptitude groups. They also indicated that they made use of both positive and negative selection in such situations. The Dunnette model (4) permits the possibility of the predictor's being differentially useful for predicting the behavior of different subsets of individuals. Further, it shows that similar job behaviors may be predictable by quite different patterns of interaction between groupings of predictors and individuals. At the same level of performance, such predictors can lead to substantially different patterns of job behaviors by different individuals. Finally, the model recognizes the annoying reality that the same or similar job behaviors can, after passing through the situational filter, lead to quite different organizational consequences.

Then, there is increased interest in the relationship of organizational characteristics to job performance. Porter (9) has pointed out the need for combining the studies of organizational effectiveness with the concepts of individual differences. This amounts to saying that we are not taking into consideration enough of the total work performance situation. I haven't seen it tried, but let us consider a situation in which one can obtain a validity of .60 in prediction of on-the-job performance, using characteristics of the individual as the independent variable -- such as aptitude test scores, educational history, work history, biographical information, and so on – characteristics belonging uniquely to the individual. Then it should be possible to obtain a correlation of at least .50, using the characteristics of the worker's work situation, i.e., using only variables outside the skin of the individual worker. If we could use both sets of predictors, we might be able to improve our prediction; certainly we would be better able to understand the nature of work performance.

In constructing criteria we still use, too often by default, ratings of performance. Frequently these are the only measures available. Two kinds of effort to improve this situation are appearing.

There is an increased effort to evaluate on-the-job performance in terms of the accomplishment of specified goals. This is derived from performance appraisal in terms of objectives. Wallace (13) reflects the importance of using the attainment of key performance goals as criteria, i.e., as measures of job performance. Far too frequently these have not been defined nor spelled out for the individual.

The second kind of effort is the increased use of objective, observable employee behavior as criteria. Such observable employee performance must itself reflect value to the employer. In the 1950's, Fryer used one-year retention as a criterion for validating selection tests for employees in the Arabian-American Oil Company. Using this criterion, he screened and validated adaptations of our simpler oral intelligence test questions and problems.

The study reported by Katzell and Merrihue (8) was an attempt to obtain operational answers to the question often asked of managers, "Can you estimate employee morale without giving them questionnaires to fill out?" Their answers were universally, "Yes". They indicated that they used such evidences as absenteeism, disciplinary actions, participation in company benefits, posture of the shoulders, suggestions, grievances, accidents, gleam in the eye, and so on. This study started with the hypothesis that "employee morale" is reflected by whatever is common to the observable, measurable evidences suggested by the managers as their bases for estimating morale. The index, Employee Relations Indicator (ERI), is the largest centroid factor common to the behavioral evidences noted above. While this index was derived for work groups, the most important finding was that in a sample of 40 industrial plants, the index was shown to correlate between .40 and .55 with plant productivity and profit. The components of the index were not themselves direct evidences of specific performance objectives but were interpreted as associated qualities or behavior surrogates.

While the index was used more as a measure of organizational performance, it suggested the merit of more intensive exploration of objective behaviors to reflect the quality of individual on-the-job performance.

In developing criterion measurements of quality of performance of mine foremen, a count of the number of work delays on each shift in each mine covering a one-month period was obtained. The number of work delays showed a correlation of -.40 with average managerial rankings of the quality of performance. Work delays correlated only -.13 with peer rankings.

In the study of the performance of salesmen mentioned above, where the sales book could not be accepted as the evidence of productivity, it was found that entertainment of customers and prospects was deemed important. In this situation, the annual cost for entertainment was found to correlate substantially with managerial rankings of quality of salesmen.

In constructing a criterion for quality of performance of two samples of chemists, rankings of the quality of performance of these chemists were made by their immediate and next immediate superiors. Then it was found that the industrial engineers, in trying to better define the duties of chemists, had quite recently collected ratio-delay observations for these same chemists. The data of the industrial engineering study were reorganized to reflect two categories of percentage information, namely, the percentage of observations in which the chemists were found doing something technical and the complementary percentage in which they were doing something non-technical (i.e., making routine computations, washing laboratory glassware, smoking, etc.). The percentage technical activities showed correlations of .70 and .76 with the average of manager's rankings of quality of performance; and .48 and .73 with average salary increase for the preceding five years.

There have been several studies touching on the question of "style" of the individual, indicating that different kinds of individuals might be able to perform the same job equally well, and that there is no one fixed pattern of aptitudes or style which would produce or be required to produce adequate performance. There may have been somewhat divergent approaches to this matter. Studies of individual differences in style have reflected the concept that there can be differences in the way an individual sees his role and differences in approaches to a job which can yield comparable results. These do not seem to be closely related to studies of the relative performance of democratic group organizations as compared to authoritarian patterns of action.

The organizational or group approach shows that different patterns of group organization and communication do produce different results, particularly in such areas as decision making and in the organizational locus of decision. These studies include those reported by Bass (1) and by Likert. Porter (9) comments in the 1966 *Annual Review of Psychology* that these studies need to be linked to individual differences (and vice versa).

Over the years, evidence has accumulated which shows that the aptitudes for performance on-the-job are related to the ways in which performance is evaluated, i.e., the performance criterion. This points to the need, in constructing criteria, to looking very carefully at the organizational goals and especially at the objective evidences by which these may be operationally reflected.

REFERENCES

1. **Buss, B.M.** Further evidence on the dynamic character of criteria. *Personnel Psychol.*, 1962, 15, 93-97.
2. **Biesheuvel, S.** Personnel selection. *Annu. Rev. Psychol.*, 1965, 16, 295-324.
3. **Dunnette, M.D.** Personnel management. *Annu. Rev. Psychol.*, 1962, 13, 285-314.
4. **Dunnette, M.D.** A modified model for test validation and selection research. *J. appl. Psychol.*, 1963, 47, 317-23.
5. **Guion, R.M. & Gottier, R.F.** Validity of personality measures in personnel selection. *Personnel Psychol.*, 1965, 18, No. 2, 135-164.
6. **Kephis, D. & Glickman, A.S.** The prediction of job performance. *J. appl. Psychol.*, 1962, 46, 50-56.
7. **Little, J.W.** Human response to jet engine noises. *Noise Control*, 1961, 7, 11-13.
8. **Merrihue, W.V. & Katzell, R.A.** ERI - yardstick of employee relations. *Harvard Business Review*, November-December 1955, 33, No. 6, 91-99.
9. **Porter, L.W.** Personnel management. *Annu. Rev. Psychol.*, 1966, 17, 395-422.
10. **Primoff, E.W.** The coefficient approach to jobs and tests. *Personnel Admin.*, 1957, 20, No. 3.
11. **Ruch, F.L.** Fair testing for fair employment. An extension of remarks made at the Fifth Annual Personnel Administration Workshop of the National Association of Independent Insurers. San Francisco, California, January 1966.
12. **Sells, S.B.** Personnel management. *Annu. Rev. Psychol.*, 1964, 15, 399-420.
13. **Wallace, S.R.** Criteria for what? Paper read at Am. Psychol. Assoc., Los Angeles, 1964.

III

RECENT PERSONNEL RESEARCH IN THE BUREAU OF NAVAL PERSONNEL: A REVIEW

Victor Fields
Personnel Research Division
Department of the Navy

Introduction

Although this paper, as the title indicates, is concerned with personnel research conducted by the Bureau of Naval Personnel, this does not imply that only the Bureau of Naval Personnel conducts or supports personnel research in the Navy. The Bureau of Medicine and Surgery has a vigorous in-house program of personnel research which is addressed largely to the problems of adjustive behavior, character and emotional disorders, and behavior under stress and isolation. The Office of Naval Research supports an extensive program of behavioral research under contract. The materiel commands support research concerned with human effectiveness in weapons systems. Nevertheless, because of the limitations of space and time, this paper will deal only with personnel research of the Bureau of Naval Personnel. No attempt will be made to give a complete account of this program. The items discussed in the following pages will serve to illustrate the problem areas of current interest.

Early Origins

In keeping with the spirit of the twenty-fifth anniversary of the Air Force personnel research program, it is proper to point out that the Navy, too, should be celebrating its twenty-fifth anniversary of personnel research. While official Navy testing for selection purposes started about 1924, a Navy research program did not exist; the development of the then current selection test was accomplished with the aid of the U. S. Civil Service Commission in which Dr. O'Rourke was then director of research.

The advent of World War II brought with it a recognition in the Bureau of Naval Personnel of the inadequacy of the existing personnel research and development resources for meeting the changing demands of an expanded Navy already girded for war. As a result, the seed which became the present Personnel Research Division was planted, circa 1941, and an active program of test development was conducted during the war years. The post-war demobilization considerably reduced the size of the organization, but in 1951 personnel research in the Bureau of Naval Personnel received a "shot in the arm" in the form of systematic and continuing support with R&D funds. This funding has been steadily increased over the past fifteen years so that the Bureau of Naval Personnel now possesses an organization of 12 to 15 professionals at the headquarters office in Washington and about 30 in two field laboratories located in Washington, D. C. and San Diego, California. This is a result of increasing attention on the part of the Navy to problems of personnel administration in a military world characterized by increasing complexity in weapons and weapons systems technology.

Selection and Classification Research

Since the aim of selection research is to improve the identification of differential abilities and thereby to optimize the matching of men and jobs, the selection and classification program has been directed toward four major goals: increasing the overall validity of the enlisted classification tests, streamlining test procedures, identifying new dimensions of human behavior that are predictive of military performance, and exploiting automatic data processing techniques in the assignment process.

The Navy Classification Battery

The composition of the Navy enlisted classification tests has changed considerably since World War II. The average intercorrelation of the tests has been reduced, and the overall validity of the tests for assignment to about 70 schools has been raised.

In the most recent study to improve the differential validity of the Navy Classification Battery (2), a Naval Knowledge Test (NKT), a Biographical Information Blank (BIB), and a Navy Activities Preference Blank (NAPB), i.e., an occupational interests inventory, were added to the existing tests for experimental evaluation. The criterion performance was final school grades.

While the advantages of unique sets of predictors and weights to optimize assignment to each of about 70 schools were fully recognized, considerations of economy in processing of test results precluded such an approach. For practical reasons, a limited number of prediction equations based on groupings of similar courses was the best alternative. Groupings were determined by similarity of curricula, similarity of interests, patterns of validities, and results of a factor analysis of predicted grades derived from actual regression equations. The analysis yielded four occupational groupings: Clerical, Electronics, Electro-mechanical, and General Technical.

The major findings of this development and validation study (2) were essentially as follows:

1. The Naval Knowledge Test, expected to measure interest in Navy occupations, did not add any significant variance not already accounted for by other tests in the battery.

2. The Biographical Information Blank showed very little independent validity.

3. The Arithmetic Computation section of the Arithmetic Test tended to depress the validity of the test (and was consequently removed).

4. The Shop Practices Section of the Electronics Technician Selection Test proved to be a highly valid predictor of grades in the mechanical and electrical schools. (Therefore, the Shop Practices Section was lengthened and incorporated in the battery as a separate test.)

5. The Navy Activities Preference Blank, an interest inventory, did not prove to be a useful addition to the battery.

6. The median intercorrelation among composites of the battery was reduced.

7. The new battery yielded an increase of approximately 5 per cent in the number of recruits eligible for assignment without any loss of validity.

The search for new and independent dimensions of behavior to be measured in Navy classification batteries has not been very rewarding, although there appears to be some promise of improving classification by means of tests developed by Kipnis which tap noncognitive abilities (3;4;5;6).

Automatic Test Scoring and Reporting

The automation of test scoring and reporting of results has been under investigation for a long time, but many administrative problems conspired to delay any effective progress in this direction until very recently when automation was given considerable impetus by the emergence of practical techniques for computerized assignment of recruits to schools. At present two optical scanners, installed in the Naval Training Centers at Great Lakes and San Diego, provide the basis for rapid strides toward automating test scoring and reporting of results.

Computerized Assignment

A significant portion of the management resources of the Navy is committed to the procurement and processing of more than 100,000 recruits each year. This aspect of personnel management is so vast and costly an enterprise, involving policy decisions at the highest level of management in the Navy, that even a slight gain in efficiency can result in savings of millions of dollars. The means to achieve such a gain in efficiency is at hand as the result of a project conducted by the Naval Personnel Research Activity, San Diego. This project bears the acronym COMPASS (Computer Assisted Assignment of Recruits).

The central aim of COMPASS is to apply computer technology to the complex problem of matching men and jobs at the school-assignment level. The problem is complex because it involves assigning each week, at each of two training centers, 1,000 or more men of different abilities and qualifications to about 70 different occupational categories on the basis of various combinations of about 30 prerequisites. The optimal distribution of human resources under such circumstances is obviously beyond the capabilities of human beings.

The inability of personnel classifiers to consider simultaneously all men, all jobs, and all prerequisites is evidenced in many comparisons that have been made between manual assignments and computer assignments of the same input population under the same quota requirements and under the same constraints of qualifying prerequisites and personal preferences of recruits. The computer always effects a superior match of men and jobs as measured by more adequate filling of quotas and by a better qualitative mix of assignees. All of which increases the probability of success in school and on the job, the probability of greater job satisfaction, the probability of more effective utilization of school facilities and curricula, and the probability of better meeting our quantitative requirements for skilled personnel.

The problem of optimal assignment of all personnel to all jobs is not a new one, nor is the basic solution a new one. All the military services have been devoting a considerable amount of thought and effort to this problem over the last 25 years (7;8;9;10;11;12;13). The suggested solutions were not quite practicable prior to the present day computer, however. In 1964 the Navy was experimenting with an Army version of the solution and was conducting simulated assignments when the Marine Corps became interested in the problem. As a result of cross-fertilization of ideas among Army, Navy, and Marine Corps, the Navy was able to initiate COMPASS operationally at the U. S. Naval Training Center, San Diego, in November 1965.

COMPASS at San Diego consists essentially of the following four computer programs:

1. PROFILE. This program provides a printout for each recruit which indicates the schools to which he may be assigned, considering both his classification test scores and the restrictions imposed by the recruiting program under which he entered the Navy.

2. SCREEN. This program reviews the assignment recommendations made by the interviewer to insure that each recruit meets the various qualifications of each school for which he is recommended. The program detects errors made in recording and key-punching information obtained during the interview.

3. OPERATE. This program is the actual school assignment procedure. It uses as input each recruit's test scores, the recruiting program under which he entered the Navy, months of obligated service remaining, assignment recommendations made during the classification interview, special qualifications for specific schools (such as security clearance and physical requirements), and school quotas for each of the more than 70 schools to which assignments must be made. The major criterion for assignment to schools is the sum of selection test scores on, usually, two of the classification tests, e.g., Arithmetic plus Mechanical tests. This sum is multiplied by a weight according to the recommendations made in the classification interview. This program then maximizes the sum of weighted scores for all available men, for all schools, within the limitations imposed by quotas, requirements, and special enlistment programs. The program does not insure that each recruit will be assigned to a school for which his selector score is highest; nor the one for which he is most strongly recommended. The solution reached is such, however, that the best overall use is made of the available manpower.

4. GEN-DET. This program identifies, by specific occupational codes for assignment, those recruits who were not sent to schools by Program OPERATE. The program recommends assignment of specific Navy enlisted occupational codes for about 80 per cent of the non-school men, using criteria established by classification personnel. The remaining 20 per cent of fleet-destined men must be assigned enlisted codes by hand when interviewer recommendations and rating entry quotas do not match.

As previously mentioned, COMPASS was introduced operationally at the Naval Training Center, San Diego, in November 1965. Programs PROFILE and SCREEN had already been in operation for several months, and their value in freeing the classification interviewer from routine clerical operations and performing these operations more accurately had been established. Comparisons of COMPASS with hand assignment procedures for the same input population and quotas indicate that COMPASS fills more school seats while maintaining or increasing the level of quality of personnel assigned to schools.

Experimental comparisons were made between manual and computer- assigned procedures on a typical week's input of 905 recruits classified at the San Diego Naval Training Center in April 1964. For this sample of 905, which had a total school quota of 545, final school grades were predicted both for recruits actually assigned to the schools by classifiers and for recruits for whom hypothetical school assignments were made using the COMPASS assignment program. The entire quota of 545 would have been assigned using COMPASS, but only 487 men were assigned by the conventional method. The comparison proved that the average predicted final school grade of men who would have been assigned by COMPASS was about 1½ points higher than that for the hand assignments actually made. This difference in average predicted final school grade would have been even greater in favor of COMPASS had only the top 487 men from the 545 assigned by COMPASS been considered in the comparison.

The current version of COMPASS is an interim system, operational only at the San Diego Naval Training Center. While this operational demonstration was being evaluated it became apparent that a more comprehensive system encompassing both training centers (San Diego and Great Lakes), and controlled at the Bureau of Naval Personnel in Washington, was not only feasible but more desirable. In the event a third training center were added, it could readily be accommodated into the system at little extra cost. The

principal advantages of a centralized system are two-fold: (1) It can consider *all* the available personnel to match *all* the quotas and job assignments simultaneously; (2) it can take into account transportation costs from the training center to the next assignment and can trade off such costs against other relevant considerations. Centralization will therefore provide improved optimization of recruit assignment through a single, efficient solution in any given period.

In view of the plans for a more comprehensive computer-based system of assigning recruits to schools and to the fleet, the original version of COMPASS is now referred to as COMPASS MOD 1. The more comprehensive version is referred to as COMPASS MOD 2.

Development of COMPASS MOD 2 was launched on 30 September 1966 with the expectation that it will be operational by September 1967. Designed into the system will be data transmission capabilities between the Great Lakes Naval Training Center and the Bureau of Naval Personnel and between the San Diego Naval Training Center and the Bureau of Naval Personnel. If and when a third training center is established at Orlando, Florida, transmission capability between that center and the Bureau of Naval Personnel will also be developed.

Retention

In war time, with the nation fully mobilized and the mobility of personnel under control, retention of personnel in the armed services is taken for granted. But under conditions of cold war military operations, the personnel turnover rate is a serious problem. What the optimum rate should be is a question that is confounded by questions of economics. Nevertheless, it is quite certain that in some categories of personnel, the rate is far too high and in others it may not be high enough. Thus, we have a problem of personnel research not faced by our predecessors in World War II. Consequently, a good deal of effort has been devoted to devising techniques by which career proneness may be identified.

Enlisted Retention

About all we know today is that of all the personnel who enter the Navy voluntarily, a small proportion of them ultimately decide on a career in the Navy. When the obligated period of service ends, reenlistees come from those who initially indicated a negative interest in a career as well as from those who indicated a positive interest, although the reenlistments from the positive interest group is about twice those from the negative interest group (14). But developing instruments by which we can identify the career-prone individual with any reasonable degree of certainty still eludes us. It appears that one question concerning the enlistee's interest in reenlistment for a career is just as good as an elaborate questionnaire in predicting career potential (14).

A study currently underway (15) was designed to determine what variables are related to retention and to what degree. The variables include socio-economic, service, and personal factors including attitudes and opinions. The criterion measure was whether or not the subjects in the sample of 5,000 reenlisted. Analysis of the data showed that 59 out of 85 items in the questionnaire were significantly related to reenlistment. Results from this research will be utilized to identify current personnel-related policies and practices which may require review, either to increase their already favorable effects upon retention or to minimize their negative effects. But conclusion as to the predictive nature of these data must await multivariate analyses.

Officer Retention

In connection with the selection of candidates for the NROTC (Regular) selection program, which includes a subsidized 4-year college education, the need to identify applicants with strong Navy career intentions led to development of a self-description blank, the Career Motivation Questionnaire (CARMOT). While this test showed considerable promise of discriminating career from non-career motivated applicants on the basis of its concurrent validity, longitudinal studies demonstrated the inadequacy of this questionnaire when subjected to the more rigorous test of predictive (temporal) validity. That is to say, over a span of eight years, when the career decision criterion became available, the predictive validity of the test was substantially zero. This led us to seek other avenues for measuring career motivation. Currently we are studying the validity of a Navy officer key developed by our San Diego laboratory for the Strong Vocational Interest Blank on the assumption that vocational interest is tantamount to career interest. This key shows considerable promise for identifying occupational interest and choice in a Navy career (16). Whether its validity will hold up over the long run (four years of college and four years of obligated service) is still a question to be decided by longitudinal studies currently underway.

Training Research

The aims of Navy training research have undergone a gradual evolution since the inception of the Navy personnel research program in World War II. In those early years training research was viewed largely as a program of developing achievement tests and evaluating training effectiveness. Today we are concerned with bigger issues. The aim of current training research in the Navy, as reflected in the program of the Navy Training Research Laboratory of the Personnel Research Activity in San Diego, is two-fold: To conduct research and development which will yield measurable improvements in Navy training efficiency and, in the long run, to develop generalizations or principles which will constitute a theory of training technology with wide application and utility in planning, designing, and executing Navy training programs. More specifically, the first aim is to find ways and means by which to:

1. Increase the proficiency, skills, or levels of knowledge of Navy trainees.

2. Reduce training time required to meet specified objectives.

3. Reduce training complexity and prerequisite aptitude levels.

4. Reduce failure rates for a given ability level of the trainee population.

5. Develop techniques which will facilitate shipboard training.

6. Define training objectives and standards in relation to demands of Navy jobs.

The second aim of the training research program is supported by studies designed to:

1. Investigate problems in transfer of training such as transfer from training devices to equipment on the job, from knowledge to performance, and from one equipment to another.

2. Investigate ways in which instruction can be programmed to promote maximum effectiveness in training.

3. Investigate relationships between individual differences in aptitudes and effectiveness of training.

4. Investigate problems of job simulation for training purposes.

5. Develop methods for validating training programs.

6. Investigate relationships between training content and training methods.

7. Develop appropriate measurement methodology for use in training research.

Consistent with these plans an experimental training program in cooperation with the Fleet Anti-Submarine Warfare (ASW) School resulted in significant improvements, without any increase in training time, in the mastery of standard test equipments used in electronics maintenance (17). Studies in programmed instruction demonstrated that both programmed instruction without an instructor and programming the material presented by the instructor were superior to conventional instruction used by the ASW school (18). For long-range purposes we have engaged in a series of classified studies on individual differences in ability of sonar technicians to discriminate, assess, and use cues from auditory and visual displays in target classification. The findings from this research will have a bearing on instructional methods in this domain of training.

Recent studies in Anti-Air Warfare training include an experimental comparison of techniques for teaching computer program flow-chart design to officer trainees in computer programming (19). Using two methods and two lengths of practice on four groups of college students, the only significant variable was found to be quantitative aptitude. It was concluded that selection is more important than methods of training when dealing with computer programmer training of officers.

An important part of any training system is the design of new courses or redesign of existing courses to keep up with changing requirements. In order that this may be accomplished effectively by available instructional personnel, an approved methodology of course design is required. Accordingly, the Navy Training Research Laboratory at San Diego developed a very useful manual for instructors who design or redesign courses in Navy schools (20).

In the field of simulation of warfare environments and conditions, efforts are being directed toward developing problem exercises for team training. The vehicle is a computer-based complex simulating tactical combat direction and advanced electronic warfare.

Experimental Training

The increasing demands for men of high intellectual caliber to fill billets in the electronics maintenance occupations led to a consideration of the feasibility of expanding the potential pool of eligible men by lowering aptitude qualifications and reducing the complexity of training.

Consequently, an experimental course for electronics technicians was initiated at the Navy Training Research Laboratory of the San Diego Personnel Research Activity with the following rationale:

1. The experimental course would be aimed at personnel who possessed aptitudes about one-half standard deviation lower than the minimum for the regular course, who would not be likely to undergo any advanced training, and who would not be likely to advance beyond pay grade E-4.

2. A considerable amount of mathematics and theoretical material often included in electronics instruction would be eliminated without sacrificing the effectiveness of the graduates in performing the assigned duties of the E-4 pay grade.

3. The course would be aimed at teaching the practical factors of electronics maintenance with emphasis on preventive, routine, and corrective maintenance tasks assigned to technicians at pay grades E-3 and E-4.

4. Because the emphasis would be on practical rather than theoretical factors in electronics, and the aptitudes required for admission would be lowered, greater numbers of men could become eligible for electronics training.

Accordingly, five classes of 16 students each, with aptitude scores about one-half standard deviation lower than normally required, were trained under a curriculum of electronics training revised in accordance with the foregoing rationale. The length of the course was 8 weeks less than for the usual 38-week course for electronics technicians. The 80 graduates of this course have since been assigned to destroyers in the Pacific Fleet where they are currently the subjects of followup evaluations. Preliminary information from several ships is very encouraging. The immediate usefulness of the graduates aboard ships has been acknowledged. They appear to be working out well and are performing corrective and preventive maintenance expected at the E-3 and E-4 pay grade. The formal analyses of the followup data on these men will determine the effectiveness of the experimental curriculum, although administrative steps have already been taken to modify the training of men who do not wish to obligate for more than 4 years of service (21).

In the area of electronics maintenance, the Bureau of Naval Personnel has long debated the desirability of giving long and difficult training to men whose service in the Navy will be limited to one enlistment. It has been argued that a 38-week course in electronics maintenance, coupled with subsequent advanced courses of from 6 to 30 weeks, provides an investment in time and money from which the return in the form of journeyman service in only one enlistment is hardly commensurate. For men who do not intend to stay in the Navy more than 4 years, a much briefer course would be more desirable if it could be demonstrated that such a course could produce technicians at an acceptable level of competence. Therefore, an experimental curriculum of 20 weeks' duration was developed. In addition, it was posited that simulation of radar and communication equipments in the laboratory phase of electronics training could be substituted for the real and expensive equipment. For this purpose a readily available, off-the-shelf, and relatively inexpensive analogue computer was employed together with standard panels of dials. The computer generated the desired functions which were displayed on the panels. Mock-ups and life-size photographs or drawings of the front panels and internal circuitry of various pieces of equipment were affixed to the standard panels to provide some realism. The students used appropriate test equipments on appropriate test points of the mock-ups which approximated the test points on the real equipment. Theory and mathematics were minimized in the curriculum that was devised for this experimental training course. The aim of the course was to train the students in practical procedures of preventive and corrective maintenance and fault localization in 20 weeks (22).

Of the 16 students who entered the experimental course, 14 were assigned to destroyers and 11 have been available as subjects in followup evaluations. Three months after their assignment, their supervisors were directed to evaluate them and to evaluate a comparable control group in respect to the need for unusual amounts of on-the-job training and supervision, as well as the adequacy of their preparation for the jobs they were occupying. An analysis of the supervisors' ratings showed that, with one exception, they did not regard the control group, which had received 38 weeks of training and exposure to real equipment, as superior to the experimental group on any of the variables. The exception was a supervisor who pre-judged two experimental subjects as lacking necessary training in electronics theory and thereupon had promptly proceeded to provide that knowledge (24).

About six months later a series of proficiency tests developed by the Electronics Personnel Research Group at the University of Southern California was administered to the same subjects and to a control group. Again the performance scores of the experimental and control groups were not significantly different except on a test heavily loaded with theory which favored the control group (24). Subsequently the experimental subjects and a control group were administered a series of troubleshooting performance tests on actual equipment. While the results of these tests have not yet been analyzed, a perusal of the data conveys the impression that the experimental group again did as well as the control group, although neither group did as well as might be desired.

The results of these experiments seem to indicate that there is considerable promise in shortened training and in the use of relatively inexpensive simulation techniques in the training of electronics technicians.

Performance

Modern conceptions of system effectiveness take into account the human component as well as the hardware components. The needs to determine the "design specifications" of the human component has led to a need for determining and specifying human performance standards related to total system effectiveness. This need is especially apparent when military commanders and planners require information concerning the degree of readiness of their weapons, units, or forces for combat operations.

The need for a method to measure and predict the human contribution in a "total system" context led to a study under contract with Dunlap and Associates (25; 26) entitled Techniques for Establishing Personnel Performance Standards (TEPPS). The major aim has been to develop a model, with general applicability to any system, which will provide the following capabilities:

1. Identify the personnel/equipment functional units comprising a system.

2. State the contribution of each personnel/equipment functional unit to system operation.

3. Determine the performance standards for each personnel/equipment functional unit that are related directly to system effectiveness requirements.

4. Test the effects of deviations from performance standards on system effectiveness.

Figure 1 outlines the data and procedures required by TEPPS. The technique involves analysis of a system into describable and useful units of personnel activities, i.e. performance/equipment functional (PEF) units. All the available data of a sytem are studied in order to define the required system states which are implied or specified by the equipment design (27). Findings are organized and diagrammed in detail to form the Graphic State Sequence Model (GSSM). This serves as the basis for generating a mathematical model of the system, the Mathematical State Sequence Model (MSSM).

By means of the MSSM, standards are derived on the basis of assumptions concerning intrinsic relations among PEF units. Standards in terms of probability of performance and/or time to perform may be derived for each PEF unit and for every potential configuration of the system if there are alternative ways of performing the system's mission.

The basic method of TEPPS has been defined (25), although since its inception, it has undergone a continual process of extension and refinement (28). Plans have been made for evaluation of the model by operating it in the setting of a simulated system. Assuming a successful outcome of this evaluation, the next step will be an evaluation of the technique on an operational system.

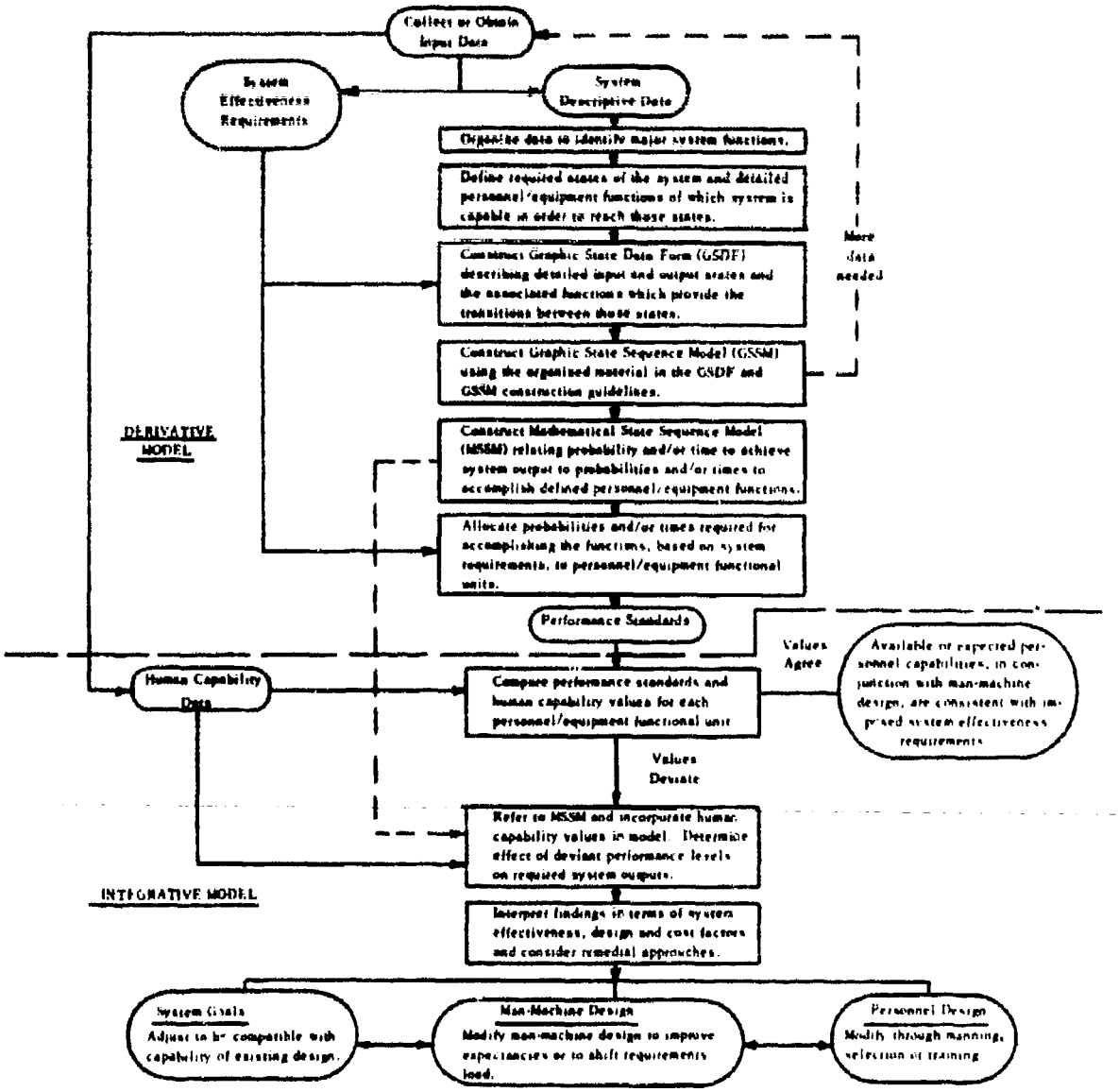

FIGURE 1. OUTLINE OF TEPPS APPLICATION.

This technique for determining system-related performance standards for the human operators offers considerable promise in the following applications:

1. To quantify human performance levels demanded by system effectiveness requirements.

2. To determine the optimum trade-offs between hardware and human performance capabilities.

3. To validate predictor tests and batteries against job performance measures.

4. To provide a basis for defining training objectives.

Electronics Maintenance Research

Related to the foregoing selection, training, and performance research is a program of electronics maintainability research supported by the Office of Naval Research in which the Bureau of Naval Personnel also has a heavy investment. The program is conducted by contract with the Electronics Personnel Research Group (EPRG) of the University of Southern California. One of the interests of the Bureau of Naval Personnel in this program of research is to develop a model by which trade-offs may be effected between equipment design and personnel training. Such a model would be very useful in planning as well as in procurement of weapons systems.*

The Electronics Personnel Research Group (EPRG) maintainability program is a multidisciplinary attack on both personnel and hardware problems which must be solved to improve the maintainability of electronic equipment in the Navy. The program staff includes psychologists, engineers, computer programmers, and electronics technicians. Project teams work on interrelated problems of assessing the proficiency and improving the training of technicians, developing models and other quantitative tools, and improving hardware and software design.

It is axiomatic that the engineering design of electronic equipment also results in a specifiable maintenance workload for that equipment. As a consequence of the way it is designed, electronic equipment has certain corrective maintenance requirements which must be met by technicians performing corrective maintenance tasks. The number and kind of these tasks will depend upon the corrective maintenance requirements built into the equipment, and upon the way the technician-equipment interface is designed, where interface is broadly defined to include hardware, test equipment, and software.

A fundamental assumption of the maintainability research program is that the corrective maintenance workload, represented by the tasks that must be performed to fulfill corrective maintenance requirements, is the starting point, on the one hand, for teasing out hardware variables which can be manipulated to reduce this workload, and, on the other hand, for identifying training and proficiency variables which can be manipulated to improve the performance of it. The research strategy is to have different project teams develop specific tools in their areas of responsibility and to combine these tools into an integrated model.

As the work has progressed, the orientation of the psychology side of the program has become more and more concerned with problems of technician proficiency and training in the present-day Fleet. This is because of the fact that much of the equipment that now exists in the Fleet will continue to be used for 5

---

* The author gratefully acknowledges the assistance of Dr. Joseph Rigney, Principal Investigator of EPRG, in developing the following discussion of electronics maintenance research.

to 10 years. This equipment cannot be redesigned [illegible] Fleet must live with it. Therefore, any improvements in the availability of this equipment brought [illegible] reducing Mean-Time-to-Repair figures must come about through improvements in the effec[illegible]s of the personnel who are performing corrective maintenance on it and in improvements in the [illegible]urces for maintenance management.

Job proficiency tests have been developed for [illegible]resentative equipments currently in the Fleet and have been administered to samples of technicians to assess their corrective maintenance capabilities. The results of these performance tests indicate very strongly that the technicians who actually are responsible for performing the maintenance on equipment aboard ships are weak in certain essential skills, such as the use of test equipment, the ability to recognize whether or not equipment is performing up to standards, and the ability to do fault localization by using front-panel symptoms. These deficiencies are in large part attributable to the fact that, although all technicians go to a basic school (Class A School) and receive training in basic electronics, not enough of them can go to enough different special schools on specific equipment (Class C School) to learn the characteristics of all the different types of equipment that they find aboard their ships.

Any one technician on a destroyer may be responsible for the corrective maintenance of four or five different kinds of equipment and a total of perhaps a dozen sets. The picture is even worse on a ship like a Destroyer Escort. A Class A School graduate simply does not have the opportunity to learn enough about all these different equipments to meet effectively their corrective maintenance requirements.

This situation led to what might be called a cognitive unburdening concept. Using the AN/URC-32 transceiver as a vehicle, a fault locator in the form of a small, plastic, hand-held job-aid with an accompanying manual was developed. This device, called the XFL (for Experimental Fault Locator) was evaluated in two studies. In the first, radio operators, who had almost no electronics training, were taught to do fault localization by using the XFL (31). It was demonstrated that Radiomen could do fault localization with the aid of the XFL with a high degree of success and within reasonable times. For six malfunctions the mean times to solution required by these unskilled personnel varied between 3 minutes for the easiest problem to 17 minutes for the most difficult problem. Even on the most difficult problem, 80 per cent of the Radiomen were able to obtain a correct solution within the one-hour time limit. In a second study, three groups of Electronics Technicians were used (32). One served as a control group. It was given no training of any sort and was not shown the XFL. A second group was given training in checking out front-panel symptoms on the transceiver but was not shown the XFL and did not use it. The third group was given training in checking out front-panel symptoms and was taught how to use the XFL. The group that was taught to use the XFL did significantly better at fault localization, circuit isolation, and component isolation than did the two other groups. The group that was taught to check out the front panel but was not taught how to use the XFL was able to do troubleshooting at about the same speed as the group using the job-aid, but they were not able to solve the harder problems. The control group did very poorly, both in terms of time and the number in the group solving the problems.

The XFL represents an implementation of cognitive unburdening. It is a way of improving the corrective maintenance of equipment now in the Fleet, within the context of the realities which exist; that is, the fact that most of the corrective maintenance that is done is done by Class A School graduates who do not have the breadth of knowledge of the different equipments that they need. A job-aid like the XFL fills that gap by making it unnecessary for them to know the symptom-malfunction relationships at the fault localization level for each of the equipments for which they find themselves responsible. This concept can be implemented in a number of different ways; the XFL device represents only one of these ways. Another is represented in problem-tree charts which portray the entire logic for fault localization on one large sheet.

A considerable amount of work has been done on the question of how training of technicians might be improved. This has taken the form both of analysis of training as a process and of the development of specific recommendations. Examples of the latter are training objectives for corrective maintenance on the AN/URC-32 and on the AN/SPS-40 radar. These training objectives have been developed in extreme detail. They are a specification of the kinds of terminal behaviors that a technician should be capable of as a consequence of Class C School training on this equipment. While training objectives are a first step in the development of a training program, a second step consists of deriving learning objectives from these, and a third step consists of developing methods for accomplishing the training. Development of learning objectives involves identification of the skills and knowledge that must be taught technicians to enable them to meet the training objectives. Although these recommendations will be supplied to the appropriate individuals in the Navy, it is anticipated that they will also be used as a basis for experimental training courses to be developed and administered by the EPRG staff, utilizing recent developments in programmed instruction and computer-aided instruction.

As was pointed out, a major objective of the maintainability research program is to develop the tools necessary for dealing with such questions as how engineering design configurations affect the corrective maintenance workload and the training that must be given a technician who is supposed to perform that corrective maintenance. In effect, what is needed is the identification of the relationships across training, task, and design domains, and the development of tools and models which can be used to quantify their effects.

The first of these tools developed by EPRG was in the area of troubleshooting. Those parts of the corrective maintenance task spectrum which involved troubleshooting have long been known to be very difficult for technicians and responsible for a large part of equipment downtime. A quantitative model of the troubleshooting process was needed in order to evaluate technician troubleshooting proficiency and to reduce troubleshooting to an orderly process that could be related to design and training variables. A model incorporating a Bayesian processor was developed and was implemented by a digital computer program (33). This computer model, called BETS, can troubleshoot electronic equipment if it is provided with what is called the symptom-malfunction matrix for that equipment. This matrix consists of the symptom patterns produced at test points by each possible malfunction in the equipment. The BETS program steps through test points one at a time. At each step, it computes the probabilities for all of a set of defined malfunctions. As it examines the symptoms at each additional test point, it recomputes these probabilities. The process continues until it converges on the solution; that is, one of the probabilities becomes one and the rest become zero. At that point, the malfunction has been located, and the program stops. In looking for the next best test point to go to after each step, the program examines all possible next moves and selects that one which will be most efficient; that is, which will move it the greatest distance toward a solution.

This model has been used for several purposes, among them to score technician troubleshooting behavior. A special type of performance test was developed in which the exact moves and the sequence of moves of each technician could be recorded with IBM cards. The efficiency of these human troubleshooting sequences then could be compared with the efficiency of the BETS sequences for solving the same problems. In two studies using this technique, it was found that Class A School technicians are only about a fourth or a third as efficient as the BETS model.

Other data collected in these studies resulted from the administration of a symptom-malfunction matrix (S-M) test. These tests required the subjects to complete each cell in a symptom- malfunction matrix by inserting a subjective probability. If the subject thought the designated malfunction could have caused a particular symptom at a particular test point, he inserted a high probability in the corresponding cell. If he thought that the malfunction could not have caused a symptom at that test point, he inserted a low probability. If he did not know, he inserted a probability of .50.

These data can be evaluated in various ways. One of the interests was to estimate the reliability of the S-M test. It was found that technicians can give subjective probability estimates of this sort with a relatively high degree of reliability. The stability, or test-retest, reliabilities were around .85. These data also indicate the kinds of symptom-malfunction relationships that are relatively difficult for technicians to understand and the kinds that are relatively easy. An analysis of the errors that technicians made in taking these tests has revealed that AC and DC circuit relationships are difficult for them to understand and that simple resistance relationships are very easy. A factor analysis of these error data has revealed that the major factor is a complexity factor which evidently corresponds to a complexity dimension identified in earlier multidimensional scaling studies conducted by EPRG. Evidently, the primary cause of errors in completing symptom-malfunction matrix tests lies in the number of electronic events going on in the circuit which must be considered before it can be determined that a particular malfunction could have caused a particular symptom.

An interesting result of error analyses done in conjunction with the analysis of performance test results was the identification of the relationship between symptom patterns and malfunction confusions. When technicians are troubleshooting, they often confuse one possible malfunction with another. It was found that malfunctions most often confused with each other have very similar symptom patterns. This would seem to be an obvious result. But the demonstration of this result stability among basic and advanced students leads to an important design principle: Provide front-panel indications and test points that will unambiguously discriminate among all major malfunction possibilities. Furthermore, the identification of this relationship points to some intriguing possibilities for improving training. Error matrices derived from S-M tests can be used to identify circuit relationships which a given population of trainees do not know very well and to guide remedial training.

The BETS program is being used in a number of ways in addition to its use as a criterion measure of troubleshooting proficiency. For example, it has been used to produce optimum fault localization sequences for the AN/URC-32. In general, if a symptom-malfunction matrix is available for a piece of equipment, the program can very quickly determine a set of optimal fault localization sequences for each identifiable fault area in the equipment. It is interesting to note that in the case of the AN/URC-32, the optimal sequences that were generated by the program very closely resembled those which were incorporated in the XFL job-aid discussed above. The chief difference between the two was the fact that in the job-aid it was practically more efficient to group the very easy tests to be done first, simply because they were so easy to do. An example would be a fuse indicator light. Even though looking at a fuse indicator light first might not be the most efficient first step, from the standpoint of a Bayesian processor, it can be done so quickly that in a practical job-aid it should be one of the first steps. Once these quick-and-easy checks were grouped in this way, it was found that the remainder of the sequences incorporated in the job-aid and those identified by the BETS program were almost identical. Thus, the BETS model is a tool with a considerable amount of versatility. It will be integrated with other models as they are developed.

A second tool, developed by the EPRG engineering staff, is a method for synthesizing times from design configurations. This represents an extension and an automation of industrial engineering Methods-Time-Measurement procedures and is called ARMAN (34). ARMAN is another computer program containing on the order of twelve hundred FORTRAN IV instructions. It was developed to serve as a tool for estimating the corrective maintenance time cost of a particular design. These time costs are one measure of the Inherent Corrective Maintenance Workload (ICMW) because they are based upon the generation of task sequences comprising the ICMW, the assignment of time units to the smallest actions in the task sequences, and the synthesis of these units into total times. In other words, the ARMAN program not only synthesizes times, it also generates the action and task patterns that consume the time.

The time units that are used by ARMAN to synthesize task times are the general units that have been developed by the Method-Time- Measurement Association. Their method is recognized and used widely in industrial engineering. Because the units they have developed are perfectly general, the ARMAN procedure is applicable to any equipment whose dimensions and specifications can be defined. This is not the case with those time estimation procedures that are based upon times derived from the observation of maintenance of particular kinds of systems in the field by a particular sample of technicians.

This program is being used to evaluate two different radar configurations, one a hybrid, transistor-vacuum tube model, and the other a hybrid, microminiaturized integrated circuit-discrete component model. Microminiaturization techniques have resulted in a shrinkage in the volume and weight of the radar to a factor of about two or three. The questions that are being asked of ARMAN relate to the effects of these design changes on the corrective maintenance task cycle and on the minimum times which would be required to perform the corrective maintenance task cycles. These minimum times are based on the assumption that no errors are made in the performance and that the performer does not have to stop to read technical manuals or look up other information.

Using ARMAN, it has been found that the microminiaturized design will cut the ICMW time in half. Furthermore, it has been possible to explore the effects of certain design philosophies and certain design variables. For example, current integrated circuit (IC) technology calls for using IC chips in combination with discrete components as the most economical application. When this is done, the question of the time cost of removing and replacing discrete components, which requires unsoldering, cutting leads, and resoldering, is an important one. It was found in the microminiaturized radar that repair times for discrete components were a significant part of the total inherent corrective maintenance workload (ICMW).

It is also possible with the ARMAN program to vary certain design factors hypothetically and to explore their effects upon inherent corrective maintenance times. For example, if modularization is carried down to the level of one IC chip with its associated two or three discretes, so that these can be put on a small printed circuit board and the board can be substituted without soldering and unsoldering leads, a dramatic drop in the inherent corrective maintenance workload time results. The exploration of variables of this sort with the ARMAN program is currently continuing.

A third model that is under development by EPRG is concerned with the treatment of errors in the performance of serial-action tasks, which are characteristic of corrective maintenance tasks. This model, now under development, will contain two parts: a task taxonomy for corrective maintenance tasks and a theory of errors. The objective here is to develop a tool for dealing with the problem of human unreliability in performing corrective maintenance. It is anticipated that a throughgoing analysis of the factors involved will reveal the principal sources of these errors, and also will provide the basis for developing models implemented by computer programs to explore important aspects of error generation, error detection, and error correction in relation to major maintenance interface configurations.

Ultimately, it is anticipated that these various tools can be brought together into an integrated model which can be used to establish trade-offs between equipment design and personnel training and to point to the most feasible ways to improve the maintainability of electronic equipment. For example, the BETS program can be used to identify optimal sequences for fault localization and component isolation; and the ARMAN program can be applied to these sequences to estimate the time costs of performing them. Two theoretically equivalent sequences may have quite different time costs, depending on the kinds of front-panel controls or test equipment that must be manipulated in order to acquire information about the state of the equipment. Part of the work on a theory of errors involves finding ways to estimate the costs of errors, for example, in terms of time lost. Again, it is anticipated that the ARMAN program can be employed to do this in conjunction with models now under development.

Computer Applications in Personnel Planning and Management

The development of modern computer technology has had a powerful impact on development of improved personnel planning and management techniques. Research initiated only a few years ago is now beginning to bear fruit in such functions as job classification, personnel distribution, enlisted advancement planning, fleet-shore rotation planning, and planning of personnel requirements. The COMPASS Project, a computerized technique for classification and allocation of qualified men to school and job assignments has already been mentioned.

The U. S. Naval Personnel Research Activity at San Diego, California has been directing a continuous and concerted effort to develop and promote the use of computer techniques in reducing the burden and increasing the accuracy of personnel planning and control. As a result of this effort, a method for computing optimal petty officer ratios in the enlisted strength of the Navy was developed and demonstrated on the Radarman rating (35). In this method current trends in retention rates, advancements, and petty officer requirements were taken into account. Another development along the same lines was the development of a computerized algebraic model for use in enlisted personnel planning (36). The model simulates the enlisted advancement system and predicts the values of six variables for nine pay grades in the enlisted ratings. More recently (37) a network flow technique has been developed to optimize personnel on board by pay grade. This technique simulates the movement of personnel into and out of a pay grade of a rating. Taking into account the constraints of attrition, advancements and demotions, and personnel at the E1 to E3 pay grades, the model optimizes the allocation of personnel to pay grades in the light of existing requirements. This is done for a 5-year period and considers the personnel presently on board. The adoption and implementation of this model for use in routine planning is being considered. Another model has been developed for computerized distribution, rotation, and assignment of personnel into the fleets. It simulates the distribution process of the Enlisted Personnel Distribution Office, Pacific Fleet (38).

The Navy is also interested in the way computer simulation models can be used to investigate personnel performance considerations in the early development cycles of new weapons systems, especially in respect to man-machine interactions. One such simulation model has been developed under contract with Serendipity Associates (39). The principal feature of interest of this model is its capacity to determine, within the total system, the results of interaction between human performance variables and other system variables such as equipment reliability, spares, and related support equipments. This model is now at the validation stage. If the results of the validation are favorable, the model should provide a valuable means to develop trade-off information concerning human and hardware performance characteristics.

Conclusion and Summary

The foregoing review of recent research of the Bureau of Naval Personnel is by no means a full and comprehensive one. Many studies were omitted from these pages which merit attention. However, the purpose of the author, in selecting the studies that were included herein, was to present them as illustrations of the kinds of problems that have persistently been emerging in recent years.

In the area of selection and classification research the traditional focus on the improvement of differential and predictive validity continues. But there has been a growing interest in automating testing and the entire classification process, including allocation of personnel to billets. As a result, the Navy's Project COMPASS has come to fruition and will be operational on a centralized basis in a year. Retention problems, which is the Navy's term for personnel turnover problems, have received considerable attention. Attempts to develop tests to identify the career-prone person have thus far not been successful. Some promise, however, lies in the potential value of an officer key for the Strong Vocational Interest Blank. In the meantime, however, data are being collected to identify numerous variables that are related to a criterion of retention *vs.* non-retention.

Training is such a vast enterprise in all the armed services that attention to it can hardly be avoided. The Navy has experienced an expansion of its training research program in recent years. In this area the emphasis has been on finding or developing ways to improve training, to make it more efficient in terms of the investment of time and dollars in certain courses of instruction as well as in terms of the terminal product. Hence, a major effort has been directed toward improvement of training in electronics maintenance. Studies have shown that Electronics Technician training curricula can be pared down in some respects without affecting the proficiency of the graduate.

Research on electronics maintainability has been directed toward developing techniques for measuring the performance of electronics personnel. A principle of cognitive unburdening has led to successful development of job-aids for Electronics Technicians. A computer model to troubleshoot electronic equipment has been employed as a criterion against which to evaluate human troubleshooting performance. Another development emerging from research on electronics maintainability is a program by which certain design factors in electronic equipment can be hypothetically varied in order to explore their effects on corrective maintenance workload. These models should ultimately serve to establish trade-offs between equipment design and personnel training.

Although normative performance standards are not too difficult to develop in some instances, criterion-referenced performance standards are quite difficult to obtain. A technique for estimating personnel performance standards in the context of a total system has been developed and is in the process of being evaluated under operational conditions.

Much progress has been made in application of computer technology to simulation of personnel management and man-machine interactions. Such simulations yield large quantities of information rapidly and will serve to facilitate the decisions of manpower management planners.

REFERENCES

1. **Stuit, D.B.** (Ed.). *Personnel research and test development in the Bureau of Naval Personnel.* Princeton, N.J.: Princeton University Press, 1947.
2. **Alf, E.G., Gordon, L.V., Rimland, B., & Swanson, L.** *The development and validation of Form 7 of the Navy Basic Test Battery.* Bureau of Naval Personnel, Technical Bulletin 62-20, 1962.
3. **Kipnis, D. & Glickman, A.S.** *The development of a non-cognitive battery to predict enlisted performance.* Bureau of Naval Personnel, Technical Bulletin 58-9, 1958.
4. **Kipnis, D. & Glickman, A.S.** *Validity of non-cognitive tests at Nuclear Power School (Submarine).* Bureau of Naval Personnel, Technical Bulletin 59-6, 1959.
5. **Kipnis, D. & Glickman, A.S.** *Development of a non-cognitive battery: prediction of radioman performance.* Bureau of Naval Personnel, Technical Bulletin 59-14, 1959.
6. **Kipnis, D. & Glickman, A.S.** *Development of a non-cognitive battery: prediction of performance aboard nuclear powered submarines.* Bureau of Naval Personnel, Technical Bulletin 61-5, 1961.
7. **Brogden, H.E.** An approach to the problem of differential prediction. *Psychometrika,* 1946, 11, 139-154.
8. **Dwyer, A.S.** Solution of the personnel classification problem with the method of optimal regions. *Psychometrika,* 1954, 19, 11-25.

9. **Brogden, H.E.** A simple proof of a classification theorem. *Psychometrika*, 1954, 19, 205-208.
10. **Brogden, H.E.** Least squares estimates and optimal allocation. *Psychometrika*, 20, 249-252.
11. **Dwyer, A.S.** The detailed method of optimal regions. *Psychometrika*, 1957, 22, 43-52.
12. **Holdrege, F.E.** *OSCAR: optimal selection, classification, and assignment of recruits.* USAF Personnel Research Laboratory, Technical Document Report 62-13, 1962.
13. **Boldt, R.F.** *Development of an optimal allocation.* Army Personnel Research Office, Technical Research Report 1135, 1964.
14. **Steinemann, J.H.** *The prediction of reenlistment among Class A School trained men.* Bureau of Naval Personnel, Technical Bulletin 63-11, 1963.
15. **Nickey, B.B. & May, Jr., R.V.** *Analysis of Navy enlisted personnel retention variables.* U.S. Naval Personnel Research Activity, San Diego, Research Report SRR 66-1, 1965.
16. **Abrahams, N.M., Neumann, Idell, Rimland, B., & Githens, W.H.** *The Strong Vocational Interest Blank keyed to predict NROTC officer retention.* U.S. Naval Personnel Research Activity, San Diego, Research Report SRR 67-4, 1966.
17. **Pickering, E.J., & Abrams, A.J.** *A program for improving the training of Sonarmen in the use of electronic test equipment: III. The evaluation of an experimental basic test equipment training program.* Bureau of Naval Personnel, Technical Bulletin 62-8, 1962.
18. **Abrams, Alvin J.** *Experimental evaluation of programmed materials on the multimeter AN/PSM-4.* U.S. Naval Personnel Research Activity, San Diego, Technical Bulletin STB 65-9, 1965.
19. **Meyer, J.K.** *An experimental comparison of instructional techniques for use in teaching computer program flow chart design.* U.S. Naval Personnel Research Activity, San Diego, Technical Bulletin 65-10, 1965.
20. **Navy Training Research Laboratory.** *Course design manual for job training courses.* (Preliminary ed.) U.S. Naval Personnel Research Activity, San Diego, Research Report 66-17, 1966.
21. **Pickering, E.J. & Anderson, A.V.** *A performance-oriented electronics technician training program: I. Course development and implementation.* U.S. Naval Personnel Research Activity, San Diego, Technical Bulletin 62-2, 1966.
22. **Washington Engineering Services Co.** *Design, development and conduct of an experimental basic training course for Navy electronics technicians.* Technical Report W-1480, Kensington, Md., 1965.
23. **Personnel Research Division, Bureau of Naval Personnel.** *Evaluation of graduates of a 20-week experimental ET course.* Unpublished report. Washington, D.C., 1966.
24. **Rigney, J.W. & Fromer, R.** *An evaluation of the WESCO experimental basic training course for Naval Electronics Technicians: I. Phase one results.* EPRG, University of Southern California, Memorandum Report 2, 1966.
25. **Dunlap and Associates, Western Division.** *A method for deriving job standards from system effectiveness criteria: I. Method development. Santa Monica, Calif. 1964.*
26. **Dunlap and Associates, Western Division.** *Development of a technique for establishing personnel performance standards (TEPPS): Phase II Final Report.* Santa Monica, Calif., 1966.

27. **Mitchell, M.B., Smith, R.L., & Westland, R.A.** *Technique for establishing personnel performance standards (TEPPS): Procedural Guide* (Preliminary Ed.), Dunlap and Associates, Santa Monica, Calif., 1966.

28. **Mitchell, M.B., Smith, R.L., & Verdi, A.P.** *Development of a technique for establishing personnel performance standards (TEPPS): Phase III Final Report.* Dunlap and Associates, Santa Monica, Calif., 1966.

29. **Blanchard, R.E., Mitchell, M.B., & Smith, R.L.** *Likelihood of accomplishment scale for a sample of man-machine activities.* Dunlap and Associates, Santa Monica, Calif., 1966.

30. **Verdi, A.P. & Smith, R.L.** *Personnel performance data store for Navy systems: Preliminary program concept.* Dunlap and Associates, Santa Monica, Calif., 1966.

31. **Rigney, J.W., Fromer, R., Langston, E.T., & Adams, H.C.** *Evaluation of an experimental fault location device: I. Fault location by radio operators.* EPRG, University of Southern California, Technical Report No. 43, 1965.

32. **Rigney, J.W., Fromer, R., Langston, E.T., & Adams, H.C.** *Evaluation of an experimental fault location device: II. Fault location and isolation by experienced Electronics Technicians.* EPRG, University of Southern California, Technical Report No. 44, 1965.

33. **Rigney, J.W., Cremer, R.H., Towne, D.M., Bond, Jr., N.A., & Mason, A.K.** *Measurement and prediction of cognitive tradings in corrective maintenance tasks: I. A Bayesian approach.* EPRG, University of Southern California, Technical Report No. 46, 1966.

34. **Rigney, J.W., Fromer, R., Langston, E.T., & Adams, H.C.** *ARMAN: A computer program for generating inherent corrective maintenance workload times: I. Description, evaluation, and application.* EPRG, University of Southern California, Technical Report No. 45, 1966.

35. **Conner R.D. & May, Jr., R.V.** *A method for developing optimal petty officer ratios.* U.S. Naval Personnel Research Activity, San Diego, Research Report 65-3, 1965.

36. **Conner, R.D. & May, Jr., R.V.** *Applications of a computerized model in enlisted personnel planning.* U.S. Naval Personnel Research Activity, San Diego, Research Report 66-3, 1965.

37. **Hayter, D.F. & Conner, R.D.** *A network flow technique for optimizing personnel on board by pay grade.* U.S. Naval Personnel Research Activity, San Diego, Research Report 66-12, 1966.

38. **Thorpe, R.P. & Conner, R.D.** *A computerized model of the fleet personnel distribution system.* U.S. Naval Personnel Research Activity, San Diego, Research Report 66-13, 1966.

39. **Spencer, G.J., Kolkowitz, H., & Dibelka, G.S.** *The COGAG simulation model (COSIMO) pilot study.* Bureau of Naval Personnel, Technical Bulletin 67-3, 1966.

BLANK PAGE

IV

UNITED STATES ARMY PERSONNEL RESEARCH IN JOB SYSTEMS SETTINGS

J. E. Uhlaner and A. J. Drucker
United States Army Personnel Research Office
Department of the Army

Historically, the personnel research activities of the U. S. Army Personnel Research Office have been largely the activities of industrial psychology. We came by this trend naturally since our pioneers were essentially industrial psychologists, steeped in the tradition of the psychology of jobs - the psychology of the world of work. The methodological mainstay of industrial psychology, psychological testing and test development, became APRO's mainstay (and in many ways still is) to predict success in training, success in total job performance, or success in slices of jobs, tasks, or job functions. As a case in point, APRO has worked intensively on general mental ability testing for screening enlisted personnel, first for the Army alone and from 1950 to the present day, as Executive Agent for the Department of Defense. Similarly, APRO from World War II to the present has been heavily engaged both in the methodology and substance of differential classification. Today, under a directive from the Assistant Secretary of Defense/Manpower (Secretary Morris), APRO is spearheading a joint-service effort to develop a common aptitude test battery to be utilized by all services for screening and classification purposes in the high school testing program.

Perhaps because of the great emphasis the Army has placed on individual combat performance and individual combat officer performance, APRO has for a number of decades also emphasized noncognitive research. To review the dozens of projects in the area would go well beyond the scope of this paper. Most significant perhaps in the early period was a heavy concern with the development of possible predictive measures of successful combat performance from the point of view of the man's will to fight as a vital adjunct to cognitive capabilities required in such situations.

The finding that combat performance could be predicted through noncognitive measures was turned into operational use. The Classification Inventory was introduced into the Army Classification Battery in 1958, as well as into the Army Aptitude Area System (13). Combat performance itself was measured largely by means of ratings, the promise of "objective" criteria being fruitless. We capitalized upon years of research to establish methods for obtaining reliable and valid ratings appropriate for use as job criteria. For example, in repeated studies, we found that a large general factor dominates the rating even where concerted attempts have been made to measure different aspects of job performance by using a number of specific rating scales. At least once we tried what might be called contextual suppressors (12). In predicting combat performance we tried once again to delineate different aspects of combat proficiency but had to settle on how good a man's overall performance in combat was (8). We had also found very little difference in validity between ratings by raters classed as "hard" and those called "easy" raters and stopped worrying about the rater difference (1; 2). The intelligence of the rater had been found to make little difference, except possibly, where raters were in the lower portion of the distribution on mental ability (3). We

accumulated empirical evidence that raters agree more in their evaluations of job success if they have had more opportunity to observe the individual performing on the job; e.g., peer ratings have superior validity over cadre ratings (4). Fellow trainees or fellow workers on the job are generally in a better position to observe performance of another. Perhaps this is an obvious finding, but in the basic training situation, even a period as short as three or five weeks is sufficient to enable the rater to make the gross judgment required. Equally obvious, perhaps, is the finding that the rating based on the judgments of more than one rater is better than the single rating. This is true to the tune of a rating validity of .53 for single ratings compared to the average of ten ratings with a validity coefficient of .84 (5). Little improvement can be expected beyond ten. How many actually to combine is a matter of administrative convenience. These findings we capitalized upon quite readily.

In more recent years it became apparent that for predicting differential officer success, the kind of philosophy which we had used successfully in differential enlisted classification could be readily translated into action in the Officer Prediction Research Program. The study to differentially predict officer success in combat, administrative, and technical assignments is approaching completion (14).

We said earlier that the personnel research psychology effort in APRO has been based in industrial psychology. To characterize our current research effort best is to say that during the last decade we have moved even more vigorously into the psychology of jobs and the world of work while taking full advantage of methodological developments such as operations research (including cost effectiveness concepts); we have placed emphasis on mission output; we have taken greater and greater advantage of sophisticated computer capabilities and its software; and we have added systems analysis to our stock of techniques.

What has all this meant? For us at APRO it has led to the development of a research program designed to improve the total mission output in a work setting, be it a simple job, a slice of a job, jobs in the aggregate, or a group of jobs interrelated in conjunction with sophisticated equipment. APRO's current effort is characterized by a concern with each of these areas (Table 1). However, there is much less preoccupation with finding the best man for each job. With greater and greater shortages of skilled manpower, whether military or civilian, and in fact with greater shortages of any manpower, selection must often take a hind seat. We have recognized that industrial and personnel research psychology have, after all, classically concentrated not only upon selection and training but upon work methods and interface problems with equipment and the man in his total work environment.

With the methodologies and computer equipment available, we could begin to ask more profitable research questions, taking into account real world variables. First, how can the military effectiveness in a systems setting be improved? Second, considering more directly the military user's problem, how can we improve the military mission effectiveness, assuming that we can realistically measure this effectiveness? In fact, much of recent effort has been to borrow psychometric theory for criterion development and apply it to systems output measurement (11). Once that is accomplished, additional questions can be asked. If selection and training are held constant (Table 2) and work methods varied, what effect is there on the measured output? Using a measured output of performance of accuracy and completeness in an information system, we examined a variety of Work Methods for image interpreter teams (Table 3):

Work Method I, in which image interpreters worked and reported independently and in which their performance was scored for right identifications, wrong identifications and omits for teams of two.

Work Method II, same as Work Method I but scored only for responses on which the two interpreters agreed.

Work Method III, wherein interpreters worked in teams and discussed their identifications freely before recording them.

Table 1. **APRO Research in the World of Work**

Job Orientation	Research Approach
Slices of Jobs	Human Performance Experimentation
Group of Jobs in Conjunction with Sophisticated Equipment	Manned Systems Research
Jobs in Aggregate	Selection Research

Table 2. **Improving Military Effectiveness in a Systems Setting**

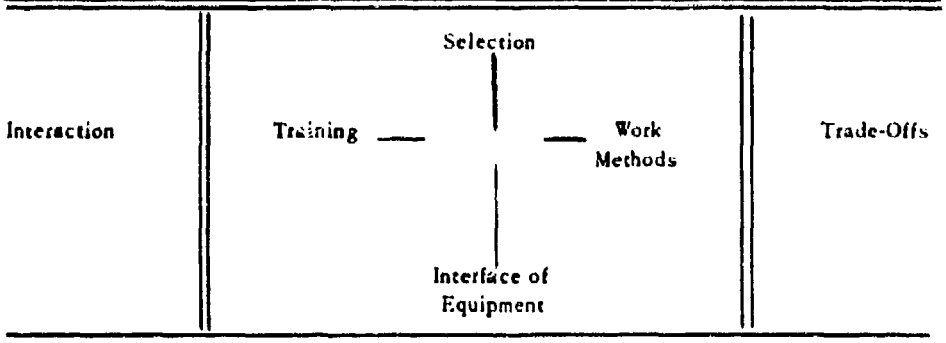

Table 3. **Total Mean Performance Scores for Image Interpreter Teams Working under Different Methods and for Individuals Working Alone***

Team Methods	Right	Wrong	% Accuracy
I Independent: All responses (N = 15 teams)	15.1	112.9	14
II Independent: Agreed responses (N = 15 teams)	3.9	7.1	33
III Cooperative (N = 30 teams)	10.1	40.6	22
IV Individuals Alone (N = individuals)	8.1	50.8	16

* Mean Right, Wrong, and Accuracy scores between team methods are significantly different ($P<0.01$).

Method I gave the largest number of correct identifications but also the largest number of wrong identifications. Method II gave a much better accuracy record but had many fewer identifications of either kind. Method II had some of the advantages of both I and II. Each method was thus judged to have potential usefulness to the commander, depending upon his needs: Work Method I for maximum identifications (however, including many wrongs), Work Method II for maximum accuracy (however, fewer identifications), and Work Method III for a reasonable compromise (6).

If we provide sophisticated equipment such as stereo to the image interpreter at a known cost (which might be considerable), is stereo worth its dollar cost in terms of its accuracy and completeness gains, if in fact there are any gains? What would be the effect on the measured output? In short, the current thinking in APRO is to approximate the world of work at least in part in a systems setting where interaction effects can be systematically examined and trade-offs evaluated from a cost- effectiveness point of view.

Let us call it the criterion concept in new dress. If we are going to put more money and time in training for a given level of performance, how much poorer quality of personnel can we utilize? It seems to us crucial that a beginning be made to attempt research where a mix between different kinds of training and varying qualities of people available are evaluated and the different mixes are studied in terms of payoff, cost matrices, and trade-off considerations. The military might specify its objectives in terms of time and cost of training. Eight weeks are now required for basic training. Perhaps four or twelve would be more useful. Perhaps we might use a track system and achieve better results in the same length of time. Computerization and simulation techniques at least make possible experimentation with these potential advances. Of course, many of these projects may involve getting several organizations together.

By way of refocus, the research program of APRO falls at this time under three major categories, with efforts under each as indicated in Table 4.

Under *Human Performance Experimentation* we seek principles applicable to many Army systems. General behavioral functions common to many Army systems to enhance man's functioning in those systems are studied. These activities have been conducted with regard to such behavioral functions or slices of jobs as vigilance required of monitors in many systems, ability to communicate more efficiently via radio-telephonic means in combat operations in a modern tactical environment, and other functions now under exploration such as ability to perform efficiently in night operations. To do the research, we obtain parameters and variables from the military systems under study. We simulate in the laboratory, using special program equipment for data input and data reduction, and check out the experiments in the field. Typical independent variables used in vigilance research are shown in Table 5.

Typical of the *Selection Research* effort, where we study jobs in the aggregate, is the attempt to determine how to make optimum use of enlisted personnel of lower mental standards. A host of studies in the military services over several decades has shown the utility of AFQT scores in measuring the potential value of young men for performance in military jobs. As we go down the AFQT scale, however, we are harder put to find ways of training the men, on the one hand, in a reasonable period of time, and on the other hand, to a level of performance where they can function without close supervision. Cumulated research experience dictates we must hold to a quality level with a reasonable success probability.

The use of manpower models in personnel resource allocation research is another example of what needs to be done about jobs in the aggregate. In a series of studies conducted at the U. S. Army Personnel Research Office, a sequence of computerized simulations of personnel systems has been developed. In such models sequences of normally distributed random numbers are generated and then transformed so they have the statistical characteristics of scores on the variables represented in the model for the input population. Particular operations are then applied to these simulated personnel samples, the optimal allocation is performed, and the criterion variable is calculated.

ble 4. **Major Categories of Research at APRO**

Human Performance Exp	ntation	Manned Systems Research	Selection Research
Monitor Performance (	gilance)	Image Interpretation	Screening and Differential Classification
Combat Communications		Information Assimilation	Qualitative and Quantitative Requirements
Night Operations (under exploration)		Decision Making	Officer Leadership
			Combat Selection and Special Warfare
			Manpower Information Systems

Table 5. **USAPRO Vigilance Research**
(23 Variables for Vigilance Research)

Specific Signal Characteristics

- Spatial location of display
- Background stimulation
- Programmed artificial signals
- Differential allocation to sense modalities

Immediate Task Factors

- Length of vigil
- Supervisory techniques
- Physical restrictions
- Responding proce dures

General Environmental Factors

- Ambient noise level
- Background lighting
- Temperature and humidity

Personal Factors--Changes

- Fatigue
- Muscle tension
- Search patterns

Personal Factors--Individual Differences

- Personality
- Personal history
- Perceptual acuity
- Self-report (within session)
- Time perception
- Activity levels
- Sex of operator
- Job motivation
- Stability and generality of vigilance

Example: In setting up the Aptitude Area classification system in 1949, certain operational expedients were introduced to facilitate hand computation of performance estimates, the compiling of data, and the assigning of men. The Aptitude Area score was one such simplification, even though least-squares regression estimates of performance in various Military Occupational Specialties (MOS) based on the 11 Army Classification Battery (ACB) tests would provide the most efficient basis for personnel assignment. Each Aptitude Area was based upon only two ACB variables and was used in lieu of, or as approximations of, predicted performances for specific MOS categories or occupational areas. The same Aptitude Area score is used for all MOS within each of the occupational categories. A second operational expedient was use of grouped scores from 0-9 instead of the full range of Army standard scores of 40-160. How much information was lost by these expedients? Two major simulations were conducted to estimate the amount of gain in overall performance that would result from substituting full regression estimates based on a full range of scores.

Optimal allocation was carried out for a number of samples using both the Aptitude Area scores and the least-squares regression estimates of performance. Realistic quota restrictions were imposed on the allocation. The procedures were then evaluated by calculating the allocation averages, i.e., mean expected performance for personnel allocated by the two procedures. Average performance under optimal allocation with Aptitude Area composites was approximately 3 to 3½ performance units higher than expected under random allocation. Thus, the average gain over random assignment by optimal allocation would be twice as much by using regression estimates of performance as it is by use of the current Aptitude Area scores. The gain in the percentage of men of superior performance assigned by the use of full regression as compared with the two-variable composite is even more impressive. On the basis of this study and similar studies now in progress, recommendations have been made concerning implementation of certain regression estimates of performance for allocation procedures (9).

In the second study, the effect of score groupings on value of the objective function was studied. A significant loss occurred in the allocation average using the one-digit code, although the loss is comparatively small. Nevertheless, we would recommend use of the full range of performance estimates in the context of a newly designed automated system in which the full range could be used with no or little additional cost.

Both of these findings, of course, argue in favor of the Army's automated personnel assignment system. Considerable gain would be expected by making the operational changes indicated in the context of the computerized system now being designed and implemented. Such change would be impracticable within a hand allocation system (10).

Under *Manned Systems Research* the emphasis is on enhancing systems output as relates to man's role in small Army systems; this is one of our most exciting areas. It is truly here that the challenge of conducting integrated human factors research is finally met.

Certain aspects of our Surveillance Project are illustrative. Our effort in this project concentrates on image interpretation of surveillance systems and on information assimilation and decision-making of command information processing systems. Systems output is measured in such terms as accuracy, completeness, and time for decision. Here, particularly, we combine field and laboratory effort and use computer simulation extensively, including real-time. As we said earlier, manned systems research is characterized by an integrated research approach which might involve all the human factors aspects in the man-machine interface. Obviously such an effort is sizeable, and we make use of a dozen or more contract groups each year. However, since the chief purpose of the APRO in-house organization is to conduct research, the policy is to maintain a ratio of at least 50 per cent of scientist time for doing in-house research. We are encouraged by the potential gains to be made for the military, especially the equipment and systems designers; these gains can be couched in terms they can best appreciate. Here is one example we are particularly excited about.

Fundamental to much of the systems output, and particularly man's performance within those systems, is the total complex of better, more accurate, more timely decision-making. Based upon the greater capability of measuring systems output now at hand, this approach brings together psychometric theory, probability theory, and man-machine interface knowledge; the approach uses the computer as a major partner and considers cost in a total value (not solely economic) sense.

The research in question - a major program at APRO - deals with a rational means of improving decision-making. The feasibility and usefulness of using expected cost estimates as a means of controlling the information produced by an image interpretation facility is being investigated. This approach employs realistic systems concepts and means for implementing an expected cost matrix. One major requirement is

to obtain computer-calculated estimates of the probability that a given identification or a given observation is in fact correct. The second requirement is to develop error-cost matrices containing value indices for each type of response as it relates to the true state for particular military situations.

The decision maker's judgments and experiences must first be made explicit in the form of probabilities and value indices in the construction of a decision problem. When a given problem is formulated, the decision model can specify a course of action which is best according to well delineated and measurable criteria.

Table 6 illustrates the general case of how the decision maker, although not knowing the true conditions prevailing at the time of his decision, can nevertheless choose among response alternates by taking into consideration both the cost of responding inappropriately, given various possible true conditions, and the probability of incurring the costs. When the costs are multiplied by the probabilities, the resultant product is the expected cost, i.e., the loss which could be expected from following each response alternative. The decision maker need only choose the response alternative with the least expected cost. In Table 6, choosing Response D would minimize the decision maker's risks.

Table 6. **Deciding Among Alternatives Using Expected Costs***

Response Alternative	Cost of Each Alternative	X	Probability of Incurring Cost	=	Expected Cost of Response
A	50		.60		30
B	80		.30		24
C	25		.80		20
D	30		.40		12*
E	50		.30		15
.	.		.		.
.	.		.		.

* Decision Rule: Choose alternative with least expected cost.

The purpose of a payoff matrix is to provide a set of numbers for alternative decision possibilities which can serve as a basis for rational choice among these alternatives. For example, within the intelligence area, the error-cost matrix could serve the military commander for making decisions within the tactical or strategic setting. A means of first determining information needs and then of determining the cost of failing to provide the needed information for a particular military situation was evaluated. As you well know, military significance will be affected by factors such as attack, defense, trafficability, and the nature of enemy forces. Echelon differences also impose different requirements.

In essence, in the image interpretation system, the error-cost matrix is a table of the costs of misinformation. In the operational situation, basic cost information would be determined by the military decision maker. Since different military situations are expected to impose different information requirements, we developed a means for associating a particular situation with the appropriate error-cost matrix. This was done by identifying situation families, developing a priori matrices for them, and providing field personnel with a ready means for classifying the particular situation of concern into one of these situation families.

Table 7 contains an abbreviated sample matrix. The highest cost in the matrix is the error of omitting a missile; the lowest cost is the imprecision error of calling a truck a vehicle. The other costs in the matrix are intermediate in value. There is, of course, no cost for a correct identification.

Table 7. Use of Expected Costs to Control System Responses
(Maximum Expected Cost Set at 2.0 by G-2)

Example 1

Responses	TK	TR	M	NO	Prob. Vector	Expected Cost
Tank	0	2	5	4	.50	1.9*
Truck	3	0	7	2	.10	2.8
Missile	4	5	0	6	.10	3.7
Nothing	5	3	9	0	.30	3.7
Vehicle	2	1	6	2		2.3
Weapon	3	3	4	4		3.4

* Decision: Report Tank

Let us suppose that a given object on a surveillance photograph has been identified by an image interpreter as a tank. The computer tells him there is only a 50 per cent probability that the given object is really a tank. To determine the probabilities, the computer takes into account such variables as image quality parameters, the established ability level of the interpreter, the difficulty of the target, and the level of confidence the interpreter has in his initial interpretation. For the other targets, let us suppose that there is a 10 per cent probability that it is really a missile, and a 30 per cent probability that it is really not a target at all. Then, multiplying the cost matrix by the associated probabilities yields the expected costs of the reporting alternatives. In this example, the lowest expected cost was for the response Tank. Therefore, the object in question would be reported as a tank even though there is only a 50 per cent probability that it actually is a tank (7).

The decision-making officer to whom the report of the image interpretation facility is being sent can control the accuracy of the information produced by the image interpretation system by setting a quality standard for his acceptance of the information. In this case, this standard is a single number representing the maximum cost at 2.0. Only responses with an expected cost at or below this maximum cost would be reported. In the example just shown, the Tank response would just qualify.

However, if the probability of the object being a tank were slightly lower, the expected cost of making this response would rise. None of the expected costs would be below the maximum acceptable cost. The image interpretation system could not report the object one way or another. The system would then seek to improve the quality of the information through a number of means such as asking other interpreters to identify the object in question, obtaining better surveillance cover, or seeking corroborative intelligence information. Any or all of these steps will change the values in the probability vector and allow one response or the other to achieve an acceptable expected cost. The decision maker, by raising or lowering the maximum cost, can thus control the amount and quality of the intelligence produced by the interpretation system. Setting a high standard will tend to allow few identifications but will yield highly accurate intelligence. Setting a low standard will allow a large amount of information, but the average accuracy of the information will be low.

Efforts such as these, in our opinion, will truly be a vehicle for making the research end products for manned systems research most useful for military planners, military decision makers, military designers–the persons involved in interfacing man with sophisticated equipment, including computers, for the purpose of making rational, rapid, and accurate decisions.

We have attempted in this paper to present a view of the current APRO program as one emphasizing the interdependent aspects of military human factors problems and as one interrelating these aspects in broad research approaches. From the research situation where man was the focal point, he has come to be a link in the man-machine-computer system, and he is so regarded by the research scientist in this setting. Recognizing that to management the salient concern is effective system output, the researcher no longer asks simply: Is man A better than man B? He asks: For a given cost, do the end products of the developed system meet the military man's requirements? The greater complexity of such a research approach, we believe, will more fully meet the military user's needs.

REFERENCES

1. **Browning, R.C. et al.** *A study of officer rating methodology–IX. Validity of ratings by hard and easy raters.* Technical Research Report 908, U.S. Army Personnel Research Office, April 1952.
2. **Browning, R.C. et al.** *A comparison of the validity of officer ratings rendered by hard and easy raters.* Technical Research Report 971, U.S.Army Personnel Research Office, August 1952.
3. **Chesler, D.J. et al.** *A study of ratings from raters with aptitude area I scores below 90.* Technical Research Report 933, U.S. Army Personnel Research Office, April 1952.
4. **Kaplan, H. & Willemin, L.P.** Effect of shorter acquaintance, or of adding cadre raters on the validity of early buddy ratings of combat aptitude. Paper read at Am. Psychol. Assoc., September 1956.
5. **Karcher, E.K., Jr., Winer, B.J., Falk, Gloria H., & Huggerty, Helen R.** *A study of officer rating methodology–V. Validity and reliability of ratings by single raters and multiple raters.* Technical Research Report 904, U.S. Army Personnel Research office, April 1952.
6. **Sadacca, R., Martinek, H., & Schwartz, A.I.** *Image Interpretation task–Status report, 30 June 1962.* Technical Research Report 1129 U.S. Army Personnel Research Office, October 1962.
7. **Sadacca, R.** Information processing in advanced computerized surveillance systems. *Proceedings of the 12th Army Human Factors Research and Development Conference,* October 1966.
8. **Severin, D.G. et al.** *Measures of combat performance in Korea– 1. Criterion measures for enlisted men.* Technical Research Report 938, U.S. Army Personnel Research Office, April 1952.
9. **Sorenson, R.C.** *Optimal allocation of enlisted men–Full regression equations vs. aptitude area scores.* Technical Research Note 163, U.S. Army Personnel Research Office, November 1965.
10. **Sorenson, R.C.** *Amount of assignment information and expected performance of military personnel.* Technical Research Report, U.S. Army Personnel Research Office. (In Press)

11. **Uhlaner, J.E. & Drucker, A.J.** Criteria for human performance research. *Human Factors,* June 1964, 265-278.

12. **Uhlaner, J.E., Goldstein, L.G., & Van Steenberg, N.J.** *Development of criteria of safe motor-vehicle operation.* Highway Research Board Bulletin 60, 1952.

13. **Willemin, L.P. & Karcher, E.K.** *Development of combat aptitude areas.* Technical Research Report 1110, U.S. Army Personnel Research Office, January 1958.

14. **Willemin, L.P.** *Prediction of officer performance.* Technical Research Report 1134, U.S.Army Personnel Research Office, March 1964.

V

RECENT PERSONNEL RESEARCH IN THE UNITED STATES AIR FORCE: A REVIEW

James H. Ritter
Personnel Research Laboratory
Department of the Air Force

Former Vice President Alben Barkley once told a story about a United States Senator whose young friend suddenly announced his support for the Senator's opposition. The Senator, quite perturbed about this decision, called on his friend and said, "John, I don't understand your recent action. Do you remember when you graduated from high school and had no money to go to college? You came to me and I gave you the funds for your education. When you graduated from college, I helped you through law school; after law school, I got you your first job with a law firm. Later, when you decided to run for office, I helped you by guiding your campaign, thereby enabling you to get support from friends of mine, and I even financed it in part. Last year I interceded for you with the Governor and had you appointed to your present job. Now all of a sudden you announce that you are supporting the opposition for my office as Senator. I do not understand, John, why?" Undismayed, John retorted to the Senator, "It's true that you did help me through college, you helped me through law school, you helped me in my first job, you helped me on my first political campaign, and last year you helped me obtain my present job through appointment from the Governor–but, Senator, What have you done for me lately?"

This morning I shall talk generally about what personnel research we have done for the Air Force lately. Sometimes in the research field, we see the results of our research on a long-range basis quite well. After the research has been finished, however, the applications are not always immediately evident, and we face the continued pressure of "what have you done for me lately?" I could cite for you a long list of the research work that the Personnel Research Laboratory has been doing. Instead, I intend to pick out representative examples of accomplishments. For those of you who would like a more complete list, I refer you to a Personnel Research Laboratory publication PRL-TR-65-23 put out in December 1965 entitled *Abstracts of Personnel Research Reports: VI - 1954-1965.*"*

Apparatus Tests for Selection

An example of recent "hardware" research in the area of personnel is in the field of selection. During World War II and for approximately ten years thereafter selection devices for pilots, and to some extent navigators and bombardiers, depended not only on paper-and-pencil tests but also on apparatus tests. In 1952 the apparatus tests were dropped because, although they added some validity to the paper-and-pencil

* This document may be obtained by request from the USAF Personnel Research Laboratory (PROI), Lackland AFB, Texas 78236. All studies described in this paper are referenced in the "Abstracts."

tests, the cost factors associated with individual or small group testing, plus the cost of the hardware itself, were too expensive in relation to the gain. It occurred to some of our psychologists in the Personnel Research Laboratory, and I am sure to many people outside of our Laboratory, that a study should now be made of perceptual psychomotor tests for aircrew selection to see if the state of the art has advanced to the degree where apparatus tests could add appreciably to the reliability and validity of our selections on a cost-effectiveness basis. Early last year, George Passey and Bill McLaurin at the Lockheed-Georgia Company made a thorough review of the problem and came up with some research which indicated that the proper type of apparatus tests might very well improve our aircrew selection techniques. The apparatus being developed is moderately portable, being packaged in four principal units each weighing less than 70 pounds and aggregating less than 15 cubic feet. It embodies 23 different tests of perceptual and psychomotor functions in 19 areas identified as critical to performance and behavior related to aircrew success. These include such things as motor skills, three-axis tracking, information processing, perceptual skills with multi-displays, attention span, attention inertia, visual perceptual speed, and motor interference. The device also includes automatic recording of responses, scheduling paced and unpaced response regimes, the capacity to endure long operations, measures of resistance to distraction, and the like.

The machine is completely electronic, and uses either punched paper tape or computer tape for programming inputs and recording detailed outputs. It is completely flexible in the sense that tests can be ordered in any sequence, and presented in combinations as desired. It is automatic in its internal timing of tests, response times, and so forth. The use of this new battery appears so promising that further research on this apparatus is being conducted independently by Lockheed without additional Air Force contract money.

Operational Paper-and-Pencil Tests

Paper-and-pencil tests continue to occupy much of our attention. Presently being used for the selection of pilot trainees is a subtest score from the Air Force Officer Qualifying Test. The entire examination is used to yield several subtest scores for selection and placement of Air Force officers. The AFOQT is only one of many operational tests which are developed, scored, and revised by Laboratory personnel. Other tests include: The Armed Forces Women's Selection Test, the Airman Qualifying Examination, a pre-enrollment test for ROTC, and specialized tests such as the Electronic Data Processing Test. This "bread and butter" operation is estimated to save about 25 million dollars yearly in training costs as a result of improved selection. Operational test research and development is a reciprocal function in which the personnel researcher prepares a device of great value to the Air Force and is rewarded with data and the opportunity to advance psychometric techniques at the same time.

Human Reliability

A most complex area of research is the prediction of human social behavior. The Personnel Research Laboratory has had extensive involvement in problems of predicting human reliability and unreliability. During the first ten months of 1958 nearly 20,000 enlisted Air Force personnel were discharged for "unsuitability." Task scientists of the Personnel Research Laboratory were directed to examine this problem, and a program of research was initiated to identify the factors associated with this type of unadaptive behavior. The favorable results from this identification project led to the adoption of more stringent enlistment criteria which resulted in a marked reduction in the number of enlisted personnel discharged for unsuitability or substandard performance.

As a result of high-level concern about the possibility that an emotionally disturbed individual might detonate a nuclear weapon, the Laboratory was directed in 1959 to pursue a research program to identify potentially unreliable personnel prior to assignment to high-risk jobs. The human reliability program

implemented in 1961 is a direct product of this research effort. The employment of Laboratory-devised initial selection procedures, together with new operational techniques for personnel assigned to nuclear weapons duties and other high-risk duties, resulted in a sharp decline in unsuitability attrition within the sensitive career fields. All of this resulted in the formation of an Assessment Branch which was activated in January 1966. Nearly 10,000 new enlistees each year are screened thoroughly in advance of initiation of background investigations for possible assignment to nuclear and intelligence duties. You might be interested to know that during the first six months of this year, the Assessment Branch pre-screened 7,800 airmen and selected a little over 3,000 of them for assignment to these sensitive career areas. Of the 3,000 men selected, not one was denied a security clearance later on due to an unfavorable background investigation. The Assessment Branch screening is much less expensive and more timely than the background investigation procedures.

Another current study in our Laboratory is addressed to the problem of evaluating both sets of procedures mentioned above with the hope that improvement in cost and efficiency can be realized.

One more area of research "fallout" should be mentioned. As you know the military services are accepting increasing numbers of men in the lower mental ability level, the so-called Category IV group. Since ability level has always been one of our primary research variables in studying reliability, we have amassed considerable knowledge about the successes and failures of low-aptitude personnel. Task scientists of this Laboratory have therefore been directed to conduct research into Air Force training and personnel utilization systems so that the benefit of military service can be maximized for these people.

Historical Data Files or Longitudinal Files

One of the primary research activities within the Laboratory for many years has been the development of historical data files. This unromantic, tedious phase of personnel research has generated the empirical base for many of the studies with which the Laboratory responds to Air Force needs. We are able to do these things now because of the earlier foresight of the "data-savers," –records management regulations notwithstanding! These data files provide the constants for understanding, describing, analyzing, and projecting the personnel system, as shall be discussed in the next section.

The officer and airman data bases are generally of three forms:

1. *Information on gains to the Air Force.* Airman classification files date from 1949 and are quite detailed and complete from 1956 on. The officer input file, Project M, contains demographic, training, and evaluation data on officer gains since 1956.

2. *Strength files.* Since December 1960 detailed data have been obtained semi-annually on the total officer force; similar files have been acquired quarterly since December 1964 on the airman force. As a new strength file is obtained, composite files are constructed.

3. *Special files.* These are many and varied. Major ones include the OER file, containing the effectiveness report history of officers since 1955; officer flying and technical training, and airman technical training files from about 1950; airman reenlistment/loss files, dating from 1956, are vital ingredients for studies of trends in retention and other topics. Other special files contain experimental testing data, job inventory data, and so on. Our data banks begin, of course, with the World War II Aircrew Classification Test Battery files.

These files have been the fount from which the response to a multitude of high-priority USAF operational studies have sprung. For example, an urgent requirement by Mr. Ronald Fox, then Deputy for Management Systems, Office of the Assistant Secretary of the Air Force, resulted in a study of great consequence to military personnel. During the extremely brief time span of 13-22 February 1965,

Laboratory personnel developed new methods for assessing and reporting officer professional skills, airman reenlistment rates, and other personnel data. This maximum effort contributed markedly to the success of the Air Force Pay Study which led to increases in the military pay structure.

Project M, which permits comparative studies of input to the Air Force officer corps, was the data source for the development of an AFROTC management control system.

The files have been used to provide demographic, training, career progression, and quality analyses of scientific and engineering officers, transportation officers, civil engineering officers, logistic and supply officers, senior noncommissioned officers, and other groups. They are used routinely in studying such topics as trends in promotion and retention, officer loss rates, and the recruitment and utilization of low-aptitude personnel. The literature reveals a growing awareness among personnel researchers in the academic and industrial worlds of the need for such longitudinal bases. We are proud that we are among the first to see the need and to actually collect such data!

Job Analysis

A fifth area of research that I would like to mention is that of occupational structures research. There are four main activities: First, the development of methods for collection, analysis, and reporting of information descriptive of Air Force jobs; second, the development of methods for organizing jobs into specialties, career ladders, utilization fields, and broader management categories; third, the development of methods for evaluating the capability of personnel to perform jobs in their career ladder or utilization field; and fourth, the development of methods to determine the appropriate grade levels for officer and airman jobs.

For the first time in history, we believe that we are in a position to define precisely the work being performed by Air Force personnel. This is done through the use of the job inventory in which the tasks in a career ladder or utilization field are listed. Task inventories can be prepared which permit accurate and reliable collection of detailed information as to the tasks done at each skill level in each specialty in a career ladder. Although there may be from 300 to 600 individual tasks in a specialty, data resulting from these inventories can be processed through the computer, with reports emerging in formats describing the exact work performed by any group of interest, as for example, airmen at the journeyman skill-level, assigned to the Strategic Air Command, presently in the zone of the interior, who did, or did not, attend a technical school. Such a detailed description is ideal for planning a course curriculum, and it makes possible training aimed at specific tasks encountered on the job. It also can be used to dictate the content coverage of the Specialty Knowledge Test appropriate to a career field.

We have recently finished an additional set of computer programs to be applied to data collected through the task inventory procedures. These programs are expected to contribute to the accurate description of work actually accomplished, in a format useful to Air Force managers. Additionally, they will provide basic data for effective exploration of the second major activity—the development of methods for organizing jobs into specialties, career ladders, utilization fields, or other appropriate management groupings. These computer programs will permit the identification and quantification of differences in task performance among individual jobs, among members of homogeneous job types, or among members of different specialties within the same career ladder.

The combination of a task inventory and a computer offers a powerful new tool for exploration of the third activity—the development of methods for the evaluation of the capability of persons to perform their jobs. We have enjoyed some success in relating supervisory trait ratings to overall criteria of success, but we believe that opportunity for a major breakthrough exists in the collection of performance

evaluations at the task level. A research program is under way to evaluate this latter possibility. The Purdue Research Foundation is attempting to clarify the details incidental to this scheme. Success in collecting performance evaluations should offer us a more objective procedure to explore in the fourth activity–the determination of appropriate grades for officer and airman positions.

Another by-product of occupational research is that the development of the job inventories has given us a tool to assist in the construction and validation of our Specialty Knowledge Tests. The SKT Division makes the tests used for the upgrading of airmen from the 3- to the 5- to the 7- to the 9-skill level. The test construction outline is a plan prepared by subject-matter specialists for writing and selecting the test questions to be used. By comparing the inventory developed by the Occupational Research Division with the subject-matter specialists' test construction outline, the SKT people have an independent source of checking the completeness of the SKT. This should offer good prospects for increasing the reliability and validity of the SKTs.

Our effort to devise a methodology for the scientific determination of appropriate officer grades has been successful. This method of a mass procedure, provides data descriptive of the proper distribution of grades for a given number of officers in a given utilization area. We have recently extended this research to make it applicable to single positions at the local level by manpower officers. The rating scales use a technique whereby job requirement factors are closely aligned to a typical position. This procedure provides built-in controls to permit accurate job evaluation. The identification of typical positions to define scale levels serves also to reduce inflationary tendencies in ratings.

This type of research will enable Headquarters USAF to derive guidelines to the proper grade distribution for an Air Force of any size and with any desired mix of specialties.

Policy Capturing

Another area of research is that of capturing policy. In 1948 and 1949 when I was a student at Ohio State enroute to the PhD, we used to play around with a cute little device known as the "multiple regression equation." The multiple regression equation to us was somewhat of a scientific plaything. However, at the Personnel Research Laboratory, Joe Ward and Bob Bottenberg have developed techniques for use of the multiple regression equation as a method to capture policies of human decision-makers. This technique enables the researcher to determine the unique or interactive value of several variables in accounting for the predictable variance of the criterion. We do not depend upon abstract, theoretical assumptions but program, instead, variables in quantified form.

The machine services are not limited to processing information. If the digital computer has the same information as the human decision-maker, the predictable essentials of human information- processing in reaching decisions can be captured. Given this information and a sample of several decisions satisfying to management, the machine "learns" the policy and applies it in similar situations.

Before I am accused of promoting excessive anthropomorphic concepts, let me explain. The basic technique is simple. The computer treats the sample of decisions as a dependent variable to be predicted; the empirical information on which these decisions are based is used in various combinations to form the predictor variables. There is no problem of asking people how to weight these variables. The regression equations developed to predict the dependent variable constitute a quantitative statement of policy that is often more consistently accurate than is available from human decision-makers.

This is the method that was used by Christal in studying the proper grade level for Air Force officers which I mentioned previously. Similar techniques have been used to capture policy of selection boards. The

decisions of the promotion board for FY 63 majors was replicated with the same level of accuracy as could be expected in repetitive decisions of the board itself. I know of no method which permits capturing randomness in human judgment. The technique is being tested by the Air Training Command to provide the basic assignment policies which are included in the automated airman assignment program. This program developed by the Laboratory also includes an optimization algorithm which insures that these policies are maximally implemented. A possible advance in this area is the work of Dr. Carl Kossack at the University of Georgia under contract. He is hopeful of replacing the multiple regression analysis technique with a statistical procedure which will give a score based on (1) continuous variables, (2) dichotomous variables, and (3) discrete but numbered variables. This technique, if it works, will be especially important in developing scientific decision-making models, manpower utilization models, and assignment of personnel models. Kossack, as you remember, initially pointed out the possible application of Markovian processes to movement through personnel systems. As his own paper at this Symposium will demonstrate, we rely on him for theoretical thinking about our personnel systems research at all levels.

Mathematical Models

Mathematical models, particularly when set up for "war-gaming," are always interesting playthings. I can well remember in Air Research and Development Command days, General Lee Davis used to inveigle my boss, Colonel George Long, into playing war games on an apparatus he had devised. The Statistical Methodology and Analysis Division of Personnel Research Laboratory has adapted modern modeling techniques to immediate operative needs. An excellent example of this is a study made for General Momyer, previous Commander of the Air Training Command. General Momyer was concerned with the number of man-days lost in waiting time by pipeline personnel during the sequence of events which occur when coming into the Air Force, completing basic training at Lackland, moving through technical schools, and finally being given on-the-job assignments. Staff personnel made General Momyer aware that a lot of man-days were *lost,* but nobody had a real "feel" or "handle" for the problem. We were asked to look into this PATS problem, that is, personnel awaiting training status–and see if we could model it for use by General Momyer and his staff to war-game various permutations on the themes of efficiency and increased responsiveness to the changing needs of the Air Force. It was not a simple problem because many people come into the Air Force each month–approximately 10,000 on the average–not only from all walks of life, but from all sections of the country and from different statuses. There are those with no prior service, then there are those who have had Air National Guard training, those who have had Air Reserve Training, WAFS, ROTCs, medical trainees, and so forth. The normal six-week basic training situation complicates the problem still further. Although recruits assigned to technical schools normally take six weeks of basic training at Lackland, in times of emergency they take only four weeks of basic training at Lackland and two weeks more at technical school. Those not going on to technical schools either complete their six weeks of basic training at Lackland, or take four weeks at Lackland and two weeks at some OJT location, if basic training is offered at the OJT point. To complicate the problem still further, technical training of different types is offered at Lowry, Chanute, Keesler, Sheppard, Lackland, and Amarillo Air Force Bases, and the same training may be offered at two and even three of these bases. Training is not offered around the clock, nor even every other week. It has to be programmed for the size of the unit and length of training. After training, some of the people remain at the technical school for various lengths of time before they are shipped off to a job. In some cases people go directly from one technical training course to another technical training course. The quotas are all different and change frequently. All of these possibilities had to be considered for inclusion in the mathematical model.

Technical training is given to about 80,000 airmen per year out of the little over 100,000 who come into the Air Force. Also, at each stage of the pipeline, there is some attrition; for example, the attrition from the technical training courses is approximately 8 per cent; from basic military training, approximately 6 per cent. These facts were also "cranked" into the mathematical model which showed quite clearly where the pipeline flow could be increased at various points, and again where some of the flow perhaps could be decreased.

The PATS model is only one of a number of types of mathematical models which have been developed at the Laboratory and which have found practical use. At the present time, you might be interested to know that the Laboratory has been asked to model the entire Air Force flow and operation. This, of course, is a tremendous undertaking.

The job is to be accomplished in phases, and a two-year study effort appears reasonable before a workable model can be expected. When we are able to do this, the planners in the Pentagon, particularly in Deputy Chief of Staff/Personnel, will have a most powerful tool.

We could mention probably another hundred examples of what we have done for the Air Force lately, long-range research, surveys, "brush fire" operations, etc., but I think that these examples give you some idea of the type of work that we are doing in our various divisions.

I cannot help but feel a great pride for the accomplishments of the Personnel Research Laboratory over its many years of existence. The Laboratory is populated by a highly competent staff of civilians and military, professional and support personnel. There is an intense desire to undertake new and more complex tasks, to develop new and more powerful tools, and to serve all communities of the world in new and more exciting ways.

VI

COMPUTER TIME SHARING*

Bert F. Green, Jr. and L. Rowell Huesmann
Carnegie Institute of Technology

The computer world changes so rapidly that a generation lasts for only about five years. Each generation has provided larger, more powerful, and more versatile computers. In each generation computers have been used by more and more people in an ever-widening range of applications. Dramatic changes are also occurring in procedures for using computers. In the present generation, a procedure called time-sharing is revolutionizing computer usage.

People who have not suffered through the discouraging process of using a digital computer may wonder what the excitement is all about. A few years ago the senior author had an enlightening conversation with the science editor of a national news magazine, who was remarkably well informed about current developments in artificial intelligence, computer simulation, and other novel uses of computers. But toward the end of the largely "wild-blue yonder" conversation, he came sharply back to earth by saying, in effect, "One thing I don't understand is how you fellows go about using a computer. What do you do on a typical day? Do you sit at your desk and think, or do you write things on paper at your desk, or do you punch cards, or do you wire connections with a soldering iron?" Plainly, even astute laymen know very little about how one actually goes about using a computer.

The procedures that were described in response to the science editor's query are still used today at most computing centers (1). Nevertheless, we shall call these procedures "the old way" and shall sketch them briefly to provide a contrast with the new. The old way to use a computer is first to plan the program, perhaps drawing a flow diagram if the program is very complicated. Then the prospective user writes his program on special programming sheets, using the appropriate computer language. These sheets are transformed into punched cards either by the programmer himself or by a keypunch operator. The deck of cards, after being checked for errors, is left at the computer center to be run on the computer. Some time

---

* This paper is based on an address "Computers: The Present State of the Art" presented by the senior author at a symposium arranged by Division 21 at the 1965 convention of the American Psychological Association. Preparation of this paper was supported in part by Grant MH-07722 from the National Institutes of Health, United States Public Health Service and in part by the senior author's Public Health Service Fellowship 1 F3 MH28, 05601, from the National Institute of Mental Health. The paper was written at the Center for Advanced Study in the Behavioral Sciences.

Ed. note – A very interesting and informative part of the Symposium program was the demonstration of computer timesharing conducted by Captain George H. Walther of the Personnel Research Laboratory. A copy of the text of his presentation, together with illustrations of the information retrieval through items with systems at the USAF Military Personnel Center at Randolph AF Base and the System Development Corporation in Santa Monica, California may be obtained from the Personnel Research Laboratory upon request.

later - often three or four hours later - it will have been run and will have produced some sort of an output. Typically the program will not have operated correctly, and the output must be examined for clues to the errors in the program, a process known as "debugging." Appropriate cards are changed in the program to correct the errors, and the deck is left at the center to be run again. This cycle repeats as often as necessary until the program finally works properly. Then the so-called production run is made in which the program does what it was intended to do initially.

Because of the piecemeal nature of this work, and the intervening pressures of other business, the time between writing the program and obtaining useful results is often measured in weeks. If someone does it all, from start to finish, in two or three days, he is usually regarded as some sort of intellectual freak. For most mortals, so much effort is required that computer programs are seldom written if they are to be used only once. The pain of getting a program to work, coupled with the general feeling that a program is worth writing only if it will be used often, has kept some people from taking advantage of computers. Others of us write the program anyway, because we think it's fun. We justify our behavior by offering to share our program with others, but the offer is seldom accepted; using someone else's program turns out to be about as time-consuming as building one's own, particularly if it has been prepared for use at a different computer center.

All this is changing. A few lucky places now have remotely accessed, time-shared computing facilities where many people can use the computer simultaneously via teletype links, with nearly immediate feedback of results. A program may be written and debugged in one short session. To use one of these new systems, the programmer begins, as he always did, by planning his program. Then he finds a convenient station having a teletype or similar keyboard instrument, very likely located remotely from the computer center. He calls the computer with a normal telephone dial on the console, thus connecting his keyboard with the machine. When the connection has been made, the computer will cause the keyboard to type an acknowledgment. The programmer gives his code number and/or a secret password which, if valid, is followed by a computer request for further instructions. The programmer then types his program directly into the computer from the keyboard, bypassing the writing and subsequent card punching in the old system. As he types, his program is automatically stored in the computer auxiliary storage - probably a disc file. When he finishes typing his program, he may ask the computer to repeat it to him as a check, and then to execute it immediately or to file it away. If he asks for execution, he will get the results in a few minutes. The output produced by the computer will typically indicate that an error has been made, at which point he can correct his mistake, changing the copy of his program that is still stored in the computer. He can then tell the computer to run his program again, and the cycle can continue until the program runs correctly. The entire process may span only an hour.

The new procedure is economically feasible because the computer can service many users working at the same time. Each user actually needs the computer only a short fraction of the time that he spends at the keyboard. When he is typing the program or making a change in the program, the computer is needed only to receive and store the input, and, at the end of each typed line, to do some simple syntactic checks, consuming no more than a few milliseconds. Thus, many users can be typing programs at the same time without producing any noticeable strain on the machine. Almost all of the computing time remains available for executing other programs. When several users request programs to be executed at the same time, they share the available computing time. The computer parcels out its time in short periods called quanta. Each program waiting to be executed gets a quantum of computer time. Then its execution is suspended and the computer turns to the next program, in a round-robin fashion. When a program is completed, or when it must pause to produce an output, it is removed from the queue, and the results are transmitted to the user. The time that a program must spend in the round-robin queue to get its needed amount of computing time depends mainly on the *number* of others in the queue, but not on the time requirements of the competing programs.

Of course the old system also parcels out computer time to many users, but in a manner better suited to the machine than to the user. The old procedure is called batch processing. When a program is submitted to be run, it is first held until a batch of programs is collected. Then the batch is made ready for input to the computer. Some time later, the computer executes each program in turn, and records the results on a tape. Still later, the tape is fed to a printer which prints the results of each program in the batch. With such a scheme, the total time to process any given program can seldom be less than an hour. Usually it is two to four hours, during which time the programmer must twiddle his thumbs or otherwise busy himself. In time-sharing, the computer works at the same time as the programmer, translating his program as he writes it, and running it as soon as it is completed. Because the computer is so much faster than the input-output devices or than the programmer's own thought processes, it can service a large number of people "simultaneously," giving each the illusion that the computer is his while he is connected to it.

Requirements of Time-Sharing

Time-sharing procedures require three things: first, the hardware for handling communication with a large number of remote stations; second, a large, readily accessible auxiliary storage device; and third, a sophisticated monitoring program that responds to signals from the remote stations, keeps the round-robin computation going, and produces outputs for the remote stations, always insuring that the several concurrent programs do not get mixed up with each other. During the last decade all three of these capabilities have undergone rapid development.

Procedures for handling digital inputs from several uncorrelated sources are at least as old as the SAGE system. Procedures for transmitting digital information over telephone lines are even older; teletypes have been using such a scheme for years. Demands for virtually error-free performance, however, are new and are now being met. It has become standard practice for computers at regional offices of large companies to send data directly to the main computer at national headquarters. Airlines reservation systems use similar procedures. Devices have now been built consisting of a keyboard, encoder, and cradle for a telephone hand set so that an ordinary telephone can be turned into a teletype.

The second requirement, large amounts of storage, is being met by disc files. Very large disc files as well as very large core memories are now available. Still, storage is one of the severe problems in time-sharing. All of the experimental time-sharing systems ran out of storage space very swiftly because the designers forgot that programmers hate to throw away old programs. Before time-sharing, programmers cluttered their offices with card decks; now they clutter computer storage. In practice, of course, there is continual pressure to purge useless programs; unused programs are relegated to a storage tape in the "deep-freeze" that can be recovered only with manual intervention and considerable delay. Nevertheless, the working file must be huge.

The third requirement, the complicated monitor that controls everything, is a natural outgrowth of two lines of software development - programs for on-line real-time control systems like SAGE and the collections of utility programs, compilers, and monitors that have grown up in computer centers. In the last decade, programming languages like FORTRAN, ALGOL, and JOVIAL have gained universal acceptance. Moreover, techniques for producing compilers for such languages are so well understood that new languages with their compilers are continually being developed. With so many compilers on hand, it becomes a tedious task to get the computer ready for each new program - a task of the sort computers do better than people, so supervisory programs called monitors were developed to run the programs automatically. It is but a short step from such monitors and from real-time multiconsole control systems to general purpose time-sharing systems.

In 1959 all of these developments were well along the time was ripe. According to A. L. Samuel (2), at about this time Christopher Strachey (3), John McCarthy (4), J. C. R. Licklider (5), F. J. Corbato (6), and J. C. Shaw (7) started work in time-sharing. The earliest systems in operation were the CTSS system at the M.I.T. Computation Center (8), the system on the PDP-1 computer at Bolt Beranek & Newman (9) and the JOSS system at RAND (7). Shortly thereafter followed the two large-scale experiments, Project MAC at M.I.T. (10) and the Q-32 time-sharing system at the System Development Corporation (11) both supported by the Advanced Research Projects Agency (ARPA) of the United States Department of Defense, largely through the efforts of J. C. R. Licklider. Other systems were put into operation at Dartmouth, Stanford, and Carnegie Institute of Technology, and the list is growing rapidly. Several commercially available systems have been reported recently. IBM offers a version of FORTRAN called QUIKTRAN, GE offers BASIC and ALGOL; Adams and Associates, Bolt Beranek & Newman, and a variety of new companies are offering services.

Technical Problems

Time-sharing can be implemented on any general purpose digital computer. Special in-out equipment is needed to manage the remote consoles, but nothing special is required of the central processor. The Psychology Department at Carnegie Tech has a computer-based laboratory built around a small (16K) machine, a DDP-116 (12). Several independent experiments can be run at the same time, or several subjects can be run in one experiment. But we could do better, and the large systems could too, with a machine specifically designed for time-sharing.

The main problem in implementing a time-sharing system is precisely how to share the time. At any point in time there may be inputs to be processed, programs to be executed, and outputs to be transmitted. The inputs and outputs are not time-consuming, but they are insistent. Some large systems now use a separate small computer for input and output processing so the main processor can spend all its time executing programs. It seems likely that this will be the best solution.

There remains the problem of executing the programs waiting in the round-robin queue. How long should the computation quantum be? Short quanta are needed to provide fast response to users, but short quanta mean large amounts of "swap time," the time it takes to switch from one program to the next. On present-day machines, the current program must be transferred to auxiliary storage, and the next program in the queue must then be transferred into the main memory before computation can resume. The next generation of machines will have a main memory sufficiently large to hold several programs. Then programs can be swapped at the same time that another program is being executed.

Another difficulty with short quanta is that the monitor program itself must operate between quanta to keep its tables up to date. The time spent in this way is called overhead. Some installations plan to eliminate monitor overhead by having a separate processor for the monitor. This is expensive but does permit the prime processing unit to spend all of its time in executing programs in the queue.

The most legitimate complaints come from users with long programs, especially production programs, who were well served by batch processing. Short quanta reduce machine capacity, which seems always to be in short supply. Noninteractive programs have no need of time-sharing, and their authors do not like to pay the price of reduced machine capacity in order to provide special services for others. Such complaints will be less valid when swap time and overhead are virtually eliminated. Another possibility is to provide a separate processor for long production programs. Of course, no matter what is done, some programmers will be dissatisfied. The same programmers who complained about four-hour delays in batch processing now complain about four-minute delays in time-sharing.

A different problem arises because several programs are in the same computer. Each program must somehow be prevented from accidentally, or deliberately, interfering with the others. When several programs are to be in primary memory at the same time, hardware devices called memory-protection registers allow the monitor to serve as referee, sealing off each program from the others and from the monitor. Software features in the monitor can also guard against inadvertent access to someone else's records in auxiliary storage. Although today's systems can provide protection from accidents and honest mistakes, they are probably no match for a clever spy. New developments are needed to guarantee privacy which is essential in some applications.

Developments in communication devices are also needed. The keyboard is not the best way for men and machines to communicate. Graphic input and output are needed. Current graphic displays, in the form of cathode-ray tube devices, and corresponding graphical input procedures are wide-band devices that cannot conveniently be controlled at remote distances from the computer. Methods of handling such equipment remotely will undoubtedly be developed. Further in the future is the possibility of transmitting computer programs by voice rather than keyboards, and even more time will elapse before we are able to talk to the machines in natural language.

Currently the language that must be used in man-computer conversations is far from natural. Time-sharing systems have not altered computer languages much; they have merely added editing features. Considerable improvement can now be made by paying close attention to designing better languages. Yntema (13) argues forcibly that the user must do too much of the bookkeeping that could as easily be left to the computer. In doing matrix manipulations, for example, the user should need to report the dimensions of a matrix only once and then let the computer keep track, as JOVIAL does, but as FORTRAN and ALGOL do not. He further argues that such languages must be interactive. The scientist typically wants to try something, look at the results, consider his next move, and try again. In fact, time-sharing systems have spawned several interactive languages oriented toward the problems in a particular discipline. The list includes computer-aided design (14;15), operations analysis (16), structural engineering (17), algebraic computation and analysis (18), and mathematical analysis (19;20). For personnel research, the most interesting and useful such language is TRACE (*T*ime-shared *R*outines for *A*nalysis, Classification, and *E*valuation) designed by Moore, Meeker, and Shure at System Development Corporation (21). Problems of data analysis in the social sciences can be expressed easily in the TRACE language, and many scientists who are unwilling to learn ALGOL or FORTRAN will now be able to use the computer.

Advantages of Time-Sharing

Time-sharing systems are very popular, and programmers who have used them are extremely reluctant to go back to the old method. Although the experimental systems were supplements to the computing power at their locales, they quickly became overloaded. Sk[illegible] [illegible] this is just an instance of Parkinson's Law: Work expands to fill the time available for its completion. A more likely culprit is the first law of the instrument: Give a boy a hammer and he will soon find that everything needs pounding. Time-sharing has given a computer, in a sense, to many people who would not trouble to learn computing the old way.

But there are more cogent reasons for the popularity and very favorable acceptance of the time-sharing systems. First, the time-sharing systems are especially well suited to one-shot programs in which the concept of a production run is not applicable. The production program is expected to be run over and over again with different sets of inputs, whereas the one-shot program will be run only once with one set of data, and that will be that. A great many calculations in research projects at academic institutions

and at research laboratories turn out to be of the one-shot variety. Many such programs take only a few seconds of calculation, but every new program, no matter how short, requires a certain amount of debugging. Under the old system, if a job could be done by hand in a few days by using a desk calculator or its equivalent, then it might as well be done that way because the computer is not likely to beat that time. Under the new system the break-even time is a few hours rather than a few days. In fact, in many cases it would pay an engineering student to do his homework on a computer, if he were allowed access to it.

The conversational mode of time-sharing not only shortens the overall process, but it actually saves computer time. A major source of programming errors is debugging itself. Correcting one error often results in a second error because the programmer has forgotten the details of his own program. When the writing and debugging can all be done at one sitting, the programmer can "stay in context" throughout the process, he will make fewer errors, and less computer time will be used in debugging. Also, the programmer can stop the execution of his program if it starts producing obviously erroneous results, thus saving computer time that would have been wasted in a batch-processing mode.

An important by-product of the conversational mode is that the computer can be used as a teaching machine for beginning programmers. Trial and error, followed by immediate feedback, results in swift learning, again promoting the use of the computer by persons who have one-shot problems. Further, the ability to correct errors immediately reduces the onus of error. Computer languages require such precision in stating problems that humans inevitably err. When each error, however trivial, costs hours of lost time, new programmers easily become discouraged; even experienced programmers have difficulty coming to terms with their own fallibility. The conversational mode drastically reduces the cost of errors, programming becomes less painful, and more scientists are willing to invest the time to learn the art of using computers. It is not frivolous to conclude that the productivity of the scientific community has thereby been increased.

A third advantage is the considerable convenience of having programs stored within the machine, at the fingertips but not cluttering the office. Programs for standard analyses can be set up to be used as needed on new sets of data. Or these programs can be reassembled into new ones for special jobs. The storage feature is also useful when several programmers are cooperating on aspects of a large system; each can have immediate access to the other's products.

Remote access has some additional advantages. Anyone having access to a teletype with the inverted computer-coding feature may, if permitted, use any computer system with a remote facility. A scientist who needs to use a program written by a colleague at a different computation center no longer has to endure the headaches of moving the program to his own computer. With remote access he can use the program on the machine for which it was written, with a distinct saving of time and effort. Also, a scientist may use his own programs on his home computer when he is away from home–e.g., at the Center for Advanced Study in the Behavioral Sciences–rather than transplanting his programs to a new computer. Finally, it is reasonable for small colleges and universities to consider pooling their resources to establish a central computing center that can serve the needs of all of the participating universities by means of remote time-sharing facilities.

Costs

The extent to which the advantages of time-sharing are worth their cost cannot be determined yet. Time-sharing systems are in their infancy and any comparative evaluation at present would very likely favor the more well-established procedures. Certainly there are some serious questions about the economic feasibility of such procedures, but it is too soon to say anything very definitely.

At present, the costs of the time-sharing systems are considerable. In addition to the costs of necessary hardware and software, there are a number of losses in efficiency in comparison with the old procedure. First, the machine is now handling the remote inputs and outputs, being a sort of file clerk for the programmer, which is a new service not provided in the old system. Second, there are slightly longer service times because the monitor is much more complicated than in the old system. But the largest costs are the swap time and monitor overhead time spent as the computer switches from one program to the next in the execution queue. New equipment is being designed to eliminate or drastically reduce these times, but additional equipment usually means added expense.

Although the costs are substantial, the debate is reminiscent of earlier issues in computing that have all been decided in favor of convenience and power rather than speed, machine efficiency, and expense. The use of floating-point rather than fixed-point arithmetic was such an issue in the early days. Fixed-point calculations were faster for the machine but more difficult for the programmer. Programming courses used to include discussions of "scaling the problem," referring to ways of keeping track of the decimal point in fixed-point calculations. Today floating point is used almost exclusively. A similar argument arose when FORTRAN was introduced. FORTRAN greatly simplified the task of programming, at the cost of both compiling time and a little bit of running time, since compiled programs were slightly longer than an optimum machine language program. Today compilers like FORTRAN are used universally.

There is even a question whether the time-sharing system should strive to reduce operating time as much as possible. Today there is a tendency to criticize FORTRAN for spending too much time trying to write optimum programs. Today's compilers, like MAD and ALGOL, tend to work much faster than FORTRAN and to produce less elegant programs just because, in one-shot programs, the compiling time is much more important than running time. Perhaps in time-sharing it will turn out that the main problem is to reduce total time at station, so that most effort should be spent in improving debugging and editing facilities rather than in improving the efficiency of the monitor.

But the final worth of time-sharing will be in the new users it brings to the computer, the additional research it inspires. With the development of special interaction languages suited to the problems of specific disciplines, scientists in all fields will literally have computers at their fingertips. New things are sure to follow. In personnel research, as I have recounted elsewhere (22), new methodologies are already appearing, such as Shepard's non-metric methods (23), and new solutions are appearing for old problems, such as Eber's maxplane algorithm for factor rotation (24). With a computer, iterative solutions are feasible, as are solutions that depend on hill-climbing search methods to find best fitting parameters. Special programming languages like TRACE will facilitate data analysis. Even methodologists will be helped in their everyday work. How often have you gotten hung up on some algebraic expression that keeps getting longer and longer. Programs like FORMAC (25) and Formula ALGOL (26) are now available to do the manipulations and simplifications for you.

Conclusion

Much remains to be learned about time-sharing systems, and many changes and improvements will certainly be made in the next few years. However, it is already safe to say that time-sharing will become the standard way of using computers in the future. Large fractions of recent computer conventions have been devoted to time-sharing. In the fascinating fast-paced world of computers, time-sharing is definitely in. Mechanisms are already being discussed (27) to provide flexible capacity at a computing center to take care of peak loads. Eventually, as McCarthy envisaged (4), we are likely to see a computational utility akin to electric power or the telephone system, with flexible capacity and with remote access by thousands of users. Long before that happens—indeed within the next five years—research scientists will be as dependent on their teletype link to the computer as they are now on their slide rule, desk calculator, and telephone.

REFERENCES

1. Green, B.F., Jr. *Digital computers in research.* New York: McGraw-Hill, 1963.
2. Samuel, A.L. Time-sharing on a multiconsole computer. *New Scientist,* 1965, 26, 583-587.
3. Strachey, C. Time-sharing in large fast computers. *Proceedings of the International Conference on Information Processing,* UNESCO, 1959.
4. McCarthy, J. Time-sharing computer systems. In Greenberger, M. (Ed.), *Management and the computer of the future.* Cambridge, Mass.: M.I.T. Press, 1962.
5. Licklider, J.C.R. Man-computer symbiosis. *IRE transactions on human factors in electronics,* 1960, HFE-1, 4-11.
6. Corbato, F.J., Daggett, M.M., & Daley, R.C. An experimental time-sharing system. *Proceedings of the Spring Joint Computer Conference,* 21, Baltimore, Md.: Spartan Books, 1962.
7. Shaw, J.C. JOSS: A designer's view of an experimental on-line computing system. *AFIPS Conference Proceedings,* 26, 1964, 455-464.
8. Corbato, F.J. et al. *The compatible time-sharing system: A programmer's guide.* (1st Ed.) Cambridge, Mass.: M.I.T. Press, 1963.
9. Boilen, S., Fredkin, E., Licklider, J.C.R., & McCarthy, J. A time-sharing debugging system for a small computer. *Proceedings of the Spring Joint Computer Conference,* 23, 1963.
10. Fano, R.M. The MAC system: A progress report. In Sass, Margo A. & Wilkinson, W.D. (Ed.) *Computer augmentation of human reasoning,* Washington, D. C.: Spartan Books, 1965.
11. Schwartz, J. A general purpose time-sharing system. *Proceedings of the Spring Joint Computer Conference,* 25, 1964, 397-411.
12. McLean, R.M. The real-time computer in a psychological laboratory. CIP Working Paper #86, Carnegie Institute of Technology, April 12, 1966.
13. Ystema, D.B. *The software problem.* Lincoln Laboratory, M.I.T., Lexington, Massachusetts, Group Report 1964-51, 1964.
14. Coons, S.A. An outline of the requirements for a computer-aided design system. *Proceedings of the Spring Joint Computer Conference,* 23, 1963, 300.
15. Ross, D.T. & Rodrigues, J.E. Theoretical foundations for the computer-aided design system. *Proceedings of the Spring Joint Computer Conference,* 23, 1963, 305-322.
16. Greenberger, M., Jones, M.M., Morris, J.H., & Ness, D.N. *On-line computation and simulation: The OPS-3 system.* Cambridge, Mass.: M.I.T. Press, 1964.
17. Biggs, J.M. & Logcher, R.D. *Stress: A problem-oriented language for structural engineering, Project MAC.* Cambridge, Mass.: M.I.T. Press, TR-6, A64.
18. Wantman, M.E. *Calculaid: An on-line system for algebraic computation and analysis, Project MAC.* Cambridge, Mass.: M.I.T. Press, TR-20, 1964.
19. Culler, G.J. & Fried, B.D. The TRW two-station, on-line scientific computer: General description. In Sass, Margo A. & Wilkinson, W.D. (Ed.) *Computer augmentation of human reasoning,* Washington, D. C.: Spartan Books, 1965.
20. Kaplow, R., Strong, S., & Brochett, J. *A system for on-line mathematical analysis, Project MAC.* Cambridge, Mass.: M.I.T. Press, TR-24, 1966.

21. **Moore, W.H. Jr., Meeker, R.J., Jr., & Shure, G.H.** *TRACE model II: Time-shared routines for analysis, classification, and evaluation.* Santa Monica, Calif., System Development Corporation, TM-2621, 1965.

22. **Green, B.F., Jr.** The computer revolution in psychometrics. *Psychometrika*, (in press).

23. **Shepard, R.** The analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrika*, 27, 1962, 125-140, 219-246.

24. **Eber, H.W.** Toward oblique simple structure: Maxplane. *Multivariate behavioral research*, 1, 1966, 112-125.

25. **Bond et al.** FORMAC, an experimental formula manipulation compiler. *Proceedings of the 19th National Conference of the Association for Computing Machinery*, 1964.

26. **Perlis, A.J. & Iturriaga, R.** An extension to ALGOL for manipulating formulae. *Communications of the ACM*, 7, 1964, 127-130.

27. **Corbato, F.J. & Vyssotsky, V.A.** Introduction and overview of the Multics System. *Proceedings of the Fall Joint Computer Conference*, 28, 1965, 185-196.

VII

PROBLEMS IN PERSONNEL RESEARCH

Lloyd G. Humphreys
The University of Illinois

It is not very remarkable to conclude that the principal problem in selection research is that we do not predict very well. It is a modest, useful level, but a long way from our ceiling. Perhaps some of the ramifications of this conclusion will be more interesting. Additional aspects of our difficulties in prediction are that we have not been showing much improvement in accuracy, that accuracy diminishes markedly as the time interval between predictor and criterion information increases, and that further decreases in accuracy seem to occur as the subject whose behavior we are trying to predict moves from training to the job.

It is somewhat more remarkable to conclude that trait theory, and in particular that part of trait theory stemming from factor analysis and homogeneity analysis, has become and may always have been a dead end. Until better psychological theories are developed it will be more useful to concentrate on becoming better test theorists and prediction theorists than to continue the search for pure trait measures. I hold also that there is no great difference between test theory and test technology. I recommend, in other words, that we approach the prediction task in an engineering fashion. We should try out the theories that arise from basic psychological research, but we should not hang on to theories that have been tried and found wanting. Many psychologists apparently feel that to work without a basic theoretical orientation is not scientifically respectable even though the theory is useless or even misleading.

Documentation and logical support for the above generalizations are obviously needed. A number of years ago Robert Travers (1) in the Air Force Personnel Laboratory reviewed the trend of predictive validities over a number of years for common criteria. The conclusion was inescapable that little progress had been made in the prediction of these criteria and that the level was only moderate. Dr. Edgerton also reached this conclusion as we heard yesterday.

New criteria may present apparent exceptions. During World War II, in the Army Air Forces we made remarkable progress in the space of a few short months in the prediction of success in primary pilot training (2), but we reached a ceiling rather quickly. As soon as we found out we were not in the academic ball park, which is where we started in the spring of 1942, we were able to move by the following winter to the mechanical, spatial, practical ball park and to values of validity coefficients very close to final values. Rapid information feedback and good test technology were the primary bases for this achievement, not basic psychological theory.

It is also somewhat deflating to psychological theory to review the history of some of the most valid tests for pilots. The Complex Coordinator had some slight connection with an early psychologist, O'Rourke of the Civil Service Commission, but the main developmental work was done by flight surgeons (3). The Rudder Control test (3) was built originally by a pilot interested in training and was being used as a training

device at a primary school when Neil Warren and I saw it on an early field trip. It was converted into a test by psychologists because it looked promising empirically, which is another way of saying that we had a hunch it would work. Instrument Comprehension (4) grew out of Wilbur Gregory's interviews with "washed out" pilot trainees, but true to his academic training Gregory's version was a highly verbal affair that for a short while seemed to require a negative weight in the classification battery. The much more valid photographic version was the idea of a bright enlisted man, Milton Burdman, who had had only a smattering of psychology in his undergraduate work. The General Information test and its predecessors (4) can be attributed primarily I believe, to John Flanagan, but it is hard to imagine anything that departs more widely from the notions of trait-factor-homogeneity theorists.

A well known review of on-the-job validities has been done by Brown and Ghiselli (5). They are low. Thorndike and Hagen (6) had generally negative results in predicting present proficiency criteria in their followup study of World War II Aviation Cadet examinees. A long history of military personnel research shows that tests do not predict on-the-job success as an officer very much better than a set of random numbers would. If you send a group of officers back to school, however, their academic success can be predicted moderately well. Officer effectiveness on the job can also be predicted moderately well by peer ratings obtained during training. Unreliability of criterion information and restriction of range of talent are indeed attenuating factors, but these factors do not explain the very low test validities obtained for on-the-job criteria.

The stability of performance measures over time is another critical issue. John Anderson (7) summarized such data for intelligence test results a number of years ago, but little attention was paid to either his findings or his conclusion. He stated that the stability data were consistent with increments in test scores for a given period of time, being uncorrelated with the base at the beginning of the period. What is this base? It is the response repertoire of the subject which has been learned, but it is not limited to specifically acquired skills and knowledge since stimulus generalization and transfer are also principles of learning. This should have severely shaken trait theorists. More recently Juola (8) has published stability data for grades in college of several semesters and the trends of predictive validities for the independent semester grade averages. His data are consistent with Anderson's interpretation. I have tried to relate these and other phenomena to Guttman's simplex and have published additional data (9). Today, I would like to review a few sets of published data and to present some presently unpublished data of the same sort.

Table 1 presents test-retest data for boys over a period of several years. The basic data were from the Harvard Growth Study (10), though I am responsible for computing the correlations. It should also be noted that the N varies from one cell entry to another. I shall attempt no detailed analysis of this or any other table. The form of the matrix is the important thing to observe: It is apparent that the high values of the correlations are adjacent to the principal diagonal and the lowest values in the table are furthest removed from the diagonal.

Table 2 shows that a similar pattern is obtained over a much shorter time interval in the learning of a motor skill and for a constant N (11). Intercorrelations of a physical growth measure appear in Table 3, again for boys and again for a changing N. Tables 4 and 5 contain intercorrelations for height and weight for a constant N for girls. There is no sex difference in this pattern. There is also no difference arising from the type of variable.

We now turn to the public school and some commonly used intellectual measures. Tables 6 through 10 present intercorrelations based on a constant N for various time spans for children in the Champaign schools. A similar pattern appears within the time limits of a single semester, this time high school mathematics. The data are in Table 11 and again the N is constant.

Table 12 contains the intercorrelations of grades in successive semesters of foreign language instruction at the college level. Table 13 has semester grade-point averages of engineers, with each semester average being computed independently of the rest. Would an on-the-job proficiency measure simply follow the trend or would there be a basic discontinuity? I do not know the answer, but it is possible that low on-the-job validities demonstrate, at times, nothing more than performance instability over time. Table 14 has similar data for students in a liberal arts college. In all of these tables, the N is again constant.

It is obvious that unreliability does not produce the lower correlations observed almost uniformly in the upper right or lower left corners of the tables. Neither is restriction of range of talent involved. People are changing. It is also well-nigh certain on the basis of logic alone that predictive validities of tests given at the beginning of the period will decrease in a similar fashion. Juola's data, previously mentioned, confirm this.

Table 1. **Intercorrelation of Mental Ages of Boys at Various Chronological Ages**

(First Test)

	8	9	10	11	12	13	14	15	16	17	18
8		721	712	747	729	657	598	648	652	556	444
9	721		751	721	714	696	634	615	609	588	509
10	712	751		816	769	704	726	738	699	604	543
11	747	721	816		859	787	745	810	802	736	638
12	729	714	769	859		854	778	786	806	775	732
13	657	696	704	787	854		864	785	770	780	754
14	598	634	726	745	778	864		839	778	750	765
15	648	615	738	810	786	785	839		868	778	744
16	652	609	699	802	806	770	778	868		848	788
17	556	588	604	736	775	780	750	778	848		828
18	444	509	543	638	732	754	765	744	788	828	

Table 2. **Intercorrelations of Trials of Complex Coordination Task**

(from Fleishman and Hempel, N = 197)

	1	2	3	4	5	6	7	8
1		75	73	66	64	57	69	59
2	75		85	85	84	79	77	79
3	73	85		85	83	79	81	79
4	66	85	85		90	88	86	85
5	64	84	83	90		90	87	86
6	57	79	79	88	90		85	86
7	69	77	81	86	87	85		90
8	59	79	79	85	86	86	90	

Table 3. **Intercorrelations of Sternal Heights of Boys at Various Chronological Ages**

	8	9	10	11	12	13	14	15	16	17	18
8		978	975	964	943	906	858	844	850	823	760
9	978		988	979	960	923	878	863	862	827	775
10	975	988		990	975	940	889	873	873	842	782
11	964	979	990		987	957	909	883	880	843	801
12	943	960	975	987		976	928	893	877	836	808
13	906	923	940	957	976		967	919	872	811	743
14	858	878	889	909	928	967		962	889	789	687
15	844	863	873	883	893	919	962		952	836	699
16	850	862	873	880	877	872	889	952		940	817
17	823	827	842	843	836	811	789	936	940		950
18	760	775	782	801	808	743	687	699	817	950	

Table 4. **Intercorrelations of Standing Height of 275 Girls at Various Chronological Ages**

	7	8	9	10	11	12	13	14	15	16
7		987	980	957	920	897	887	866	836	810
8	987		989	969	934	914	904	882	850	824
9	980	989		986	954	927	909	881	844	814
10	957	969	986		979	947	911	865	816	780
11	920	934	954	979		974	923	855	790	747
12	897	914	927	947	974		964	887	810	763
13	887	904	909	911	923	964		961	901	860
14	866	882	881	865	855	887	961		974	948
15	836	850	844	816	790	810	901	974		989
16	810	824	814	780	747	763	860	948	989	

Table 5. **Intercorrelations of Weight of 275 Girls at Various Chronological Ages**

	7	8	9	10	11	12	13	14	15	16
7		890	880	835	810	793	755	773	744	732
8	890		920	896	871	856	825	812	771	759
9	880	920		932	906	882	840	818	773	756
10	835	896	932		958	936	892	842	777	755
11	810	871	906	958		967	921	866	790	762
12	793	856	882	936	967		954	892	816	775
13	755	825	840	892	921	954		944	880	839
14	773	812	818	842	866	892	944		953	916
15	744	771	773	777	790	816	880	953		965
16	732	759	756	755	762	775	839	916	965	

Table 6. Intercorrelations of Scores on California Test of Mental Maturity

(N = 176)

	Verbal				Performance		
	1	3	5		1	3	5
1		59	54	1		52	31
3	59		53	3	52		41
5	54	53		5	31	41	

Table 7. Intercorrelations of Vocabulary Scores from Grade to Grade

(N = 278)

	2	3	4	5	6
2		65	58	63	56
3	65		65	73	68
4	58	65		72	65
5	63	73	72		76
6	56	68	65	76	

Table 8. Intercorrelations of Reading Scores from Grade to Grade

(N = 278)

	2	3	4	5	6
2		61	64	64	58
3	61		63	71	67
4	64	63		74	65
5	64	71	74		73
6	58	67	65	73	

Table 9. Intercorrelations of Scores on Two Arithmetic Tests in Three Grades

	Arithmetic Fundamentals (N = 142)				Arithmetic Reasoning (N = 278)		
	4	5	6		4	5	6
4		59	43	4		70	63
5	59		62	5	70		70
6	43	62		6	63	70	

Table 10. Intercorrelations of Scores on Two Language Tests from Grade to Grade

(N = 278)

	Spelling				Language Mechanics		
	3	4	5	6	4	5	6
3		52	49	56	4	67	59
4	52		63	62	5	67	68
5	49	63		69	6	59	68
6	56	62	69				

Table 11. **Intercorrelations of Four Hour Exams on High School Mathematics**

(Data from M. D. Terman, N = 600)

	1	2	3	4
1		72	70	69
2	72		75	72
3	70	75		76
4	69	72	76	

Table 12. **Intercorrelations of Grades in Four Semesters of Beginning Foreign Language**

(N = 613)

	1	2	3	4
1		71	61	60
2	71		70	65
3	61	70		75
4	60	65	75	

Table 13. **Intercorrelations of Grades in the Eight Semesters in Engineering**

(N = 269)

	1	2	3	4	5	6	7	8
1		54	42	37	40	35	31	33
2	54		56	50	47	42	32	34
3	42	56		58	60	48	43	48
4	37	50	58		53	55	31	30
5	40	47	60	53		66	56	46
6	35	42	48	55	66		56	50
7	31	32	43	31	56	56		55
8	33	34	48	30	46	50	55	

Table 14. **Intercorrelations of Semester Grades in the Liberal Arts College**

(N = 414)

	1	2	3	4	5	6	7	8
1		59	56	53	47	37	42	43
2	59		52	51	48	40	41	45
3	56	52		53	51	47	35	41
4	53	51	53		51	56	47	48
5	47	48	51	52		57	48	50
6	37	40	47	56	57		47	46
7	42	41	35	47	48	47		63
8	43	45	41	48	50	46	63	

Instability of performance does not rule out the use of the trait concept, but it does complicate the theory. There are other data in the personality domain that complicate trait theory much more severely. These are the data that demonstrate situational specificity. For example, it is possible to obtain reliable ratings of personality traits in the home and in the school. Independent ratings made in each situation are positively and substantially correlated, but correlations between ratings of behavior at home and at school are much lower (12). It has also been shown that children having certain background characteristics in common will behave one way with the mother, another seemingly opposed way with the teacher (13). Peterson (14) has summarized these and other converging lines of evidence and has reached the conclusion that situation is a much more potent contributor to variance than is trait.

The clinching argument concerning the untenability of the trait concept is primarily logical, not empirical. It is applicable *in toto* to personality and ability tests narrowly defined and, in large part, to all psychological measurements.

I have pointed out elsewhere (15) the proliferation of factors from Spearman to Guilford and have interpreted this trend in terms of test construction practices, the multiplication of number of tests, the disregard for correlated factors, etc., rather than in terms of a closer and closer approximation to the nature of reality. I have also supported as a convenient descriptive device the use of a hierarchical model of factors of human ability to encompass the varying degrees of generality of factors from Spearman, through Thurstone, to Guilford, though admitting that the hierarchical model does not well describe the complexity of psychological tests suggested by factor analysis. Further analysis along these lines has confirmed me in the conviction that factor analysis of most psychological data can be only descriptive and that we must be very cautious about placing unitary traits within the organism.

I start with the nature of the psychological test. All tests measure performance. Practically all human performances require a learning base. Practically all human performances also require an incentive. While the latter is frequently bothersome in the test situation, the former is always with us, i.e., the nature of the organism and its experiential history are inextricably bound together. There are no cultures for psychological tests.

Secondly, psychological tests are typically composed of multiple items. The test constructor defines the total score on the test by deciding which items to score and which weights to assign to scorable items. These criteria are essentially arbitrary and subjective although there are objective statistical aids (e.g., item analysis) for the decisions. To require each test to scale in the Guttman sense (16) is not a viable solution since psychological data scale only when a given item is repeated with minute variations throughout a particular test.

Factor analysis is also a highly subjective procedure. The number of factors to retain and the rotation of factors have subjective elements that match the subjectivity in factor interpretation (17). True, factor analytic procedures can reveal fundamental physical dimensions from the intercorrelations of box measurements (18) or from the intercorrelations of the behavior of balls (19), but psychological data are not physical data. Since we measure performance which is typically assessed by a number of items arbitrarily grouped together, we could not tie our factors to traits even with a completely objective factoring technique. Our tests sample behavior, but we have no definition of the population or of the sampling method used. The varying degrees of item intercorrelations and/or the varying degrees of test intercorrelations allow factors to be extracted in an almost limitless number of orders. The only possible status of factors under these circumstances is as descriptive, summarizing concepts.

I can only dimly, and briefly, speculate on an alternative to trait theory. It will probably be an S-R theory with molar definitions of S and R. For prediction purposes, we shall have to introduce measurement of situations, as well as response measures (tests), into our equations. Predictions will be contingent ones,

relative to the assessment of previous situations and of probable intervening situations, as well as to the assessment of the criterion situation.

In the meantime we can make some progress, as mentioned earlier, by forgetting traits and factor-pure measures. An attempt to match the behavior sampled by predictor tests with the behavior we wish to predict is a simple but good starting point. Just remembering that both predictors and criteria sample behavior is an important step forward from the belief or hope that we are really measuring underlying traits. We should also start with a test theory that recognizes that tests are composed of items or, if you prefer, that most psychological measures are tests composed of multiple items. Classical measurement theory is not helpful in this respect at best and for the unsophisticated can actually be misleading. Internal consistency item analysis must be used with great caution also. We can write better items, we can package items in more efficient form for prediction purposes, and we can combine test information in better ways. I have not the slightest doubt, for example, that a test designed in accordance with something like a factor-type analysis will be more valid for prediction purposes than one designed in accordance with homogeneity theory. I am also confident that selection for retention is more rewarding than selection for short-range performing in training. If we find, as we might, low or even small negative correlations between retention predictions and training predictions, I would not feel alarmed. We should instead take a hard look at training content and methodology.

Let us not reify test scores and make predictions by speculation. Let us not try to sweep our inadequacies in long-range prediction and on-the-job prediction under the rug. Instead, let us stick close to our data and obtain more and better data. An empirical point of view, a good technology, an adequate number of examinees for tryout purposes, and a good record system to provide for followup will carry us a long way. The prospect is not exciting, but this is where I place my bets. In summary, let us be good engineers.

REFERENCES

1. **Travers, R.M.W.** Personnel selection and classification research as a laboratory science. *Educ. psychol. Measmt.*, 1956, 16, 195-208.
2. **Dulkois, P.H.** The classification program. *AAF Aviation Psychology Research Report* #2. Washington: U. S. Government Printing Office, 1947, 72-90.
3. **Melton, A.W.** Apparatus tests. *AAF Aviation Psychology Personnel Report* #4. Washington: U.S. Government Printing Office, 1947, 79-85 and 400-404.
4. **Guilford, J.P. & Lacey, J.I.** Printed Classification Tests. *AAF Aviation Psychology Research Report* #5. Washington: U.S. Government Printing Office, 1947, 479-486 and 350-369.
5. **Brown, C.W. & Ghiselli, E.E.** Some generalizations concerning the validity of aptitude tests. *Personnel Psychol.*, 1953, 6, 139-150.
6. **Thorndike, R.L. & Hagen, E.** *Ten thousand careers.* New York: John Wiley & Sons, 1959.
7. **Anderson, J.E.** The limitations of infant and preschool tests in the measurement of intelligence. *J. Psychol.*, 1939, 8, 351-379.
8. **Juola, A.E.** Prediction of successive terms performance in college from tests and grades. *Amer. Educ. Res. J.*, 1966, 3, 191-197.

9. **Humphreys, L.G.** Investigations of the simplex. *Psychometrika*, 1960, 25, 4, 313-323.

10. **Bruckner, W.F., Rothney, J.W.M., & Shuttleworth, F.K.** Data on the growth of public school children. *Monog. Soc. Res. Child Develpm.* III, No. 1, 1938.

11. **Fleishman, E.A. & Hempel, W.E., Jr.** Changes in factor structure of a complex psychomotor test as a function of practice. *Psychometrika*, 1954, 19, 239-252.

12. **Becker, W.C.** The relationship of factors in parental ratings of self and each other to the behavior of kindergarten children as rated by mothers, fathers, and teachers. *J. consult. Psychol.*, 1960, 24, No. 6, 507-527.

13. **Becker, W.C., Peterson, D.R., Luria, Zella, Shoemaker, D.J., & Hellmer, L.A.** Relations of factors derived from parent-interview ratings to behavior problems of five-year olds. *Child develpm.*, 1962, 33, 509-535.

14. **Peterson, D.R.** The clinical study of social behavior. Unpublished manuscript.

15. **Humphreys, L.G.** The organization of human abilities. *Amer. Psychol.*, 1962, 17, No. 7, 475-483.

16. **Guttman, L.** A basis for scaling qualitative data. *Amer. Soc. Rev.*, 1944, 9, 139-150.

17. **Humphreys, L.G.** Psychological interpretations of factors. *Int. Encyclopedia of the Soc. Sci.* (In Press.)

18. **Thurstone, L.L.** *Multiple factor analysis*. Chicago: Univ. of Chicago Press, 1947. 140-147.

19. **Cattell, R.B. & Dickman, K.** A dynamic model of physical influences demonstrating the necessity of oblique simple structure. *Psychol. Bull.*, 1962, 59, 389-400.

VIII

PERSONNEL RESEARCH FOR THE GREAT SOCIETY

John T. Dailey
The George Washington University

At the first meeting of the President's Committee on Mental Retardation in July 1966, its chairman, John Gardner, Secretary of the Department of Health, Education, and Welfare, told the Committee that its primary function should be "to find ways to enable a person to be whatever he is capable of being – to provide him the opportunities to reach his potential." The Committee's concern is not just with the mental retardates in the lower 3 per cent; rather, it is supposed to be concerned with the whole spectrum of abilities from the point of view of enabling everyone to reach his full potential.

From the point of view of personnel psychology, this is probably an ideal definition of the Great Society and what it is all about. The Great Society will be that society where every individual has the opportunity to reach his potential, whatever that might be. One of the primary concerns of the personnel psychologists in the fu[illegible]should be devising new and better ways of assessing the potentials of individuals from minority or disad[illegible]d groups.

One of the greatest problems in counseling and educating the disadvantaged is that existing talent tests do not identify latent talent that a student's environment has not permitted to develop. They measure the skills that he has developed but not the skills he might have developed in a more favorable environment. New approaches are needed to the prob'em of identification and development of the talents of the disadvantaged, and such approaches are now under way at The George Washington University.

Teaching Basic Vocational Talents

If our low-income youth are to have a full chance to emerge from poverty, the quality of their education must be massively upgraded. This means more than just keeping the students in school another year or two. It means having them learn more and develop their basic educational and communication skills to a higher level. It was found in the Project Talent study (1) that many of the youth from deprived backgrounds lack the basic skills and aptitudes necessary for many of our more highly skilled jobs which are most in demand due to the increasing pace of automation.

The George Washington University has developed a new curriculum and materials for raising the vocational talents of youth from deprived backgrounds. This work was carried out under a contract with the U. S. Office of Education, entitled "Development of a Curriculum and Materials for Teaching Basic Vocational Talents." The materials developed were tried out and evaluated in a representative sample of the nation's schools in eight school systems (San Antonio, Texas; Atlanta, Georgia; Wise County, Virginia; Washington, D.C.; New York City; Bayonne, New Jersey; Erie, Pennsylvania; and Detroit, Michigan).

The fundamental new approach on which this project was based is actually to train students in the skills of nonverbal abstract reasoning, mechanical reasoning, and spatial visualization. These types of skills have been found to be fundamental to the understanding of the technical training required to meet the demands of industry and military services for technically trained personnel. By direct training on material similar to that used in standardized tests of abstract reasoning, mechanical reasoning, mechanical and electrical information, and spatial visualization, the basic understandings and skills of the students in these areas can be increased.

In addition to the basic aptitude practice exercises, two other materials were used in this special curriculum. One is a series of basic remedial-type readers which have been developed. The readers were designed for mature eighth graders who read at the sixth-grade level or below. They are profusely illustrated and designed to fit in with the backgrounds of disadvantaged students. They provide reading material at a low-difficulty and high-maturity level, with content that is directly related to occupations and adult living. The titles of the readers are:

Transportation Long Ago
Transportation Today and Tomorrow
The Automobile
Occupations for You
Tools and Basic Machines

The other part of the curriculum consists of a laboratory course with equipment and simple demonstration devices for teaching students those aspects of mechanical ability and basic mechanical and technical comprehension that are not suitable for teaching by readers, paper and pencil, or filmed exercises.

On the posttesting at the end of the school year, substantial increases in these scores were found for the talent areas where instruction was given, but only normal growth or less was found for the areas not taught. Abstract Reasoning, Two-Dimensional Spatial Reasoning, and Three-Dimensional Spatial Reasoning were found to be especially trainable. The data confirm that important changes in vocational talent test performance can be effected by the curriculum materials that were developed. It is planned to follow the students into vocational training to determine how the increased level of skill on the tests is reflected in greater trainability.

As the development of the new curriculum and materials for teaching basic vocational talents in the eighth and ninth grades progressed, a number of additional ways of using the materials have been indicated. Among these are the following areas:

1. Use in adult basic education or literacy training.

2. Use by the Armed Forces.

3. Use in short concentrated courses to assess how well students learn various technological skills and concepts and how well they like learning them. On the basis of this, they can be guided into appropriate long-range training courses.

4. Use in special summer programs, for which such concentrated programs are appropriate. These would assess students, sharpen their talents, and guide them into vocational training programs.

The primary purpose would be to give students a chance to learn the basic talents underlying success in training for technological careers. On the basis of how well they learn the various basic technological skills and concepts and how well they enjoy learning them, they could be guided into existing high school,

junior college, or other post-high-school programs for technological careers. This represents a new approach to the problem of identifying and developing the latent talents of students from a wide variety of backgrounds, who may not have had an adequate chance to learn the skills and concepts sampled by existing occupational aptitude tests. In effect, such a program could serve as an aptitude test and also an interest test.

It is hoped that the new curriculum and materials will enable schools to strengthen the basic vocational talents of their students so that more of them can develop the basic skills essential for emergence from poverty.

Influence of Training on the Factorial Structure of Talent Tests

Mechanical talent or aptitude appears to be a skill largely learned indirectly through a variety of out-of-school experiences. A rural or small-town environment is particularly rich in such experiences; mechanical comprehension may have been well named "Barnyard Physics." One objective of the project is to determine how young people acquire their mechanical comprehension and to develop and evaluate materials to substitute for the lack of environmental stimulation in mechanics and technology that handicaps most of our young people today.

In order to study the relationship between skill levels in abstract reasoning versus mechanical reasoning, the test scores for these two variables for each school were plotted for eighth- and ninth-grade boys and girls in the eight school systems studied. It was found that all of the groups showed under-performance in mechanical reasoning as compared with abstract reasoning, except for Wise County (a rural area). The other schools were all serving low-income areas in large cities.

One approach to this was a series of factor analyses designed to shed light on how environmental influences and other formal training affect the factor content of the tests. The first of these studies has shown large differences in the factor patterns of the fourteen pretest subscores for the eighth-grade boys in rural and urban schools. The same forces that cause over- and under-achievement on mechanical tests also seem to operate to change the factor content of tests.

Table 1 shows correlations between the tests and reading comprehension for each of five school systems, for all Project Talent fifteen-year-old boys, and for Project Talent fifteen-year-old boys not in high school. As can be seen, the patterns of correlations vary markedly for the various groups. The means, standard deviations, and rotated (varimax) factor loadings for the five school systems are shown in Tables 2, 3, 4, 5, and 6. It can be seen from these tables that the five groups vary enormously in mean scores on some of the tests. A clearly defined verbal factor appeared in four of the five schools, but no other factor appeared in more than two groups. The factors that appeared were, for the most part, the ones to be expected from previous experience to be found in a matrix of such well-known types of tests.

Each of the five groups of schools had a factor pattern that was different from the results of a similar factor analysis of the correlations of the same tests based on the Project Talent sample (2) of 3,656 fifteen-year-old boys (in and out of school). In this analysis four unrotated factors accounted for 76 per cent of the variance. When rotated (varimax), these factors appeared to be Technical Information, Three-Dimensional Spatial Reasoning, Verbal Reasoning, and Two-Dimensional Spatial Reasoning.

It would appear that correlation patterns based on a nationally representative sample may not be at all representative of the students in many of the individual schools involved. Further research is needed to determine to what extent the interpretation of test interrelationships should be based on data for homogeneous categories of schools rather than all schools combined. These findings are consistent with the

hypothesis that the factor loadings of tests are highly affected by the specific experience patterns of the students who are tested. Current studies are comparing the factor structure of pretest and posttest batteries after exposure to the new curriculum. These studies should yield significant new insights into the development of vocational talents and should be a valuable source of leads in improving the materials that have been developed.

Table 1. **Correlations between the Vocational Talent Test Scores and Reading Comprehension Scores for Five School Systems (1965) and Selected Project Talent Samples (1960)**

	Proj. Talent Ident. No.	At-lanta, Ga.	San Ant., Tex.	Wash D.C.	Wise Co., Va.	Erie, Pa.	15-year-old Males	
							All[1]	Not in High School[2]
Abstract Reasoning	R-290	.08	.16	.12	.16	.38	.58	.40
Mechanical Reasoning	R-270	.03	.15	.33	.39	.46	.57	.47
Arithmetic Reasoning	R-311	.21	.07	.14	.34	.48	.66	.40
Vis. in 2 Dimensions	R-281	.38	.05	.30	.19	.29	.39	.28
Vis. in 3 Dimensions	R-282	.20	.06	.12	.14	.37	.48	.40
Vocabulary Info.	R-102	.46	.38	.39	.54	.70	.77	.61
Math Info.	R-106	.36	.08	.25	.34	.52	.67	.32
Physical Sci. Info.	R-107	.62	.21	.40	.27	.65	.66	.49
Biological Sci. Info.	R-108	.16	.25	.34	.33	.63	.63	.47
Aero. & Space Info.	R-110	.32	.16	.16	.31	.62	.60	.40
Elec. & Elec. Info.	R-111	.37	.15	.09	.13	.49	.54	.44
Mechanics Info.	R-112	.33	.28	.25	.34	.47	.56	.49

[1] From Table V-2c, pages 90-93, Project Talent Monograph: *Studies of a Complete Age Group - Age 15.*
[2] From Table V-2b, pages 86-89, *Ibid.*

Table 2. **Factor Analysis of Vocational Talent Pretests for School System A, Atlanta, Georgia**

(N = 60)

	Factors									
	1	2	3	4	5	6	7	8	M	SD
Abstract Reasoning	.04	.04	.04	-.00	.96	-.02	-.04	.00	3.78	2.22
Mechanical Reasoning	.07	-.02	.97	00	.05	.11	-.02	-.04	4.93	2.36
Arithmetic Reasoning	.26	-.50	.21	.54	-.19	.17	-.26	.08	2.47	1.84
Vis. in 2 Dimensions	.87	.19	.13	-.02	-.02	.00	.02	-.06	6.68	4.93
Vis. in 3 Dimensions	.31	.83	-.00	.13	.01	.06	-.21	-.03	5.07	2.07
Reading Comprehension	.61	.04	-.11	.18	.12	.46	.22	.23	9.23	4.24
Vocabulary Info.	.24	-.14	-.12	.61	.30	.42	.15	.06	4.02	2.66
Math Info.	.17	-.15	.03	-.18	-.10	.77	-.08	.27	2.92	1.89
Physical Sci. Info.	.30	-.13	.06	.44	.09	.60	.17	.14	2.55	1.74
Biological Sci. Info.	-.10	.22	-.00	.82	-.03	.01	.10	.19	1.85	1.44
Aero. & Space Info.	.11	-.12	-.04	.11	-.04	.13	.90	.14	1.67	1.20
Elec. & Elec. Info.	.03	-.04	-.08	.20	-.01	.22	.18	.89	2.38	1.45
Mechanics Info.	-.15	.25	.12	.31	.01	.78	.13	-.00	3.08	1.84

Table 3. Factor Analysis of Vocational Talent Pretests for School System B, San Antonio, Texas
(N = 136)

	Factors									
	1	2	3	4	5	6	7	8	M	SD
Abstract Reasoning	.17	.27	-.02	.06	-.05	-.00	-.07	-.87	5.88	2.61
Mechanical Reasoning	.09	.73	.11	.07	.18	-.08	.11	-.26	6.33	2.74
Arithmetic Reasoning	.01	.01	.92	.04	-.05	.06	-.07	.01	3.58	1.74
Vis. in 2 Dimensions	-.00	-.13	.09	.93	.00	-.02	.01	-.10	7.90	4.90
Vis. in 3 Dimensions	-.02	.83	-.08	.13	-.13	.06	.06	-.03	6.11	2.63
Reading Comprehension	.78	.01	.24	.00	-.00	-.12	.23	-.08	14.24	4.83
Vocabulary Info.	.66	-.00	-.10	.02	-.12	.22	.19	-.29	5.48	2.49
Math Info.	.08	-.00	.06	.02	-.02	.96	.07	-.02	4.16	1.76
Physical Sci. Info.	.53	.10	-.32	.44	-.12	.24	-.06	.22	4.85	2.18
Biological Sci. Info.	.54	.43	-.27	.03	-.20	.12	-.06	.00	3.15	1.65
Aero. & Space Info	.03	-.18	-.16	.12	-.62	.07	.50	-.27	2.59	1.45
Elec. & Elec. Info.	.16	.12	.11	.03	-.90	.01	-.07	.03	3.65	1.87
Mechanics Info.	.29	.17	-.08	.02	-.01	.08	.84	.10	5.47	2.03

Table 4. Factor Analysis of Vocational Talent Pretests for School System C, Washington, D.C.
(N = 159)

	Factors									
	1	2	3	4	5	6	7	8	M	SD
Abstract Reasoning	.10	.61	-.10	-.15	.18	.19	-.59	-.05	5.74	2.59
Mechanical Reasoning	-.06	.17	.14	.14	-.19	.05	-.84	.10	5.96	2.86
Arithmetic Reasoning	.07	-.00	-.01	.12	-.04	.97	-.06	.04	3.33	1.81
Vis. in 2 Dimensions	.01	.14	.08	.90	-.00	.12	-.09	-.04	8.13	5.82
Vis. in 3 Dimensions	-.05	.85	.01	.23	-.22	-.07	-.11	.04	5.51	2.63
Reading Comprehension	.54	-.16	-.06	.46	-.08	-.03	-.49	.12	14.34	6.28
Vocabulary Info.	.66	.12	.01	.20	-.02	.08	.12	.51	5.84	2.89
Math Info.	.27	.12	.06	.02	-.90	.04	-.12	-.03	4.01	2.03
Physical Sci. Info.	.72	.04	.10	-.06	-.30	.04	-.15	.17	4.44	2.41
Biological Sci. Info.	.83	-.04	.23	-.01	-.07	.07	.08	.00	3.45	2.01
Aero. & Space Info.	.33	.13	.82	-.00	-.05	.01	-.02	.01	2.37	1.58
Elec. & Elec. Info.	-.00	-.13	.86	.08	-.05	.02	-.05	.23	4.22	2.97
Mechanics Info.	.21	-.12	.26	-.07	.01	.02	-.15	.88	5.44	2.42

Table 5. Factor Analysis of Vocational Talent Pretests for School System D, Wise County, Virginia

(N = 112)

	Factors									
	1	2	3	4	5	6	7	8	M	SD
Abstract Reasoning	-.14	.81	.15	-.21	.02	.03	.11	-.20	6.78	2.64
Mechanical Reasoning	.43	.52	.11	-.22	.30	.32	-.05	-.25	8.77	3.34
Arithmetic Reasoning	.09	.10	.92	-.03	.19	.05	.02	-.12	5.27	1.94
Vis. in 2 Dimensions	.09	.16	.05	-.02	.06	.96	.09	-.02	9.88	5.06
Vis. in 3 Dimensions	.29	.79	-.03	.08	.05	.16	.16	.00	6.28	2.68
Reading Comprehension	.07	.06	.19	-.19	.86	.10	.18	-.09	18.65	6.92
Vocabulary Info.	.25	.13	.31	-.06	.48	.00	.59	-.06	8.49	3.02
Math Info.	.08	.09	.03	-.91	.19	.00	.12	-.09	5.92	2.25
Physical Sci. Info.	.30	.19	.24	-.37	-.16	.06	.65	-.14	7.30	2.78
Biological Sci. Info.	.11	.10	-.20	.01	.28	.15	.78	-.22	5.26	2.18
Aero. & Space Info.	.21	.20	.17	-.14	.11	.03	.24	-.86	3.60	1.81
Elec. & Elec. Info.	.81	.18	.11	.05	-.04	.19	.11	-.24	5.62	2.48
Mechanics Info.	.73	-.02	.06	-.26	.23	-.04	.36	.01	9.24	3.15

Table 6. Factor Analysis of Vocational Talent Pretests for School System E, Erie, Pennsylvania

(N = 169)

	Factors									
	1	2	3	4	5	6	7	8	M	SD
Abstract Reasoning	.17	.17	.08	-.91	.10	-.11	.12	.04	8.08	2.64
Mechanical Reasoning	.27	.20	.15	-.38	.16	-.21	.66	.25	9.14	3.80
Arithmetic Reasoning	.21	.06	.95	-.08	.02	-.10	.09	.08	5.02	2.41
Vis. in 2 Dimensions	.14	.95	.08	-.17	.14	-.08	.11	.08	11.97	5.27
Vis. in 3 Dimensions	.22	.14	.05	-.12	.94	-.06	.08	.13	7.03	2.67
Reading Comprehension	.72	.10	.34	-.23	.14	-.24	-.06	.13	18.26	8.59
Vocabulary Info.	.80	.09	.10	-.16	.06	-.11	.12	.29	9.82	3.84
Math Info.	.30	.06	.13	-.15	.06	-.90	.10	.12	4.59	2.49
Physical Sci. Info.	.80	.06	.09	-.16	.13	-.10	.11	.24	7.01	3.53
Biological Sci. Info.	.61	.01	.10	-.28	.14	-.24	-.35	.40	5.00	2.28
Aero. & Space Info.	.74	.12	.10	.10	.18	-.25	.26	.20	3.69	2.15
Elec. & Elec. Info.	.50	.14	.09	-.01	.12	-.02	.32	.67	7.57	3.65
Mechanics Info.	.33	.05	.03	-.06	.08	-.13	.02	.87	8.84	3.02

Measurement of Language Facility

A second major new approach to assessing the potentials of individuals from minority or disadvantaged groups has been the Dailey Language Facility Test.

Much recent research has indicated that education in basic communication skills is essential for emergence from poverty (3). Students who learn well in schools have been found to be much more apt to enter professions or occupations with a future. Those who do not respond to schooling and who drop out have great difficulty in occupational adjustment. It has been found that the key to school success is adequate use of language, and deficiencies in this are traceable to preschool years and are closely related to immediate family environment. Such studies as Project Talent (1) have indicated that basic learning at the high school level is largely a function of the basic communication skills of the students as they enter the ninth grade and not a result of the type of schooling received in high school. Accordingly, more and more attention has been focused recently on preschool programs as a method of making up for lack of a stimulating family environment. This is now being brought to a peak by Project Head Start. The fundamental objective of preschool programs should be to foster the language development of the child. If he is able to acquire basic oral communication skills by age six, he should then be able to follow through in the primary grades and acquire adequate skill in reading and written use of language.

In order to evaluate preschool programs, a test of the ability to use oral language is needed. Such a test must be individually administered, should take no more than ten minutes, and should be one which the average teacher can administer and score. It should provide a measure of language facility which is independent of vocabulary, information, pronunciation, and grammar. The Dailey Language Facility Test, which was designed to these specifications, was originally developed for use in evaluating preschool methods and programs. It has been used as part of the evaluation of the Language Arts Program of the elementary schools in the District of Columbia. It is being used to evaluate Head Start and other preschool programs, and has been or is being used in a number of evaluation studies in Washington, D.C. and in a number of other school systems in grades one to nine. It has also been administered to several groups in the Job Corps and similar training programs, to samples in adult basic education programs, and to samples in schools or programs for the deaf, physically handicapped, mentally retarded, delinquent, emotionally disturbed, and musically talented, and in tutorial programs.

In effect, this test gives a measure of how well a person can use the language or dialect to which he has been exposed in either his home environment or school. The test obtains a standardized sample of speech in ten minutes or less by means of having the subjects tell stories about or describe a series of pictures. Responses to each picture are assigned scores on a 9-point scale according to detailed scoring criteria and examples at each level.

The test can be administered and scored by personnel with a minimum of training. Two different scoring systems have been developed. The first gives a 9-point scale measure of how well the individual is able to conceptualize and communicate in his chosen language. This score is independent of standard English vocabulary, enunciation, information, or grammatical exactness. The second scoring system codes the errors or deviations from standard English pronunciation or usage and gives a diagnostic profile of ability to speak standard English. The nine steps of the basic scale represent nine levels of maturity through which each individual will progress as he grows from infancy to full language maturity.

The average three-year-old child will be able to respond to these pictures only in single words. By age five he will be able to respond in single sentences. By age nine he will be able to give a partial description of both objects and action. At about age eleven he will be giving complete descriptions. By age fifteen he should be adding interpretations or deductions about the pictures and organizing them into stories. They may often involve well-known fictional or historical characters or be tied into current events or recent

historical happenings. Large differences will be found in the levels of maturity among students of the same age. The more mature four-year-olds will be functioning at the six-year-old level, while some six-year-olds will be at the four-year-old level. An occasional four-year-old (even from an extremely disadvantaged environment) will be found to be functioning at age eleven level or higher.

It is possible for a young child to have a very high ratio between his language facility age equivalent and his chronological age since he is not handicapped by lack of experience in competing with much older students as he is on the usual intelligence test. Cases have been found where three-year-old children have the basic language facility level of a twelve-year-old or older. Throughout most of the range of the test, students scoring at the 90th percentile have a language facility age equivalent that is about 50 per cent greater than their chronological age. Mentally retarded subjects aged ten or older will usually be performing below the age six level. This scale is designed to be a measure of how well an individual uses language the way he has been hearing it. It is interesting that individuals in extremely disadvantaged situations have been found to differ a great deal in the level at which they can use the language to which they have been exposed.

In studies at the high school level by Project Talent and others, it has been found that a very fundamental factor or ability involved in success in school work, assuming that motivation is constant, is the ability to acquire, retain, and manipulate symbolic information. This is usually best measured in the later grades by tests of reading comprehension or vocabulary or tests of arithmetic reasoning with verbally presented problems at about the seventh-grade level. The basic theory and rationale for the Dailey Language Facility Test is to extend the range of this type of measure downward to assess at early ages the characteristics of the child who has the capacity to develop into one with a higher verbal IQ than another. It is hypothesized that the primary characteristic which will differentiate between two such children will be the basic level of language development.

Adequately stable measures of this language facility can be obtained from most children by age three and in many as early as age two. Mere precocity in very early speech, however, should not be given undue weight in interpretation. Those rare cases of delayed speech that later develop well will not be validly measured until the full onset of speech occurs.

Such tests as the Wechsler and the Binet have been found to have poor ability at early ages to predict verbal IQ at maturity (4). The correlations between IQ at age four and age twelve have been unsatisfactory. This has led to considerable confusion on the part of both psychologists and parents. The root of this difficulty has been that such IQ tests are a combination of a large number of separate tests, some of which are communications and language oriented, and some of which are perceptual-motor in nature. Much research at the high school or young adult level indicates that language facility aptitude is substantially independent of many other perceptual and motor factors. This shows up in differential aptitude test batteries. It also shows up in factor analyses of the various subtests of IQ tests if they are like the Binet or Wechsler. Because of the age placement of the various tests, the Binet or Wechsler type tests are heavily nonverbal at the youngest age levels and heavily verbal at the more advanced age levels. This means that the test is really measuring different things in adolescence than it is at the preschool level. The Dailey Language Facility Test is designed to measure language facility at all ages through from about three to maturity. It is theorized that the youngster who is going to have extreme difficulty in basic communication skills and be severely educationally retarded in grade six is likely to have had inadequate language development at an earlier preschool age. In a sense, the Language Facility Test may be regarded as a test of the ability to learn to read at a later age. A picture represents symbolic information in the same way as do numbers and words.

Language is an incredibly complex coding system. This is equally true of oral speech or written text. Facility in oral speech is assumed to require much the same aptitudes as facility in the use of written language. It is assumed that increasing maturity in oral language facility represents an increasing ability to

encode, decode, and communicate symbolic information of an increasing complexity. This same ability is central to the problem of learning to comprehend written text.

The test is suitable for administration by anyone experienced in dealing with children or students and eliciting rapport in an interview. No prior training is required, and the test can be given and scored after careful study of the manual. The test is untimed, and the student should be given as much time as necessary to collect his thoughts and tell the story. It is rare that the three pictures will take more than ten minutes total for a given student. It was desired to have pictures that could be meaningful to students from a great variety of backgrounds. Many pictures were tried out before deciding on the twelve plates that are used. One series is of actual photographs of preschool students and teachers in a school for children of migrant farm workers. Another series was drawn especially for the test, and these plates were purposely made stark and free of background detail. To give greater variety, the third series consists of scenes by old Spanish masters and contains a maximum amount of shading and detail. The plates are cycled to give three alternate forms (Plates I, II, and III; IV, V, and VI; VII, VIII, and IX) and three extra plates (X, XI, and XII) for use as alternates. Statistical analyses have demonstrated that the different pictures give highly similar results and measure the same thing. Their factorial content, means, and standard deviations are very similar. As a result, all can be scored by the same general criteria.

The directions for administration are as follows:

1. "What is your name?"
2. "How old are you?"
3. "Do you like to hear stories?"
4. "Could you tell me a story about a picture if I show it to you?"
5. "Good. Tell me a story about this picture."
6. "Tell me what you see in the picture."
7. "What are they doing in the picture?"
8. "What else can you tell me about the picture?"

Scoring Method I – Level of Language Maturity

The response to each picture should be scored as follows:

9.....A well-organized story with imagination and creativity. Need not be original. May use well-known fictional or historical characters.

8.....A complete story that is not well organized.

7.....A story with an interpretation of some elements of action or intentions, as deduced from or suggested by the picture.

6.....A detailed description of what is happening, but nothing about past or future action or intentions. At level 6 all or nearly all of the elements of the picture will be covered, in contrast to level 5 where only selected elements will be covered.

5.....A partial description consisting of two or more sentences with some description of movement or action as seen in the picture.

4.....Two or more sentences describing persons or objects but no verb of action or indication of interaction between a person and an object.

3.....A complete sentence that makes sense.

2.....Compound responses, two or more words at a time, a single word describing action, or more than one single-noun response.

1.....One single-noun response.

0.....No response, garbled speech, or only pointing at picture.

If the child does not respond at all to a picture or says he can't see it, he should be given one of the alternate pictures (Plates X, XI, or XII) and scored on it instead of the original picture. If six pictures fail to elicit any response, other evidence from teachers or parents will need to be appraised to determine whether the lack of response is because of excessive shyness. This is extremely rare, especially among urban children.

Each picture should be scored according to the above scale. The total score for the test is the sum of the scores on the three pictures used.

Levels 0, 1, and 2. If the subject fails to make any response whatever, or speaks in an incomprehensible manner, or simply points at the picture, the score would be 0. At level 1, the subject gives a single-noun response. Only one word is used. To receive a 2, the subject gives a compound response of two or more words, or employs a single word of action, or uses a phrase. However, at level 2, the subject need not use a complete sentence.

Level 3. If the subject uses any complete sentence, he should be given credit at level 3, even if it is "I can't tell you anything," or "That's all." However, it must be a complete sentence. If auxiliary verbs are omitted ("He riding," or "She sitting"), credit should not be given for a sentence. Sentences may be grossly ungrammatical as long as they are complete ("Dis am a hoss," "He are a man," or "They is on bed."). The ability to use a complete sentence is a most important milestone in growth in language maturity. Up to one-third of low-income Negro first-grade children will score less than 9 for three pictures. However, by grade three nearly all such students will be responding consistently to the pictures using complete sentences. Of one group of 689 low-income Negro urban nine-year-old children, only five failed to respond consistently in complete sentences.

Levels 4 and 5. Description of movement or action is used here in the sense that something is said about the movement or action seen in the picture. At level 4, the subject only describes the objects in the picture.

Levels 5 and 6. Both of these levels should include a description of the objects and action in the picture being scored, the difference being that the description at level 5 will be only a partial one. The level 6 description should be complete or very nearly complete and cover the major elements of the picture.

Levels 7, 8, and 9. The essential element here is the existence of a "story." This is defined as something made up about the content of the picture wherein the picture in effect becomes an illustration for the story. At level 7, the story should include some past or future action or intentions suggested by the picture but not shown in it. The level 7 response will often include considerable description and may not be a complete story. However, something must be "made up" about the picture. Level 9 should contain all the major elements of interpretation of objects and for setting out past or future action or intentions. It should show imagination and creativity, but it can include well-known fictional or historical characters and need not be original. The story should hang together and not ramble. Everything in it should tie in to the picture and be deduced from it. It need not be very long. Level 8 should be reserved for stories that do not qualify for the top score because of weak organization.

It should always be remembered that vocabulary, grammar, and pronunciation should not be considered when scoring the scale level. A level 9 story could be grossly ungrammatical and from a person not familiar with the correct names for some elements of the pictures. The scale measures facility in the use of the person's own language or dialect and not in standard English. The scale values are completely independent of vocabulary, information, and grammar. Length of response is not a factor at level 6 and above. The scale levels recapitulate the chronological development of language facility. Independent scorings by clerks with two hours of training correlate in the vicinity of .90 for a total score based on three pictures.

Scoring Method II – Measuring Deviations from Standard English

The samples of speech can also be scored to give diagnostic profiles of frequency of common deviations from standard English. A preliminary 24-category code was developed for this purpose and was applied to a group of 630 students four years after they had started to school together in kindergarten. Part of the group were in seven elementary schools in a special language arts development program and the others were in eight similar schools. All the students were from low-income Negro neighborhoods. Clerks were trained to listen to the tape recordings and tally the deviations as they occurred. Table 7 shows the code used for this purpose.

Table 7. Scoring Deviations from Standard English - Scoring Method II for the Dailey Language Facility Test

	Errors	Examples of Error	
A	Simple verb, wrong number	she want; they sees	A
B	Auxiliary verb, wrong number	he have waited; she are going	B
C	Auxiliary verb omitted	he running	C
D	Wrong past participle	wore (worn); came (come); flew (flown)	D
E	*s* on plural not ending in *s*	chilluns (children); geeses	E
F	Incorrect irregular plural	shelfs	F
G	*a* for *an*	-	G
H	*got* for *have* or *has*	-	H
I	Letters interchanged (*t* for *d*, *d* for *t*)	boddle (bottle); laty (lady)	I
J	*g* on *ing* pronounced	-	J
K	*in'* for *ing*	runnin'	K
L	*picture* mispronounced	pitcher	L
M	Consonants slurred	chillun (children)	M
N	Unaccented vowel slurred	fam'ly; an'mal	N
O	Verb tense changed in sentence	She is getting up and then she got dressed.	O
P	Number of verb agreeing with incorrect subject	The duck and the gull is flying.	P
Q	*they* for *there* or *their*	they shoes	Q
R	*d*, *t*, or *v* for *th*	nuttin' (nothing); muddah (mother)	R
S	*s* on possessive noun omitted	lady' watch	S
T	*r*, *l* omitted	litta gir'; gi'l	T
U	*ɛ* for *ā*, *ûr* for *âr*, *ôr*	bleck (black)	U
V	dipthongized vowels	bayid (bed)	V
W	Elongated, distorted vowels	tahde (tired); bade (bed)	W
X	Other comments (please specify):		X

The students in the Language Arts Program schools made significantly fewer errors in their speech samples than did the control group. This was true at the level of language facility (high, medium, and low) as measured by the 9-point scale on the test. When the experimental and control groups were compared on the incidence of eight major categories of speech errors at each of three levels of language facility, the experimental group did better in 18 out of 24 comparisons. This is significant at the 1 per cent level.

A factor analysis was carried out relating the regular language facility scale score on the tests and the various error scores to such variables as subscores on the Metropolitan Readiness Test, the Peabody Picture Vocabulary Test, subscores on the Stanford Achievement Test, teacher ratings on quality of homework, motivation, ability to get along with other children, and effect of home situation, as well as grade, sex, and age. Three varimax rotated factors emerged from the speech sample: One was defined by the 9-point language facility scale score, one appeared to be a cluster of errors in pronunciation, and the third appeared to be a cluster of errors in structure or grammar. With further refinement the error code should become quite useful in measuring different aspects of dialects and for evaluating various methods of dialect transformation.

Longitudinal studies are under way to relate language facility scores to later school success, to later verbal IQ, and to later performance in verbal facility. In the meantime, the following evidences of validity have been found:

1. The test shows a sharp gradient with chronological age and growth in ability to score well on the test; this coincides with results of other studies of the verbal performance of children of different ages.

2. Intact grade groups of low-income-area children show a distribution of language facility scores that is approximately normal and little different from those of children in general. However, groups of students who have not done well in academic work in school do poorly on the test. Such groups include Job Corps trainees, students in nonacademic tracks in grade eight, poor readers, and youths who have failed the Armed Forces Qualification Test. Honors track students in low-income areas do well on the test.

3. Mental retardates tend to be very poor in language facility.

4. Special training programs designed to stimulate language and intellectual development in young children have caused important gains in language facility scores of low-income-area children.

5. The test has been found to give apparently valid results when given in several foreign languages with the responses being made in the foreign language. These include Spanish, Greek, and several Asiatic languages, as well as sign language.

Relationships of School Characteristics to School Outcomes

Another important issue in the education and counseling of disadvantaged groups is determining the relationships between what schools do and the results they obtain with students of widely differing socio-economic backgrounds.

When dealing with such a complex set of measures and characteristics as a group of school input and output variables, it is most difficult to make interpretations of relationships between any single school characteristic and school performance. For example, children may read better in urban schools, but this may just reflect greater size of urban schools; or larger schools may just show up better because they spend more or serve higher income areas. No simple comparisons of this sort are likely to be meaningful. To compare school performance with a school characteristic, one must consider up to 50 variables

simultaneously if comparisons are not to be misleading. Multiple factor analysis is one way of making such comparisons.

With Project Talent data several studies have been made of the relationships between many school practices and many school outcomes. It is not possible to obtain absolute proof of cause-and-effect relationships from Project Talent data nor any other similar statistical data based on relating past measures of behavior with each other. For example, it is not possible to prove unequivocally by statistical survey methods that higher teacher salaries cause higher school achievement. On the other hand, it is possible to do analyses that make it appear quite likely that these higher salaries may be one of the most important factors in obtaining higher achievement. At the same time, it is possible to use equivalent methods to indicate that it seems very unlikely that size of school as such is a necessary prerequisite for high achievement because there is no associational evidence at all in this direction when all factors are considered. A given factor is extremely unlikely to be a causative factor if it shows no unique association with measures of school performance in a comprehensive set of data where the other important factors are being held constant.

The Project Talent data to date seem to indicate four school factors closely and uniquely associated with school outcomes (such as achievement and going to college and going to college and staying in school). These factors are teacher salaries, teacher experience, number of books in the school library, and per-pupil expenditure.

It should be cautioned that we cannot conclude with certainty that these factors are causing the differences in school outcomes. It may be that they are caused by some outside factors which are just being mirrored or reflected by these above measures. Nevertheless, their relationship is substantial even after as many as 30 of the most important school and community characteristics have been held constant in mathematical analyses.

On the other hand, we can be much more confident in concluding that many other factors are not likely to be prime causes of school excellence since they do not have any sizable unique correspondence with school outcomes. The data in Project Talent indicate that some school characteristics seem very unlikely to be prime causes of school excellence of output. Among these seem to be school size, average size of classes, age of building, and suburban location.

A similar study was made of the 129 elementary schools in the District of Columbia (5). It was found that the variable which was most closely related to school performance was the median family income for the census tract in which the school was located. The higher the median income, the higher the achievement; and the lower the income, the lower the achievement. This variable stood out above all others. The school performance was measured by the percentage of the fourth-grade students scoring below the national norm in the Reading Test of the Metropolitan Achievement Battery in 1963-64. The second most important factor which contributed to high performance when all other variables were held constant was low rate of non-promotion within schools. In addition, similar analyses were made to predict achievement in Word Knowledge, Arithmetic, and Total Language performance, all of which yielded similar patterns; high median income emerged first, followed to a much lesser degree by low rate of non-promotion within the school.

One of the least useful variables in this study was median years of education of the adult population in the census tract. This is to be expected since many studies, particularly Project Talent, have shown that high schools differ tremendously in their levels of achievement. This makes grade completed a very inaccurate measure of the actual level of education. It is often said that a person has "received" only an eighth-grade, or a fourth-grade (or some other grade level) education. However, an education is not "received" in a simple sense. A person learns in a school within a cultural setting and social environment. Often out-of-schoool family or community factors will almost completely nullify attempts to upgrade the educational skills and attainment of youth in schools in low-income areas.

It is of considerable interest that in this study ethnic data about the composition of the student body or the school staff were not needed to predict school achievement with a high degree of accuracy. The use of the median family income level by itself predicts performance about as well as any combination of all the available socio-economic and educational variables. Such factors as gross expenditure rate per pupil in each specific school and degree of overcrowding in each specific school bore little relationship to school achievement in schools with the same levels of parental income. This was also true of the age of the school building. The school factors most closely related to achievement were the presence or absence of a librarian and participation or non-participation in the Language Arts Program. The general pattern of findings was extremely similar to the patterns in the studies of the national samples of high schools in Project Talent.

The American school system is probably the next most important socializing agent after the family, since it transmits the mores, habits, and aspirations of society as a whole. Whenever the families work in harmony with the schools and the majority components of our society, the schools readily achieve their goals; but whenever they do not, the schools have great difficulty in meeting their goals. Many large city school systems are in trouble today and have difficulty in carrying out effective school programs. This is reflected by inadequate skill development in their students and by high rates of juvenile crime. It appears clear, however, that no single person, group, or policy in individual school systems can be blamed for this, since studies such as Project Talent show the same situation in all large metropolitan school systems. They experience great difficulty with the motivation and development of children from families that do not share in the mores, habits, and aspirations of the predominant groups in American society. They also find it increasingly difficult to finance adequate school programs because city limits tend to be relatively fixed, and more and more of the wealth and leadership of metropolitan areas are located outside the central city. The problem in the schools is not a matter of race or ethnic groups. Whenever the schools are serving children from intact families whose heads are regularly working at semiskilled, skilled, or white-collar jobs, they achieve their goals easily, regardless of the race, creed, or color of their parents.

Many studies have been made of the relationships between family characteristics and school performance, and it has been uniformly indicated that family characteristics are the primary factors determining the success of a given school. The success of a school is largely proportional to the types of families it serves, almost regardless of such factors as school size, age of the school building, amount spent per pupil, overcrowding, or class size.

The troubles and problems of the Washington, D.C. schools are the same as those of other large metropolitan school systems, and these problems spring from the same causes. The school performance and delinquent behavior of children in the Washington schools are a direct reflection of the extent to which families are operating to assimilate their children into modern organized society. In those areas where families are performing this function, including some predominantly Negro neighborhoods, performance in school is good, and juvenile crime is minimal. After differences in the types of families involved have been equated, Washington school performance and juvenile crime rates compare favorably with those in similar large cities.

The results of these studies suggest that a massive improvement in the school performance of the present low-achievement students is not likely to result from any combination of increases in what schools have been doing in the past. We cannot solve our school problems merely by higher salaries, more teacher experience, more books, larger schools, smaller classes, new buildings, or moving slum schools to suburban locations. New and novel methods and approaches will be required, including large-scale team approaches to reach the home environment and parents as well as the schools and children. Ways must be found to upgrade the quality of the low-income home environment, and work is now under way on the development of materials and procedures suitable for use by low-income mothers to stimulate the language development of their children.

These studies that have been reported are not meant to be representative of all the research in the area but do represent approaches by the author to the basic problem of fostering the intellectual development of disadvantaged and low-income groups. It is hoped that such research can make a contribution toward the Great Society, whose goals are for every individual to have an opportunity to develop his potential and to share in its rewards.

REFERENCES

1. **Flanagan, J.C. et al.** *The American high school student,* Technical Report to the U. S. Office of Education, Cooperative Research Project No. 635. Pittsburgh: Project TALENT Office, University of Pittsburgh, 1964.
2. **Shaycoft, M.F., Dailey, J.T., Orr, D.B., Neyman, C.A., Jr., & Sherman, S.E.** *Studies of a complete age group – Age 15.* Technical Report to the U. S. Office of Education, Cooperative Research Project No. 566. Pittsburgh: Project TALENT Office, University of Pittsburgh, 1963.
3. **Dailey, J.T.** Education and emergence from poverty. *J. marriage and the family,* November 1964, Vol. XXVI, No. 4, 430-434.
4. **Bloom, B.S.** *Stability and change in human characteristics,* New York: John Wiley & Sons, 1964, 64.
5. U.S. Office of Education contract OE-C2-6-061811-0575. *Evaluation of the contribution of special programs in the Washington, D.C. schools to the prediction and prevention of Delinquency.* Washington: Geo. Washington University, August 1966.

SESSION CHAIRMEN

COL. FRED E. HOLDREGE
1105th Military Personnel Group
Department of the Air Force

DR. EDWARD H. KEMP
Directorate of Science and Technology
Department of the Air Force

DR. RAYMOND V. BOWERS
Department of Sociology
The University of Arizona

DR. MEREDITH P. CRAWFORD
Human Resources Research Office
The George Washington University

MAJ.GEN. OTIS O. BENSON, JR.
Biosciences and Bioengineering
Southwest Research Institute

DR. CLARK L. HOSMER
Director
Florida State University

IX

THE SELECTION OF PEACE CORPS VOLUNTEERS

Abraham Carp
Peace Corps

The Peace Corps selection process has been in existence for about five and a half years. Several books and about fifty articles have been published concerning the selection process, and about thirty small-scale unpublished studies have been accomplished by the Peace Corps staff. Significant breakthroughs or the discovery of the magic wand have not been the lot of the Peace Corps psychologists.

Most of the well-known structured, objective personality tests and many of the unstructured projective tests have been administered to adequate samples of Volunteers. An omnibus biographical inventory has been administered, item-analyzed, clustered, factored, and all but discarded. Discussions about the unreliability of the criterion, restriction of range, and corrections for attenuation have raged throughout our offices. As always, disappointing correlations in the teens can be boosted into the respectability of the sixties.

Objective tests of ability, however, have proved to have some relevance to overseas performance. Both the General Aptitude Test (a short omnibus type intelligence test) and the Modern Language Aptitude Test have correlations in the low thirties with overseas performance, particularly in countries where language fluency is critical to performance. This is especially impressive since language aptitude scores are utilized in initial assignment resulting in marked restriction of range.

Similarly, peer ratings and selection board evaluations have correlated in the mid to high thirties with both the dichotomous criterion of early attrition and a five-point rating of overseas performance. Unfortunately, multiple correlations including all available variables are not significantly larger than the zero-order correlations.

Having given you the results, let me backtrack and give you a picture of the Peace Corps and its selection process.

Background

The purpose of the Peace Corps as stated in the Peace Corps Act is:

> ". . .to promote world peace and friendship through a Peace Corps, which shall make available to interested countries and areas men and women of the United States qualified for service abroad and willing to serve, under conditions of hardship if necessary, to help the peoples of such countries and areas in meeting their needs for trained manpower, and to help promote a better understanding of the American people on the part of the peoples served and a better understanding of other peoples on the part of the American people."

In order to understand the problems of selection of Peace Corps Volunteers, it is helpful to restate a series of early basic policy decisions regarding the overall Peace Corps program.

1. It was decided that the Peace Corps would go only where invited.

2. Volunteers overseas could work for the host government or for a private agency or organization within the foreign country, serving under host country supervisors, and working with host country co-workers wherever possible.

3. Volunteers would not be "advisers" but "doers."

4. Volunteers would serve for two years, without salary or draft exemption.

5. Volunteers would be provided a living allowance enabling them to live in a modest manner comparable to the circumstances of their co-workers.

6. Volunteers would enjoy no diplomatic privileges or immunities, have no PX or commissary rights, receive no "hardship" or "cost-of-living" allowances and have no vehicles unless needed for their job.

7. Volunteers would learn to speak the language of the host country, learn to appreciate its customs, be able to discuss adequately and intelligently the United States when questioned, refrain from political or religious proselytizing, and set as the standard of their success how well the requested job was fulfilled.

8. A termination allowance of $75 for each month of satisfactory service was established to help the Volunteer get started again in this country.

9. The Peace Corps would be open to all qualified, single Americans above 18 years of age and for married couples with no dependents under 18, where each had a needed skill.

10. A college degree would not be a requirement for service. A special effort would be made to attract farmers and craftsmen who possess skills and experience but no degrees.

11. The highest medical, psychological, and character standards would be established; and final selection would be made at the conclusion of training.

12. The hardships of Peace Corps life would be featured in recruitment so no candidate would misjudge the terms and conditions under which he volunteered to serve.

13. Candidates, trainees, and Volunteers would be told they could resign from the Peace Corps at any time. The Peace Corps wants only those who serve freely, a decision now made each day by each Volunteer.

14. Each Peace Corps program is treated as a separate entity. Only after a firm agreement has been reached between the officials of the United States and host country which specifies the number of Volunteers and the kinds of skills needed in a program, then and only then does the Peace Corps begin to select trainees for that program and to plan a training experience especially designed for the particular country-program. As contrasted with military services, a person does not "enlist" in the Peace Corps; instead he offers to serve; he does not become a part of the organization until he has been offered and has accepted an invitation to train for a specific program. There is nothing in the Peace Corps corresponding to the basic training program of most military services. Although all training programs include certain common elements, each training program is literally tailored to the host country's Volunteer requests.

The Peace Corps has Volunteers in 50 different countries. While almost any single activity will be performed someplace in some country, generally Peace Corps Volunteers serve in teaching, agricultural development, public health, cooperatives, small industries, and community development. Except for most of the teaching programs, reasonable fluency in the host country language is a prerequisite for effective performance.

In order to apply for the Peace Corps, an applicant must fill out a questionnaire covering his academic and work history, marital and military background, extra-curricular and community activities, as well as a medical history. He also indicates his date of availability, his area, country, and job preferences, and takes the Peace Corps Placement Tests.

From the questionnaire, 10 to 20 references are contacted by mail and about 85 per cent are returned. The reference is asked to rate the applicant on a five point scale on job competence, emotional maturity, relationships with other people, and overall competence. About four weeks after the references have been mailed the folder containing the questionnaire, transcripts, and references are transferred to the Classification Branch.

It is in this branch that the skill, training, and other qualifications of applicants are compared to the requirements of programs to be manned by Peace Corps Volunteers. Those applicants possessing the necessary qualifications are invited to enter training for specific Peace Corps projects. The decision to invite an applicant for training is based upon a careful analysis of the applicant and his potential assignment. The aim is to assign each applicant to the job which will best utilize his or her talents and to maximize the overall effectiveness of the Peace Corps.

In performing this function, the staff of the Classification Branch combines a behavioral science background with specialized knowledge of Peace Corps programs throughout the world. The Classification Officers work closely with other members of the Peace Corps staff who are responsible for the development of specific programs and, in addition, periodically visit Peace Corps programs overseas to acquire first-hand knowledge of the work being performed.

In the accomplishment of its mission, the Classification Branch performs four main functions. These are as follows:

1. Identify the kinds of Volunteers needed in each program.

2. Insure the optimum distribution of applicant talent to all Peace Corps programs.

3. Compare applicant talents with program requirements.

4. Invite applicants to participate in training for Peace Corps programs.

Identifying the Kinds of Volunteers Needed. The Classification Officer responsible for manning programs in a particular region assists the Program Development Office in determining the kinds of Volunteers needed for each program. This is based upon information about the country, the environment surrounding the job, and the kind of work to be accomplished. It is the responsibility of the Classification Officers that these requirements be realistic in terms of the work that is to be performed and with the descriptive information available on applicants. That is, the job requirements are translated into appropriate educational level, language aptitude, work experience, etc. For example, if West Pakistan wants Peace Corps Volunteers who can advise on bridge building, it is probable that they need Volunteers with a strong background in civil engineering. If it is further indicated that they must speak the local language fluently,

then aptitude for learning a foreign language becomes an important attribute for success on the job and it must be possessed by potential Volunteers. However, in many cases, the work to be performed can be successfully accomplished by Volunteers with less specialized backgrounds. In such instances, non-specialized liberal arts graduates or equivalents may be considered for assignment, and it is secondary skills gained from summer work, homemaking skills, sports, or farm experiences which are the determiners of probable success on the job.

Insuring Optimum Distribution of Volunteer Talent. While all the applicants who pass the assessment phase of the pre-screening are qualified to be Peace Corps Volunteers if their skills can be used appropriately, they vary in the quality of skills and training they possess. As might be expected, requests for Peace Corps Volunteers exceed the supply of highly qualified applicants. Because of this, one function of the Classification Branch is to distribute the applicants equitably among the many Peace Corps programs so that they may be considered as potential trainees. To do this, the Matrix and Statistics section maintains a continual control over the allotment of applicants for consideration by the several Classification Officers. As invitations are issued and accepted, there are changes in the status and needs of each program. Feedback on these changes are insured by internal controls and adjustments in the distribution of applicants are made accordingly.

Comparing Applicants to Program Requirements. The qualifications of each applicant are considered to determine if he can be invited to train for a specific program. While it may seem an almost impossible task to determine where best, if at all, to assign an applicant among approximately 100 training programs about to start, there are certain factors which reduce the choice to more manageable proportions. Mandatory requirements, such as programs needing only male Volunteers or those with college degrees or with previous teaching experience, are among those factors. In addition, the area preference of the applicant, as well as when he will be available for training, are further considerations which narrow the choice of programs for which an applicant can be considered. Whenever possible, the desires of the applicant are considered, but of course, the qualifications needed for assignment are of prime importance.

After determining the programs for which an applicant is to be considered, the Classification Officer evaluates the individual in terms of the job demands, the kinds of applicants already assigned and future applicants who may be considered for the program. If, after considering all these factors, it is decided that the goals of the Peace Corps can be achieved through the applicant in a particular assignment, an invitation to enter training will be issued.

Inviting Applicants to Enter Training. It is only after consideration of the above points that an applicant is invited to train for a specific program. Unlike a military service which issues "orders to report," the Peace Corps, being a completely voluntary organization, only *invites* applicants to report to training for a specific program. This is probably a very useful aspect of the total selection process. Each invitation to train for a program is accompanied by a Program Brochure describing the country, the culture, the duties of Volunteers, how they will live, with whom and for whom they will work, etc. The applicant is asked whether or not he wishes to train for the particular assignment and is advised that he has the privilege of declining the invitation without prejudicing his being considered for another Peace Corps program assignment.

Self-selection, the decision to accept or decline the invitation to train for a specific program, is an important phase of Peace Corps selection, but it also adds to the complexity of the task. Acceptance rates have varied from a high of 90 per cent to less than 30 per cent for different programs. Since it is impossible to predict in advance what the acceptance rate for a given program will be, it is always necessary to invite considerably more applicants than needed.

An applicant invited to train for a program is reminded that there are still four hurdles between him and overseas service as a Peace Corps Volunteer:

1. He must meet the rather rigorous physical standards established for service.

2. He will be subjected to a full background check of character, reputation, and loyalty.

3. He must maintain an adequate performance during training.

4. He must be judged in training as being the sort of individual who will be able to serve effectively in an overseas assignment while at the same time serving as a good representative of the United States.

Field Selection

The further assessment of applicants who accept invitations to training is the responsibility of the Field Selection Branch, assisted by other staff members of the Peace Corps.

A Field Selection Officer responsible to Peace Corps, Washington, is assigned to each training program and has the responsibility for coordinating the assessment and evaluation of individual trainees. The doctoral degree in clinical, counseling, industrial, personnel, or social psychology is a minimal base for appointment as a Field Selection Officer, and in addition, they are carefully selected for their maturity, judgment, and professional experience in the behavioral sciences.

During the training program, each trainee's suitability for Volunteer service is assessed in a variety of ways, primarily by staff members of the training institution. Responsibility for the collection and development of the assessment data at the training site is assigned to a professional staff member, usually a psychologist, designated as the Field Assessment Officer. Early in training he administers the Minnesota Multiphasic Personality Inventory, a screening device to identify trainees for whom a more extensive psychiatric evaluation may be needed. Other psychological tests may be administered where they are felt to be useful in further assessment. Approximately half way through the training program (six weeks) and at the end of training, he obtains peer ratings from the trainees. However, all psychological tests are voluntary, and a trainee may reject any test or test item where he feels that his right to privacy is being invaded. The Field Assessment Officer is also responsible for securing systematic evaluations from all members of the instructional staff - not only grades in courses but also judgments regarding each trainee's suitability for Volunteer service. Finally, he is responsible for preparing a systematic Training Resume for each trainee including recommendation for the conditions of overseas placement. The Training Resume is forwarded to the Peace Corps Country Representative overseas.

All applicants who accept an invitation to training are also required to complete certain forms used to initiate a full field investigation. This is carried out by the Field Investigation Branch of the U. S. Civil Service Commission. Its investigators visit each community where the applicant has lived and make inquiries about him from his neighbors, supervisors, and associates. They also check his credit and police records. At the end of about two months, the Civil Service Commission provides the Field Selection Branch with a detailed but unevaluated report on the applicant. In most cases, this report merely confirms the generally favorable picture of the applicant as revealed in his pre-training assessment and, thus, leads to greater confidence in the candidate. In a few cases this report reveals derogatory information regarding the applicant, sometimes so serious as to require immediate selection-out, at other times to require a careful scrutiny of the applicant during training before a final decision is reached about his suitability for an overseas assignment.

For each Peace Corps program, there is an Intermediate Advisory Selection Board and later a Final Advisory Selection Board; the membership of these Boards is as follows:

Peace Corps: Field Selection Officer, Program Development Officer, Corps Representative, and the Training Officer, if appropriate.

Training Institution: Program Director, Field Assessment Officer, M. Consultant, and Psychiatric Consultant.

These persons constitute the Board of which the Peace Corps Field Selection Officer serves as Chairman. Additional persons who are in a position to contribute useful information regarding trainees may be invited by the Chairman to participate in all or part of the meeting of the Board.

The Final Advisory Selection Board is charged with the heavy responsibility of advising the Peace Corps as to which trainees are judged to be fully qualified for overseas service and which are not.

The Peace Corps Director has delegated basic responsibility for the selection of Volunteers to the Director of the Division of Selection. The Director of Selection cannot participate direc in the evaluation of each trainee, so he relies on his staff of trained Field Selection Officers, each responsible for assigned programs. Each follows a set of designated procedures designed to lead to the best recommendation in each case. However, since the Director of Selection must accept the responsibility for each decision, the action of the Board is advisory to him. Naturally, he hopes that he can regularly accept the recommendation of the Board. But, if for any reason (e.g., sharp disagreement among members of the Board or failure to weigh adequately all evidence) he believes he should take action other than that recommended by the Board, he has the right and the obligation to do so.

The ultimate responsibility for the selection of Volunteers rests with the Director of the Peace Corps. This is no abstraction. For example, should there be serious disagreement or serious doubt among the members of the Final Advisory Board, the final decision on selection is made by the Director of the Peace Corps.

Implications

The question remains, is this selection program effective? A review of the Peace Corps history indicates that the Volunteers have performed amazingly well overseas. Only 7 per cent are terminated early for failure to adequately adjust to their assignment, while a fraction of 1 per cent are returned for more serious psychiatric reasons. Eighty-five per cent successfully complete their full tour of duty.

In view of this high success rate, the research questions referred to earlier raise questions, at least for the Peace Corps, with respect to the typical regression model used in most selection research. The empirical validation of tests or items makes at least two assumptions:

1. Future samples to be selected come from the same population as the validation sample.

2. The criterion for performance is highly stable over time.

Both of these assumptions, particularly the latter, are highly questionable for the Peace Corps. The educational, motivational, age, and sex characteristics of Peace Corps applicants are highly variable depending on, among other things, the season of the year, state of the economy, draft quotas, internal and external reports of conditions in foreign countries, and the nature of our recruiting campaigns. And our criteria of success are constantly changing, depending on the biases of Peace Corps staff overseas, the nature of the jobs, and many other factors. In essence, success or failure is significantly dependent not only on the characteristics of the individual but on the ecology of the environment.

Where does this leave selection? Again, I have no panaceas. Pragmatism tells me that there is no "Peace Corps" type. Consequently, we try to ensure that our selection procedures are global, judgmental, flexible, and designed to produce heterogeneity rather than homogeneity. Sometimes, I think that a selection system which minimizes the number of required characteristics for entrance and maximizes the number of characteristics permitted for graduation will, in the long run, best serve its purpose.

We have become less sure of the accuracy of prediction; we have become more concerned with our role as experts. We have also become concerned with the way in which selection is consistent with Peace Corps philosophy and American culture. If, indeed, the Peace Corps is an "open society," what is the optimum role of the psychologist as selector? Still groping, we have tentatively agreed that the selection process, too, must be open. One assumption under which we are working is that evaluation must be a two-way process. It is somehow unfair to elicit information which is not shared with the individual trainee. Consequently, we are trying to make our testing and evaluation a participative process. At various points in the training process, test results, peer ratings, and faculty evaluations are discussed with the trainee. The material presented to the trainee is not given as "the truth" but as the way he is perceived by tests, peers, faculty, psychologists, etc. He is encouraged to discuss these evaluations and to test them against his own self-image. He is challenged constantly to question the reality of his desire to serve in the Peace Corps as well as his ability to do so. With this approach, perhaps we can obtain higher validities for our personality measures. Hopefully, when the trainee takes the test he is more motivated to "tell the truth" as he sees it, since he knows he will not be arbitrarily eliminated as a result of his test responses and that he, too, will find out what his responses mean. This approach places the psychologist in a difficult role, since he must be both evaluator and counselor, but I am convinced it can be done.

In one sense, our selection process implicitly embodies the principle that the most efficient technique is one that collects as many independent evaluations of relevant behavior in relevant situations as is possible. Consequently, we constantly challenge training to imitate as closely as possible the actual job and environment. If the job is unstructured, unsupervised, perhaps the training program should be, also. Perhaps the one reason why correlations between training performance and job performance are so low is that, except for knowledge factors, many of the factors relevant to performance are not present in the training situation.

In our research efforts in the future, we shall concentrate on in-depth studies of the criterion situation to help determine the relevant pre-training and training characteristics. This may lead to both more relevant training and more relevant selection. I am convinced the two must go hand in hand.

X

THE ROLE OF INFORMATION IN PERSONNEL MANAGEMENT

Carl F. Kossack
The University of Georgia

The need for more careful consideration being given to the role of information in management systems is slowly becoming recognized. All too often the increased capability in data processing and data transmission that is realized through the use of high-speed computer and data processing systems simply results in the development of additional data generation involving larger and larger data files. Such data are often evolved independent of the need for system-type information, or at best, each set of data is developed with little regard being given to the interdependency of the data contained in different files. One writer has paraphrased the adage, "one is unable to see the truth because of all the facts." It is the purpose of this paper to consider the information needs associated with personnel management systems and to suggest an approach that may prove to be a useful vehicle through which one can consider this type of information problem. Experience has also shown that the introduction of system-type concepts to management problems almost always discloses a data deficiency relative to the information required to apply the system model to an operational situation. In this regard, it appears that the information acquired and used by individual managers in meeting their particular narrow responsibility is not too useful when a more total system approach is considered. All too often, one finds the available data is either incomplete, since it has been gathered only long enough to shed some light on a problem that has arisen, or the more continuous type information files that are available are those that are required for financial reporting or legal and tax considerations; and these types of data appear to be studiously inappropriate for use in management-type decision-making. One final introductory remark seems required, and that involves making a distinction between data and information. Although in the above disoussion there has been no attempt to distinguish between data and information, in modern usage one considers data as being the raw measures and characteristics that are acquired during the operation of any program, while the term information is restricted to processed data where the processing is oriented towards estimations required in decision-making. Using these definitions, it is the intent of the paper to discuss how personnel data may be converted into information relating to a personnel management system.

At a meeting of the Institute of Management Science held some time ago, W. O. Trueblood (1) presented a paper in which he outlined an information flow model for corporate decision-making. In his paper, Trueblood gave a broad general description of the flow mechanism involved in such an information system; however, he did not attempt to be very specific in how such a concept might be applied in any practical application. In 1963, the author presented a paper (2) which attempted to develop in more detail these concepts, especially as they would apply to an electronics manufacturing program. Generally, the Trueblood Model would also be applicable in helping to define the role of information in personnel management, and the rest of this paper will be devoted to such a development.

Let us first consider the Trueblood Model as it was initially presented. In the model three levels of management were recognized: (1) the headquarters level, (2) the divisional level, and (3) the plant level. Of course, these designations for the three levels are arbitrary, but those used by Trueblood are the ones often associated with a manufacturing-type organization.

To describe the information flow model, let us present the flow diagram as first introduced by Trueblood and then discuss the model by using the diagram (Figure 1).

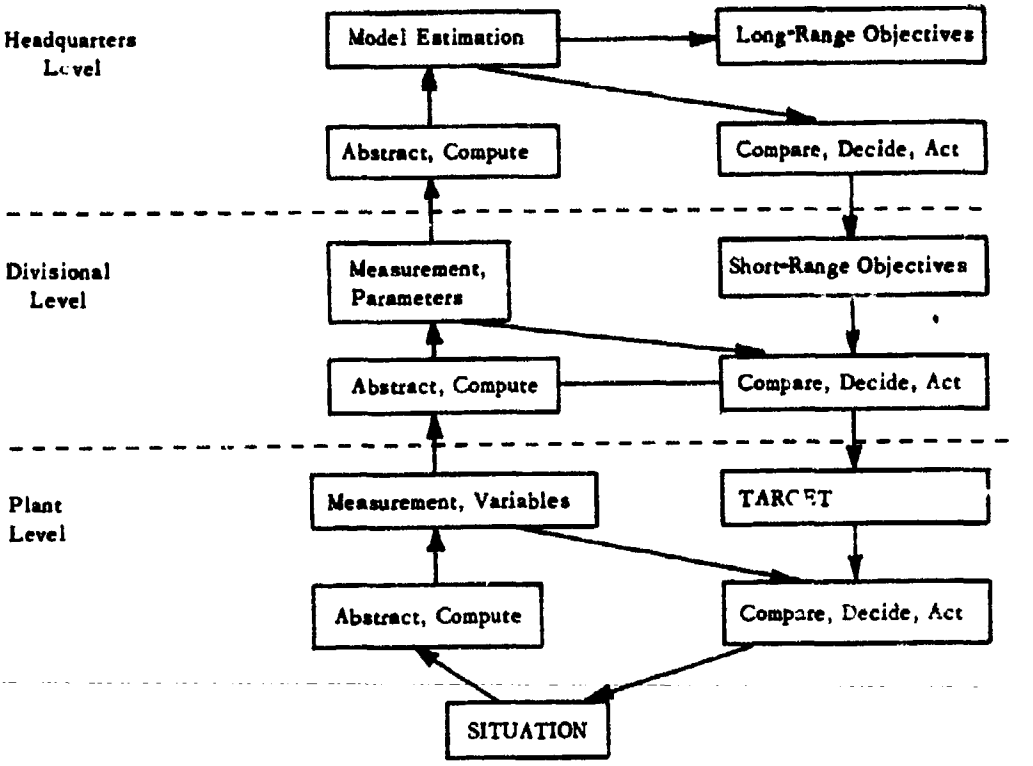

Fig. 1. **Flow Diagram - Information Flow Model**

From the above flow diagram one can observe that at the plant level values of the "variables" of the SITUATION must be measured and that the decisions made at the plant should be based upon these observed values relative to their expected or planned level; and further that the decisions made would have a direct effect upon the situation. It is expected that such variables would be operational in their general nature. At the divisional level one obtains "parameter" estimates using the measures of the variables that are transmitted to the divisional level from the plants. Decisions are made on the basis of these parameter estimates, and such decisions are concerned with short-range objectives. At the headquarters level the concern is with model estimation, and the decisions are of a long-range planning nature.

Now in order to appreciate the meaning and mechanisms that are associated with such a flow diagram, one must define such terms as variable, parameter, model, etc. To do this let us consider a definite personnel utilization system situation; for this purpose I have elected to use a simplified personnel utilization approach that one might find in a weapons system operational program. I will not attempt to

use any realistic descriptions in this exposition, first because I have none available, but also because I am afraid that the demands of realism would complicate the consideration to the extent that the general concept might be lost. I will use in the discussion the personnel utilization problems associated with the introduction of a new weapons system into the defense arsenal of a country with particular reference to the crew assignment problems of such a new system.

To consider the model that might be used in the long-range personnel utilization planning for the introduction of such a weapons system, I will utilize certain concepts that I introduced in my 1960 paper on the mathematics of personnel utilization models (3). Our model thus assumes the existence of numerical functions defined over the characterization space which, for each type of personnel required to operate the weapons system, the function estimates the relative effectiveness of an individual if given such an assignment. In addition to these effectiveness functions, some team-effectiveness model must also be available so that one can convert the relative effectiveness of each member of a team into a measure of system effectiveness.

One such effectiveness model may be in the form of a family of learning curves as shown in Figure 2. Each curve shows the progress of a crew in its system utilization training program, with the appropriate member of the family of curves being selected according to the background characteristics of the particular crew members. Thus, we may have the following representations. It should be noted that the expected final level of effectiveness of a crew may also vary according to the characteristics of the crew.

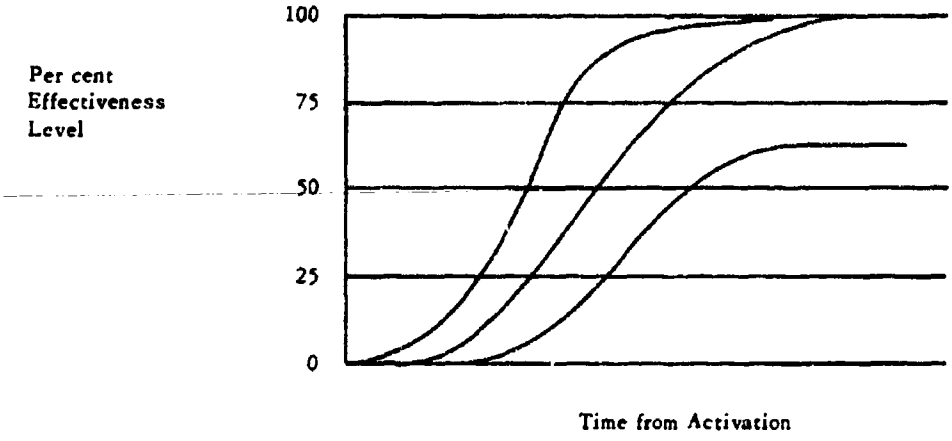

Fig. 2. Family of Learning Curves for Crew Effectiveness

These individual effectiveness functions and crew effectiveness models are research and development type problems that are the responsibility of individuals at the headquarters level.

The long-range objectives associated with the weapons system involves the introduction of the new system into the defense program in such a way that the manpower resources of the country are allotted to the several available systems in such a way that the effectiveness of the total defense system will meet the requirements of the country. These long-range plans will utilize the effectiveness model of the new system, in combination with general personnel assignment plans, to provide this interactive type of long-range plan. Contained within these long-range defense plans will be the short-range plan for the buildup of the overall effectiveness of the new weapons system. For example, these short-range plans may be characterized again by a type of growth curve, and thus the parameters associated with such a growth curve would be the parameters of the short-term plan.

In fact, this growth curve would be the short-term plan given the divisional level who would be assigned the responsibility, using the plan, of implementing the new system into the defense program. Such a growth curve may be represented by the mathematical model

$$y = \frac{1}{\alpha + \beta p^t}$$

where y is the relative effectiveness of the system, t is the time from initiation of the implementation program and α, β, and p are the parameters.

The divisional level would need to program the field (plant level) implementation of the individual system installations and for such a purpose would need to use the learning curve models for each individual installation. A member of the family would be selected to correspond to the characteristics of the crew assigned to the particular system. Each field installation would then be charged with the responsibility of developing the readiness of each system assigned it and for this purpose would be given the TARGET for each system in the form of the learning curve associated with the characteristics of the crews to be made available to the field command. It is assumed that at any given time a measure could be made of the percentage of completion of the training program.

Let us review the overall model before discussing the information requirements for the utilization of the model in an actual operational situation. The basic model considered is in the form of a learning curve whose shape depends upon the background characteristics of the crew members assigned to the particular system. This curve gives the expected time to operational readiness for the system as well as the estimate of the final overall effectiveness for the system. In the long-range planning program carried out at the headquarters level, the personnel allocations are made for the implementation of the system within the constraints of the overall service requirements. These allocations are then used to evolve the individual systems learning curves which are then integrated into the total system effectiveness model. This total system effectiveness model is then used with similar models for the other weapons systems available for the defense program to provide the total long-range defense plans. At the divisional level the responsibility of implementing the new system is assigned, and the short-term goal is transmitted in the form of the total system effectiveness model as used in the long-range planning program. Now the divisional level will take the several system installations that are required by the long-range plan and assign one or more individual system to each of several field level organizations. Each organization will be assigned crews as the plans call for and given the expected learning curve for their system as the TARGET.

With the implementation of the new system, the SITUATION at each field organization will be essentially a measure of the status of the system. We can now define the variable as used by Trueblood as being the per cent effectiveness of a given system at any given time. Thus the field level will make repeated measurements of this percentage as time goes on and will compare these with the level given in the TARGET. The field level by this comparison may DECIDE to act, with their actions being restricted to

those that fall within their own field activity. For example, if one system assigned the field organization is well ahead of TARGET while another, still the responsibility of the same field organization, is falling behind its TARGET, transfers of key personnel might be made in the hope of better meeting both TARGETS.

As these variable values for each system are being obtained by each field organization, each system's status would be reported to the division who would ABSTRACT the data and compute the best estimates of the total systems growth model parameters, say $\hat{\alpha}$,$\hat{\beta}$,and $\hat{\rho}$. These estimates would be compared with the parameter values given the division by the headquarters level, and if a significant variation is indicated, the division may make transfers of personnel between field organizations, or shift individual systems responsibilities, or simply recommend changes that may be instituted within each given field organization. Such changes are made to better meet the short-range goal.

Now these parameter estimates, as they are obtained, would be reported to the headquarters level which would be receiving similar parameter estimates for the other weapons systems involved in the long-range defense plans. These estimates would then be inserted into the long-range planning model and the results compared with the long-range plans. If there were a significant variation between the actual situations and the planned situations, changes, such as delaying the phasing out of an old weapons system, may be made by the headquarters.

In addition to the long-range planning model, the headquarters level is responsible for the validity of the personnel effectiveness models that are used in evaluating system effectiveness. Thus, in the early planning stages certain assumptions must necessarily be made with regard to the relationships that exist between the characteristics of individuals and their expected effectiveness when assigned a particular responsibility as a crew member of one of the new systems. As experience is accummulated in the form of both variable measurements and system parameter estimates, the headquarters research and development group would correspond these experiences with the expected results obtained from the model assumptions. If there were a significant discrepency between them, the R&D group would have the responsibility of initiating studies aimed at modifying the models so as to make them correspond more closely with the experience.

This information flow system, associated as it is with the personnel utilization and effectiveness models, defines quantities whose values are critical in the personnel management of the system. Thus, the variables that are measured in the field are direct indicators of the status of each system, while they also contribute to the estimation of model parameters. In addition, the interrelation between personnel characteristics and system effectiveness as considered by the system models provides critical information relative to the man-machine problem so frequently encountered in new systems.

Although it is recognized that the above consideration is lacking in realistic details and perhaps even in logical content, the purpose of this paper has been to point out the need for system modeling if one is to develop a meaningful information flow mechanism. It has been all too often the experience of the author to find that reports required by management provide, at best, information for a special study of a particular problem and that the usefulness of the data after the completion of the special study is very limited. Often these report requirements are retained by management on a sort of "just in case" basis. Until a more total system information program, such as that envisaged by Trueblood, is developed, there is little hope of having the right data available at the right time.

In conclusion, I feel it necessary to note that such ideal total system concepts seem to be exceptions rather than the rule in our society, and perhaps the complexity of the human factor as it enters such systems makes it impossible to approach information requirements of a system at such a logical level. The challenge is surely with us in this respect, and any advance made in this direction cannot help but assist management in meeting its responsibilities, even if the final total system program is never realized.

REFERENCES

1. Trueblood, W.O. Intangible aspects of management control system design. Paper read at TIMS Symposium on Management Control System, 7th Annual Meeting of the Institute of Management Sciences, New York City, October 1960.
2. Kossack, C.F. The role of information in manufacturing planning. Paper read at the Fall meeting of the Southwest Chapter of TIMS, Dallas, Texas, September 1963.
3. Kossack, C.F. A personnel utilization model. *Proceedings of Symposium on Air Force Human Engineering and Training Research*, National Academy of Science. Publication 783, National Research Council, 1960, 32-40.

XI

THE EVOLUTION OF METHOD IN PERSONNEL RESEARCH

John C. Flanagan
American Institutes for Research

This discussion of the evolution of method in personnel research makes several assumptions which should be explicitly stated at the outset of this presentation. First, it is intended that *evolution* mean improvement and that *method* be given a broad meaning, such as, a systematic way of achieving an end. Using these definitions, the title can be stated as improvement in the ways of achieving the aims of personnel research. As all personnel researchers are aware, in order to evaluate improvements in achieving ends, it is of great importance that the ends be clearly and completely defined. Focusing on the evolution of method in personnel research in this light immediately reveals that the aims of personnel research have changed very markedly over the period of the last 25 years.

Simply stated, the goals of the personnel researchers have shifted from establishing that simple relationships exist between a predictor and a criterion to the much broader aim of contributing to the solution of the problem of how a total system such as the defense system, or the Air Force subsystem of this total system, can achieve maximum effectiveness at a given level of expenditure.

The first big change, then, in method in personnel research relates to the definition of the problems which are to be researched. Experience has shown that the most critical factor in determining the success or failure of personnel research is the definition of the problem. Many large, sophisticated, and well conducted studies have, after the correct application of powerful analytical methods, presented unimpeachable evidence pointing to conclusions of only trivial value.

A second important topic under the heading of evolution of method in personnel research relates to the general area of data collection. This includes specifying the populations to be involved, the variables to be observed, the procedures and measurement devices to be used in making the observations, and the scoring or grouping of these observations for data-analysis purposes.

The third topic to be discussed will be improvements in methods for data analysis. This will include changes in statistical methods, equipment, and concepts in this field.

The final topic to be included will be a consideration of data interpretation, data presentation, and decisions.

To return to the first topic which relates to the evolution of methods in defining problems for personnel research, a brief review should provide some perspective. Twenty-five years ago the attention of most personnel researchers was focused on individual studies. Frequently the question was defined as "Will this one general ability test predict success in this training course?" In training itself, a personnel researcher spent a substantial amount of time developing measures of progress to be used in training. They discovered

early that the most important problem in training was usually the one of "Is this training directed at the most important requirements for effective performance on the job?" Only gradually did personnel researchers discover that it was profitable for them to study job design and the effectiveness of operating procedures involving personnel.

In the 1950's a very significant step was taken by personnel researchers in the direction of broadening their outlook. One of the important contributors to this development was Dr. Robert B. Miller, a former member of the Army Air Force's Aviation Psychology Program. Dr. Miller (1) working on an Air Force contract with the American Institutes for Research, developed the outlines for a procedure for evaluating training devices. However, halfway through the project, he noted that much of the information relating to the validity and effectiveness of the training device could be obtained by analyzing the design of the device before it was built. On his recommendation the purpose of the project was changed to the evaluation of the probable effectiveness of training devices at the design stage.

This, and other projects which Dr. Miller carried out at A.I.R. for various Air Force research groups (2), led to the development of task-analysis procedures and to the development in the Air Force of the human factors requirements for the design of personnel subsystems in connection with new experimental and prototype weapons systems. Doctors Walter Grether, Robert M. Gagne, Arthur W. Melton, and others were active in formulating Air Force policies and plans in this area (3).

Examples which illustrate the need for precise definition of the problem in personnel research are given for selection and classification studies and for training programs. The accepted and recommended procedure of most psychologists 25 years ago for obtaining a valid set of predictors in a selection or classification situation consisted in trying everything available and being quite surprised at which things turned out to have predictive value. At the present time this has been largely replaced by systematic and comprehensive studies of job performance. Among the techniques frequently used for this purpose are the critical incident technique and task analysis. Using these data with such a procedure as the method of explicit rationales, a relatively sound basis for defining job requirements is achieved. In addition, the method of explicit rationales presents for critical review prior to development or empirical validation the insights and inferences on which the selection or development of predictor measures is based. The same logic and procedures are used for setting up a proficiency or performance test to measure the individual's learning in terms of his performance with respect to a particular set of job requirements.

A current trend somewhat accelerated by work in the programmed learning field is the development of assessment and measurement procedures designed to measure all of the behavioral objectives of a training course before developing any of the training materials for the course. An example of this type of procedure with a few additional modifications is provided by a project just completed by Mr. David G. Markle (4) of A.I.R.'s Palo Alto office. The project was designed to give the trainee as complete an education in first aid as possible in a 7.5-hour, 1-day training period. With the assistance of the American Red Cross, a 500-item test was developed including all of the skills and information regarded as essential to such a course.

Before beginning work on the course, the test was given to a number of new employees of the American Telephone and Telegraph Company which supported the project to determine which items adults of this type already knew. The final training program consisted of 20 instructional films of 1.5 to 16 minutes in length, 17 sessions for practicing newly taught skills, and 13 workbook sessions. These latter were generally in programmed learning format. The results from various tryouts indicated that employees with no known first aid training obtained about 29 per cent of the items correct; employees who had taken one of four previously developed programs obtained about 40 per cent of the items correct; employees who had taken the 7.5-hour final program as developed by Mr. Markle and his staff obtained 81 per cent of the items correct.

Thus, the gain from the previous training program represented only an increase of 11 per cent over what untrained employees already knew, whereas the new program provided a gain of 52 per cent over this previous level of knowledge and skill. Various procedures were designed to increase the efficiency of the development program and to avoid teaching things already known to the employees. The pretesting of newly hired personnel was one method. A second was the use of a criterion frame outline in the form of subtests for various segments of the sequenced learning program. These subtests made it possible to check on the effectiveness of a specific instructional sequence and modify it if the students were found not to have learned.

A deficiency of this type of empirical test is that it does not indicate those instances where a shorter training program would have been satisfactory. Because of the tightness of the program and the large amount of learning to be accomplished in 7.5 hours, a systematic effort was made on this project to develop "lean" programs initially which were likely to err on the side of insufficient learning. Where deficiencies were found, additional instructional materials were added in small increments until the desired training results were obtained.

One of the interesting sidelights of this study was that the employee with the highest score from any of the previous training programs obtained 66 per cent whereas the employee with the lowest score from the new program obtained 71 per cent of the items correct. Thus, there was no overlap between the two groups, and the poorest learner from the new program was superior to the best learner using the old program. It seems clear from this example that the systematic and detailed analysis of the skills and information to be obtained in a training course and the systematic development of a balanced training program using empirical checks on various segments of the program in the form of subtests should greatly improve the efficiency of much of the present training materials.

One of the problems in predicting both training and on-the-job performance using selection and classification tests has been that there is frequently a wide variation in the quality of the training received by various personnel included in the program. Similarly, motivation, morale, and operating conditions can vary to such an extent that they seriously attenuate the basic relations between the predictive instruments and the performance criterion. It is believed that as training is improved and becomes more consistent in quality, and supervision and operating procedures are improved so as to obtain a higher and more uniform motivation and morale among the personnel involved, the predicted measures will show much higher correlations with the criterion than are currently found.

An extreme example illustrating these attenuating factors is provided by the experience of one of our projects in Nigeria. In this study, done under the direction of Dr. Richard R. Rowe, it was found that achievement tests given to college entrants predicted performance in the freshman class quite well. The relation between these predictor scores and subsequent grades dropped as the students advanced to each succeeding class; by the time that they were seniors, the relation between the predictor achievement scores and their grades at college in their fourth year was essentially zero. This was not at all surprising because of the very great differences in the quality of elementary and secondary schooling received by the various students within this class.

The last point to be made regarding the definition of the problem is that the new aim of the personnel researcher is to maximize the contribution of the personnel component to the effectiveness of the total system. Thus, the personnel researcher is faced with a much broader problem at this time. He is now a member of a team trying to integrate his data on cost effectiveness to maximize the performance of the overall system.

The second aspect of method to be discussed is data collection. This includes definition of the populations, specification of the variables, development of procedures and devices for making the necessary observations, and scoring or grouping of this information for subsequent analysis. Twenty-five years ago, populations tended to be rather small. Frequently the number of variables was only one. Most of the procedures used for prediction were standard published tests, and the scoring and grouping of the data from these measures was nearly always quite simple and conventional.

During World War II, many new tests were developed for specific purposes using the types of procedures mentioned in the previous section. Also during World War II, procedures were initiated to follow all aviation cadets through their courses in the training command. This continuous followup program represented a radical departure from previous studies on selected sample groups. Although much reliance was placed on simple criteria, such as pass-fail in training and training grades (some of which had doubtful validity), there were also many developmental projects aimed at improving the criterion measures. Some of the procedures which were initiated at that time and which have been improved considerably in recent years include the objective flight check, the situational performance test, and the standardized assignment.

In addition to the development of specific measuring devices and observational procedures to meet particular data collection needs, perhaps the most important trend is in the direction of much more comprehensive data collection, both initially and during the individual's training and operational activities. It is now possible to include individual item data and specific observations, as well as summary scores and grouped observations, as a basis for future decisions. It seems fair to say, however, that the development of measures and techniques for data collection has not kept pace with the rapid progress in the other three topics discussed here. Perhaps the personnel researcher's efforts will need to be focused on improving data collection procedures in the next few years.

The next topic is the evolution of methods of data analysis in personnel research. This topic will be discussed under three headings: statistical methodology, equipment, and concepts. The evolution of theory and mathematical statistics has paralleled and in most instances led evolution in the other two areas. More than 50 years ago, Karl Pearson and his students provided theoretical formulae for tests of significance with respect to statistics calculated from samples. Their function was to determine how often a deviation as large as that observed could be expected if a large number of similar estimates were obtained from comparably selected samples. These values were expressed as standard deviations, or standard errors. Since for many statistics, values obtained from samples in this way tended to be distributed in terms of the normal probability function, it was possible to calculate rather readily how often a value as large as the obtained statistical value could be expected as a result simply of variation from sample to sample under the stated conditions. The conditions were usually stated in the form of a null hypothesis that the two observed values come from the same parent population. If the null hypothesis was rejected, this was usually interpreted as confirmation that the values were not from the same population but that a real difference existed. Fiducial limits were proposed by R. A. Fisher to test whether it was reasonably likely that alternate hypotheses could account for the observed data. However, since these hypotheses need to be specified in advance, the only risk considered was usually the *a priori* hypothesis which was most frequently the null hypothesis. The probabilities reported usually related to the chances that the stated hypothesis would be rejected if true.

In the 1930's, J. Neyman and E. Pearson helped to clarify the logic of testing hypotheses by emphasizing the importance of the second type of risk or error, that of accepting the initial hypothesis when some other hypothesis was true. They referred to the two types of risks or errors as "errors of the first kind" and "errors of the second kind."

A general statistical decision theory was developed by Abraham Wald as an extension of the logic to all of the various risks involved. He first applied his theory to problems of sequential analysis of statistical data. The logic, however, applies to all types of statistical analysis. The essence of decision theory is to

consider cost considerations and consequences of d isions in all situations. The decision maker is required to estimate the cost of an error of the first kind an lso the cost of an error of the second kind. A strategy is sought which would yield decisions which would optimal in terms of costs and benefits.

The importance of decision theory was th forced the personnel researcher to look not at the question "Does a relationship exist?", but to tal n many more factors in the situation and consider the consequences of making various decisions on the sis of the available data.

A somewhat parallel development over the past years has been the work of Doctors Hubert Brogden, Paul Horst, and Lee Cronbach, and Goldine leser (5). These studies have related to maximizing the benefits from selection and differential prediction and the effects of variations in testing costs on net gains. The value of the concept of decision theory is that it enables a rigorous comparison of alternate testing, training, and operating procedures when all of the possible payoffs and costs under each of these alternates have been estimated. The development of decision theory is certainly one of the most significant events in the evolution of method in personnel research.

The evolution of equipment for use by the personnel researcher has been dramatic. From written calculations, to hand adding machines and calculators, to punched-card tabulating equipment, to electronic computers, to large capacity, remote-input, time-shared computers has involved an almost unbelievably rapid increase in ability to use sophisticated and complex concepts in the personnel research field. Very important advances were made in the Air Force in World War II in terms of using large numbers (i.e., whole classes), using standard procedures and automatic followup, using punched-card tabulating equipment and scoring machines, and also introducing built-in evaluation of the possible contribution from new measures and instruments. However, the new electronic computers make possible very important forward steps in terms of collecting much more detailed data regarding each person by following him in his career and evaluating his performance in great detail and obtaining data for maximizing the benefits from personnel decisions in a way which would have been quite impossible without the new equipment.

It is interesting that as decision theory was being developed to provide the broad outlines for a precise solution to maximize specific systems from a theoretical point of view, the electronic equipment essential to any such analyses was also being developed, and at the same time new concepts for overall system planning were being formulated. Some of the basic concepts which have had great influence on the personnel researcher include systems engineering and project definition. Perhaps the most important aspects of these concepts are the requirement for establishing primary mission goals and major performance goals and insisting on the careful and systematic estimation of total costs and comprehensive measures of effectiveness for the system. A typical requirement is the limiting of the functional flow block diagrams to pure functions to permit man and equipment trade-offs later in the system analysis program. These concepts force the personnel researcher to be a part of the design of the system which specifies performance goals and systematically explores all avenues for minimizing cost. The personnel researcher must be an expert on human behavior under various conditions and be able to make realistic estimates of the long-term performance of an individual on a particular job proposed as a part of the personnel subsystem. Thus, the framework for data analysis has increased its sophistication to keep pace with the development in the broadening of the problem of the personnel researcher.

The last topic to be discussed relates to data interpretation, data presentations, and decisions. This topic is closely related to some of the concepts discussed under the last heading. The interpretation must be made in terms of a larger setting and in terms of a functioning of a complete system rather than as an isolated item. There is a great need for more precise cost accounting and information on personnel reliability and personnel performance under unusual conditions. Data must be presented simply and in terms of system effectiveness rather than individual performance. Some of the more powerful concepts of multivariate analysis are important for presenting the increased volume of data in a simple and easily

interpreted form. One of the problems in data interpretation which is a source of difficulty at the present time is the validity of the measures and observations obtained. It is difficult to make the correct decision with respect to a broad, well-stated problem even with the aid of sophisticated theoretical formulations and advanced hardware if the techniques for data collection are not entirely trustworthy and are lacking in precision.

In summary, the problem of the personnel researcher has changed a great deal in the last 25 years. Personnel research is now much more clearly one part of a team working to optimize the results in an entire system. Interaction of personnel with other components of the system must be an integral part of the new research program. This change in the definition of personnel research has added requirements for collecting data on many more variables and in much greater detail. It seems that this area of data collection offers the personnel researcher his greatest opportunity to contribute to the increased quality of decisions. More precise and valid techniques for measurement and observation would clearly make possible a sharper focus on the optimal decision.

The development of decision theory, new electronic computer equipment, and sophisticated concepts in the areas of systems engineering and related management problems have kept pace with the new definition of the problem of personnel research. Another area in which important contributions are needed relates to data interpretation and decision making.

It appears that the conceptualization of the problem of personnel research, statistical theory, computer equipment, and management concepts have all developed at a more rapid pace than the data collection and data interpretation methods of the personnel researcher. This field will need intensive efforts for some years to enable management to take full advantage of the opportunities for improved operations that these developments have made possible.

REFERENCES

1. Miller, R.B. *Handbook on training and training equipment design.* WADC Technical Report 53-136, Wright-Patterson Air Force Base, Ohio, 1953. (a)
2. Miller, R.B. *A method for man-machine task analysis.* WADC Technical Report 53-137, Wright-Patterson Air Force Base, Ohio, 1953. (b)
3. Gagne, R.M. et al. *Psychological principles in system development.* New York: Holt, Rinehart & Winston, 1962.
4. Markle, D.G. *Final report: The development of a first aid training and personal safety course for Bell System employees.* American Institutes for Research in the Behavioral Sciences, Instructional Methods Program, Palo Alto, California, 1966.
5. Cronbach, L.J. & Gleser, Goldine C. The dollar dimension in testing decisions. *Psychological tests and personnel decisions.* Urbana, Illinois: University of Illinois Press, 1957.

XII

MILITARY PERSONNEL RESEARCH: AN HISTORICAL PERSPECTIVE

Leland D. Brokaw
Personnel Research Laboratory
Department of the Air Force

Because so many members of the audience have spent major and minor segments of their professional life dealing with the nuts and bolts of the design phases of the present structure, this paper will not dwell upon the specifics of the development of a test, the rationale for a psychomotor apparatus, or the procedures followed in the collection of critical incidents. The evolution of types of research will be traced in the hope of making more clear the enduring pressures and priorities that underlie military personnel research.

It is interesting to attempt to set the date from which this review should proceed. This is the Twenty-fifth Anniversary Symposium; and it is quite true that November 1941 is an historic date in the United States Army Air Corps Aviation Psychology Program. To imply that our present circumstance had its origin at that precise instant would be misleading. It is not appropriate to set the zero point at the time that Dr. Yerkes and his colleagues assisted the War Department with the testing program leading to Army Alpha in World War I, because they depended upon still earlier researchers in the domain of mental measurement. Wilhelm Wundt, and the psycho-physiologists of his day, in a sense mark the beginning; it is improbable that profit would come from the identification of the precise instant and place whence sprung our current program. Better far that we should treat it as a mental measurement scale, identify the origin as undefined, and attempt to work along from the arbitrary point established by this birthday party.

This review will be biased. Since 1943 the writer has worked within the Aviation Psychology Program, in those organizations that have evolved into the present Air Force Personnel Research Laboratory. Formal sources for the material of this review could be the program documentation, the technical reports, and the papers we have placed in the journals, and at the professional meetings. However, this discussion will reflect experience; the program documents, the technical reports, the journal articles, and APA papers are like the exposed portions of icebergs--indications of general shape and size, but concealing major portions of interest.

Investigation of the characteristics of the published materials of the Laboratory reveal some continuity of certain program aspects, certain cycles of investigation that ebb and flow, and isolated explosions of specific activity.

There is a steady tom-tom beat of production, calibration, and validation of operational tests. Indeed the tap root of personnel research springs from the exploitation of individual differences in selection and classification programs.

The dramatic relations between stanine level and elimination rates for pilot trainees who had been screened on an aptitudinal basis in World War II are well documented, well regarded, and well worn as justification for other projects in the personnel research area. Incidentally, the matter of pilot selection continues. We find the same types of paper-and-pencil tests to be valid now that were valid in 1942, and that the same principles of supply and demand apply.

The considerations which led to the use of the aircrew measures provided a basis for the development of relatively sophisticated tests to use in the selection of airmen for service in the Air Force and to define their area of occupational assignment within the Air Force. These measures demonstrated very useful levels of validity against criteria of school success, and here we find the eternal question in personnel research. Our field of significant achievement is also our field of significant need. Our tests have had solid success in the identification of those individuals who might be expected to do well in training, yet attempts to relate test scores to operational measures of success on a job have yielded insignificant results.

This dilemma, recognized throughout the profession as "the criterion problem," has produced tons of ratiocination, justification, and explanation. It is difficult to accept the concept that measures related to success in training for a certain job should have zero relationships with success on that job. It is improbable that zero relationships *do* obtain between those aspects of the job that are trained and the selection score. It is probable that other factors obscure the relationships if direct measures are made of global factors of job success. Evidence in support of this contention can be inferred from McReynolds' (1) finding in 1958 that successively higher aptitude levels were associated with successively higher grade in Air Force enlisted populations. This is not clean, because policies of elimination of non-productive airmen are correlated with aptitude scores.

Let us specify that objective identification of comparative levels of achievement on the job is a basic need of the science. Encouragement comes from studies based on the concept that homogeneous work on the part of all members of a criterion assessment sample enhances the probability of objective evaluation of the members of that group. Positive results from these studies were reported by Wiley (2;3) in 1964 and 1965. There is reason to believe that evaluation of proficiency at the task level may produce useful result. This, however, devolves upon future research.

What are the ingredients of the twenty-five years of personnel research under review? Casual inspection of the titles of hundreds of reports emanating in that period suggests a heterogeneity like unto an oriental bazaar. More careful screening reveals that there are certain areas which receive attention on either a continuous or a repetitive basis.

The normal human characteristic of naming that which is not understood is well known. An example or two will clarify this point. An infantry platoon is pinned down by enemy fire. The officer and sergeants are killed. Private Joe Doaks rallies the men, redeploys them, escapes the trap with them. Private Doaks demonstrated "Leadership."

A factory worker shows his supervisor how to change the design of a sub-assembly so that it involves fewer parts, is less expensive to manufacture, and is more reliable in performance. He has demonstrated "Creativity."

An individual of limited intellectual resource has spent every available minute studying and after a number of years is awarded a graduate degree by a leading university. He has demonstrated "Motivation."

These three concepts, along with others, have triggered numbers of studies over the years. Some of these studies have been generated from within, when a research worker has had an idea that suggested to him that here, at last, was the key to the problem. Others have been imposed from above, or requested by other organizations.

Consider the recurrence of some of these topics in the Laboratory program.

In 1947 and 1948 John T. Cowles and John T. Dailey, assisted by Bob Keller (4), were called upon by the Officer Candidate School (OCS) to assist in the evaluation of the candidates, in terms of dimensions like "leadership," "motivation," and "creativity." They devised a program wherein a group of officer candidates would be confronted with a problem and left to work it out under their own resources. During the time they worked on the problem, the officer candidates were observed by tactical instructors who had been counseled by the psychologists on the kinds of behavior to rate and who had been provided with forms for the collection of data.

During the experimental and training phases, everyone involved was delighted with the program. Some unrevealed leaders were discovered; some bluffers were unmasked; the judgment of relative capabilities of individuals previously held by tactical officers was supported consistently enough to please them. The system was adopted for use and then met the problem usually encountered when evaluative systems go operational.

The officer candidates' unofficial motto was "Cooperate and graduate." And they did. The word was out, and various cadets manipulated their behavior so that everyone got an equal chance to shine; the means of the ratings crept together; the variances equated; the system failed.

In 1951 Ernest Tupes (5) started collecting peer ratings in the OCS. Candidates were forced to nominate and order their fellows on various trait dimensions. These data were collected across a number of classes; they proved valid for tactical officer's ratings in the school; they were found significantly related to later Officer Effectiveness Reports (OER). The collection of such ratings is onerous, and a policy judgment came from higher headquarters that peer ratings should not be a part of the evaluative process in any officer training unit.

Factor analyses applied to trait ratings by peers in a number of different populations, and reported by Ernest Tupes and Raymond Christal (6), demonstrated a stable structure, with five common factors, across samples as diverse as kindergarten students, officer candidates, fraternity brothers, and college students.

The next cycle of this investigation involved the development of questionnaires touching the factor areas identified by the trait rating analyses by Warren T. Norman under contract at The University of Michigan (7;8;9). This work, completed in 1963, is approaching its final evaluation, in terms of validation of test scores for OERs achieved in the field.

And so it is seen that the fairly straightforward concept of evaluation of would-be officer personnel, along dimensions of behavior which are quickly recognized by name, has been treated in three distinct phases over the years between 1948 and 1966, although as yet substantive findings are not at hand. This set of circumstances highlights a major problem in the relationships between personnel research workers and operational personnel officers. Eighteen years have been spent in productive effort, and even though it now looks more promising than ever before, a method for the quick and objective evaluation of personality traits which are widely regarded as basic to desirable officer behavior is not yet available.

Impatience for results, coupled with a need for answers to pressing problems, has made administrators reluctant to take problems to personnel research workers. Not long ago a senior Air Force officer commented that by and large personnel research was worthless because the answers were always behind the problems. In a momentary sense this can be true, but in a larger view it is quite false. Consider a case which frequently rises. The general says to the psychologist, "I need the answer to Question X."

The psychologist says, "Yes sir, I can provide the answer to Question X in four years."

The general says, "My tour is three years. I cannot wait four years. I will take action Y, and you ignore the problem."

In three years and six months a new general says, "I need the answer to Question X."

The psychologist says, "Yes sir, I can provide the answer to Question X in four years," and the merry-go-round continues.

Fortunately, many persons in the system recognize this phenomenon, so that a major portion of the charter of a research unit is properly devoted to documentation of work that continues without basis in such momentary requests.

The apparent slowness of response observed in personnel research has its basis in facts which argument cannot alter. If you would educe an objective measure of some desirable trait, as for example, leadership, you must not only devise an instrument for the collection of data, but you must provide for the collection of other data which reveal the operation of the trait in question. You must give your sample of individuals upon whom data were collected the opportunity to display the variations in leadership which will demonstrate that you did, or did not, collect initial data which were predictive. This is the reason peer ratings collected in the Officer Candidate School in 1951 are not evaluated in any real sense until 1957; this is why an experimental test administered to thousands of basic airmen at Lackland in 1952 cannot be validated across one hundred technical schools until 1956; this is why a reenlistment index devised in 1960 cannot be evaluated until 1965.

There are ways to shorten these times, but in each case costs go up, and the scientific defensibility of findings goes down. We have shortened the time for validation by administering tests to airmen as they entered the technical school, accumulating a minimal acceptable sample, and correcting the derived validities for the restriction caused by administering the experimental test in an already selected group. We have collected statements of reenlistment intent simultaneously from airmen with one, two, three, and three and a half years of their first tour accomplished, and compared response patterns across the sample. These have been compared across the samples in terms of the reenlistment behavior of the first eligible sample. The succeeding samples have been added as they became eligible.

Now, let us contemplate another problem. We have discussed the objective measurement of personality traits believed basic to desirable officer behavior; we have considered the validation of a selection test for a technical school; we have contemplated the evaluation of reenlistment intent. Each of these topics is presented in a context of personnel research. At this point a number of our academic colleagues will draw their PhD cowls down tightly on their heads and snarl that the foregoing is not research and should not be dignified with such title. In-service applied research programs are but the application of recognized principles to the solution of administrative problems, and properly accomplished by technicians, not scientists.

The vicious aspect of this point is that it contains some truth; and the overt support philosophy for the in-service program has reflected it at times. There has been a period when our program documentation carried a specific indication of the weapons system toward which the research was oriented. At other times this wording has entirely disappeared, and phrases like "investigate the fundamental characteristics of thus-and-so" have taken their place. However, the concensus of senior personnel within the Personnel Research Laboratory over the years has been that the differentiation between basic and applied research in artificial, that without basic elements, the applied program is sterile and cannot last. If you force competent research personnel into a program of techniques where only momentary problems are considered, quality will evaporate, and the whole program will not be far behind.

The Personnel Research Laboratory has been fortunate that there has been a consistent sprinkling of senior officers at our intermediate and high levels of command who have recognized the imperative requirement for investigation of fundamental matters in concert with the solution of the more overt needs of the Air Force Personnel System. They have been responsible for actions such as initiating a research program in the Laboratory in the area of occupational analysis with full knowledge that nothing would come of that program during their own tenure of position. In fact, that program, initiated in 1958, produced few major findings of direct use to the Air Force until 1965, when our study of the appropriate grade for all Air Force officers reached a reportable stage (10). It is quite true that study was done in response to a direct request. That request could not have been serviced without the prior period of uninterrupted research.

It is not appropriate to refer to personnel research as if it were a univocal process, but it is beguiling to attempt to do so. Has there been any change, has there been any progress between 1941 and 1966? If so, what?

There are two factors that make such exposition difficult. The first such factor is people; the second is history. By people, it is meant that different persons in positions of authority operate from different policies, and emphasis is placed in certain areas as a function of the bias of the leader. Being in a governmental unit, we sway with the politics of the party in power; for four years we emphasize in-service research, with definite pressure to broaden our program and make it more genuinely basic science; for the next four years we minimize the in-service capability, contract for development work, and concentrate on supporting the immediate needs of weapons systems. Although such pressures have been felt, their effect within the Laboratory has not been marked. This is true because of the ponderous inertia of governmental budgetary procedures. It takes effort to induce much change in the budgetary configuration of an operation of small size, as we have been. The stability of our key staff has tended to introduce and maintain a desirable balance in the way we conduct our business.

The second factor making difficult the assessment of the evolution of the program in personnel research lies in history. If in a vacuum, with constant economic framework, with unchanging military commitments, without technological change in weapons, vehicles, computers, and communications, the evolution of the process could be observed. Since 1941 several factors external to our control have impacted on the program. First, World War II ended. A number of projects, mostly focused on development of methods of performance evaluation were inconclusively abandoned. The manning of the aviation psychology program was reduced from several hundreds to about five persons.

Starting in 1946 John T. Cowles and John T. Dailey began to assemble a peace-time research capability in military psychology. Their program grew and by 1947 monitored an aircrew testing program fostered by Reuben Baer and developed paper-and-pencil classification tests for enlisted Air Force personnel. Effort was spent identifying and elaborating measures of officer quality and broadening the coverage of aptitudes for enlisted persons by closer attention to personality factors, with data collected through biographical surveys rather than from classic personality scales. Tests were scored by machine; scores were key-punched into IBM cards; validation studies were run by computing basic summary data on IBM tabulators, then computing the Pearson correlations by hand, with desk calculators, and multiple correlations in the same way. Pedestrian and primitive by standards of today, it was a forward step from calling off scores from answer sheets to prepare scatter plots for the individual hand computation of Pearson correlations as was done early in World War II.

And then the Cold War led into the Korean conflict. Under the liberal manning policies thereby induced we found a considerable increase in Air Force psychologists. The Air Force psychology program grew rapidly and came to include research programs in operator training procedures, crew research,

maintenance techniques, methods for the use of radar observers, and a host of others. The Personnel Laboratory program grew to include research in manpower, a program in motivation and personality measurement, a program in aptitude test research, and improved support functions in experimental test administration and data analysis.

In 1957 the larger organization was curtailed, leaving the Personnel Laboratory with a program of aptitude and evaluation research, and beginning a new program of occupational analysis.

And then, one day in August of 1958, a not very prepossessing truck backed up to building T9016* at Lackland and unloaded a few tons of hardware known as an IBM 650 Data Processing System. We didn't know it then, and we didn't recognize it for the next five years, but that was a most significant date in the Laboratory's history. We were like the cave man the day he discovered a wheel. Our load was marvelously lightened; our capacity was immeasurably increased.

At first our new computer was used as a large desk calculator. We are not unique in this respect–every organization with its first computer has gone through the cycle of automating the previous operation and of being slow to understand the true power of the computer and to apply that power economically to their need. But there had been, as there are still, men of insight in our program, and in the Air Force Personnel System. John T. Dailey, in 1948, had contracted with David Votaw (11), a mathematician at Yale, to study the personnel assignment problem, the best fit of N men in N jobs with attention to all characteristics of the men and of the jobs. Earlier Robert L. Thorndike (12) had worked on the same problems; most of the reaction to their work was amazement at the variety of symbols in the equations in their reports. In the early 1950's, John Leiman had contracted with Carl Kossack to contemplate Markov chains in the light of the flow of personnel through a system.

As we moved into the 1960's, we came of age. We modernized our computing procedures as urged by Joe Ward and Robert Bottenberg. Our analytic procedures moved five years ahead of the profession then and are five years ahead of it still, if the papers read at professional meetings are to be trusted. We have developed machine applications for the assignment models and flow models contemplated by Votaw, Kossack, and Thorndike. John Merck and Frank Ford of our staff have polished them and supervised the detail work of programming them for the computer, which was upgraded to 7040 in 1962.

More than this Joe Ward and Robert Bottenberg have gathered themselves, together with their native intellect, the analyses they encountered on the way to their doctoral degrees, the problems of a modern personnel system, and the characteristics of the high-speed electronic computer, and they have furnished the profession with a more powerful tool for the accomplishment of personnel research than was dreamed of in 1941, or 1946, or 1960 (13). Their application of the multiple linear regression model as a tool for behavioral research represents a major breakthrough in the state of the art. Precise evaluation of the role played in any personnel action by all the variables incident to that action is one manifestation of the power of the model. Data for selection of an optimal set of selection variables to meet several criterion groups is another.

As is often the case in research, minor elements of bygone programs are vital to current major efforts. During and since World War II, every bit of data or personal information on a member of the Air Force that the Laboratory happened to acquire in either a research program or from support of an operational program has been retained. Records Management types have complained of our failure to discard records more than a certain number of days or years old. Laboratory commanders have striven manfully to clean the Augean stables of our archives; "throw the junk away" has been the cry. And we did throw some few things away–after we had copied them on magnetic tape.

---

*Ed. Note.–Official designation for the building which housed the Personnel Research Laboratory and housed many of its antecedent forms.

But in November of 1966, the old data are suddenly a treasure beyond price. It is possible to trace the relationships between selection scores and the actions of persons who have profoundly distinguished themselves, in either a positive or a negative way. Longitudinal studies involving selection variables and career progression become matter of course. The data bank has been augmented by the collection of information from active Air Force personnel records and by compiling them into chronological histories. Mated with aptitude, training, and career progression data, the computer can now accomplish studies of populations, not samples.

A flow model in a computer, showing numbers of persons in every category of interest, permitting the Deputy Chief of Staff for Personnel to wargame such policy changes as he may elect, cannot exist without some mystic values called transitional probabilities. It would be possible to estimate the numbers of, for example, staff sergeants in the administrative career field who have exactly seven years service who will, during the next year, become technical sergeants, remain staff sergeants, accept discharges, be killed, get married, have another dependent, switch career fields, and so forth. When the estimate is complete, it will be data of a sort, but accuracy is bound to be low. With the data bank one can develop an empirical matrix. Then one can say to the general, "Yes, sir, that is what happened."

The combination of high-speed computational facility and files in depth on personnel data permit entirely new kinds of endeavor. We have joined our files with the craft of the economist to accomplish studies which have been of financial benefit to every member of the armed services.

The facility engendered by the computer makes small work of arduous computational chores. It can accomplish the simultaneous comparison of two thousand airmen, in all possible pairs, across their recorded performance of 900 tasks in their career field. It can identify groups of airmen of homogeneous duty in a career field, even though scattered throughout the Air Force, and it can write an accurate description of the content of their jobs.

Twenty-five years have seen us move from the validation of a test battery through a process of hand-tallying of scores and the computation of correlations on a scatter plot to an automated data processing capability that can accomplish a thousand man-years of hand-done clerical work within the twinkling of an electronic circuit. The power of investigative resource has gone up millions of times with the advance in computer technology. Personnel research has broadened from the identification of aptitude variables associated with success in training to the modeling of Air Force careers and the accomplishment of demographic studies which make the chore of the personnel planner and administrator an objective science rather than a mystic art. Precision in the identification of the capabilities of persons, and the requirements of positions can now be welded to humanitarian and equitable policies for the effective manning of our military service.

REFERENCES

1. McReynolds, Jane. *Aptitude levels in the enlisted manpower pool of the Air Force.* WADC Technical Note 58-63 (I) and 58-63 (II), Personnel Research Laboratory, Lackland Air Force Base, Texas, 1958.
2. Wiley, L.N. *Relation of job qualification ratings to performance ratings of basic training instructors.* Technical Documentary Report 64-21, Personnel Research Laboratory, Lackland Air Force Base, Texas, 1964.
3. Wiley, L.N. Developments in analyzing airmen job performance. *Proceedings, 12th Annual Air Force Science & Engineering Symposium, Vol. II.* Aerospace medical Division, Brooks Air Force Base, Texas, 1965.

4. Cowles, J.T., Dailey, J.T., & Keller, R.J. *Development of evaluation and predictive measures in the AF Officer Candidate School.* Headquarters Air Training Command, Barksdale Air Force Base, La., Research Bulletin 47-1, 1947.

5. McClure, G.E., Tupes, E.C., & Dailey, J.T. *Research on criteria of officer effectiveness.* Human Resources Research Center, Lackland Air Force Base, Texas, Research Bulletin 51-8, 1951.

6. Tupes, E.C. & Christal, R.E. *Recurrent personality factors based on trait ratings.* Technical Report 61-97, Personnel Research Laboratory, Lackland Air Force Base, Texas, 1961.

7. Norman, W.T. *Problems of response contamination in personality assessment.* Technical Note 61-43, Personnel Research Laboratory, Lackland Air Force Base, Texas, 1961.

8. Norman, W.T. *Development of self-report tests to measure personality factors identified from peer nominations.* Technical Note 61-44, Personnel Research Laboratory, Lackland Air Force Base, Texas, 1961.

9. Norman, W.T. *Validation of personality tests as measures of trait-rating factors.* Technical Documentary Report 62-4, Personnel Research Laboratory, Lackland Air Force Base, Texas, 1962.

10. Christal, R.E. *Officer grade requirements project: I. Overview.* Technical Report 65-15, Personnel Research Laboratory, Lackland Air Force Base, Texas, 1965.

11. Votaw, D.F., Jr. & Dailey, J.T. Assignment of personnel to jobs. Research Bulletin 52-24, Human Resources Research Center, Lackland Air Force Base, Texas, 1952.

12. Thorndike, R.L. *Research problems and techniques.* Washington: U.S. Government Printing Office, 1947, 121-127.

13. Bottenberg, R.A. & Ward, J.H., Jr. *Applied multiple linear regression.* Technical Documentary Report 63-6, Personnel Research Laboratory, Lackland Air Force Base, Texas, 1963.

XIII

OPPORTUNITIES FOR RESEARCH IN THE MILITARY SERVICES

Ernest J. McCormick
Purdue University

I believe I was targeted for this topic because I have not learned the art of keeping my mouth shut. In particular, I guess I have come to resemble a cracked record, having repeated on various occasions the same theme—namely, my deep-down conviction that the opportunities for personnel and related research in the military services are significantly greater than they are in industry or universities. Basically, this conviction is predicated upon two factors: first, the very respectable, and occasionally outstanding, achievements in this area that the various services have racked up in the past; and second, the psychologist's usual assumption that future behavior can best be predicted on the basis of past behavior.

As a point of departure, then, I would like to offer the opinion that many of the landmarks in personnel research and related areas have been made by the military services themselves, or have evolved from beachheads that the military services have established in their research activities. I hasten to add, however, that the road to these accomplishments has not always been a freeway; it has frequently been mountainous, curved, and strewn with rocks. (I will return to a discussion of the rocks later on.)

If you will pardon some reminiscences about the obvious, I would like to cite a few "for instances" of the significant research contributions that have been made by the military services.

Dominating the landscape of the past is of course the wide assortment of innovations made by the military services in the broad area of personnel classification and assignment. Basic to this effort is the requirement for the measurement of human attributes by the use of tests, scales, questionnaires, and related devices. In this connection, there is hardly a text in tests and measurements that does not, in its introductory chapters, cite the Army Alpha and Beta examinations of World War I as the first significant step in the whole domain of group tests. Concurrent with the development of these tests was the development of procedures involving their use on a wide scale in the classification and assignment of Army personnel. Although this quantum step in personnel selection and placement procedures served as the springboard for the development of corresponding tests and procedures in industry some years later, its impact within the military services was relatively nominal for the following couple of decades during which the military services, as such, passed through a period of partial eclipse. The experience developed during World War I, however, plus intervening developments of test technology in the civilian sector, plus that of certain Government agencies such as the U. S. Employment Service, provided the backlog of experience and technology that made it possible for the Army and Navy in World War II to embark immediately upon what was then the most extensive program in personnel classification and assignment the world had ever known. Despite the criticisms to the contrary, both well-founded and not, it is at least my own personal conviction that those classification programs contributed significantly to the effectiveness with which the military services girded themselves for that unhappy period. Among the several very effective programs during those days, for example, was the (then) U. S. Air Corps program for pilot selection, an unusually comprehensive and thorough test-validation program initiated by Dr. Flanagan and his fellow cohorts here at Lackland Air Force Base.

Aside from the demonstration of the effectiveness of personnel classification programs during, and since, World War II, the research staffs of the military services have made significant contributions to the entire domain of tests and measurements in various ways: for example, in the development of a wide variety of tests as such; in the domain of what I will call test technology (such as the contributions of the Army Personnel Research Office); and in special, unique approaches to the use of tests, such as in validating tests for certain subgroups of subjects. As a logical followup of the original classification programs of the military services, the three services are now applying computer technology to the personnel classification and assignment problem in the assignment of recruits to various types of training and duty. I might add, parenthetically, that such elaborations place a much heavier burden upon the researcher than the now archaic scheme of having such personnel decisions made by people.

Another facet of personnel research in which the military services have had a dominant input is that of personnel appraisal and rating methodology. The work of the military services in this particular area was initiated by the Army during World War I with the rating system that was developed by Scott and his associates. Another milestone was the development of the forced-choice methodology, sparked by Richardson and others of what is now the Army Personnel Research Office, and pursued further by others, including Berkshire (a former member of the staff of the PRL, now with the Naval Aerospace Medical Institute). The forced-choice rating technology has generated over the years a chain reaction into other areas such as personality measurement, attitude measurement, etc.

Related to performance appraisal are other aspects of performance measurement. In this connection the military services have explored, more than anyone else, the use of performance tests of various types, both for research and administrative use. The Army, for example, is in the last stages of completing an ambitious research undertaking that has involved a testing program in which 4,000 lieutenants were ordered to a test site for a week of situational performance tests.

Still another research area in which the military services can take a front seat is that of training. Collectively, there is vast ignorance about the human learning process in real-life situations. However, the military services have done more than their fair share in reducing this void. Perhaps a very significant, although not spectacular, forward step along those lines has been the development of training content and methods which are designed to fulfill actual job requirements. A couple of examples of this are those of HumRRO in their TRAINFIRE program (dealing with leadership training). A related effort with an additional twist is the research program at the U. S. Naval Personnel Research Activity, San Diego, reported by Dr. Fields, in which the objectives were (1) to train as electronics technicians new recruits whose aptitudes were lower than those usually selected for such training, (2) to train them in a shorter time, and (3) to train them so their skills would be more immediately useful on the job. Evidence to date indicates that there are no differences in the shipboard job performance of the experimentally trained men and those in a control group, although time may later reveal differences in promotions, performance in advanced schools, etc.

Various aspects of personnel research have dependence upon some type of job analysis information. Job analysis, of course, has been something a bugaboo to some personnel research efforts, since the problem of quantification or categorization of job information has been a fairly knotty one. But here, too, the military services have ventured to tread, with reasonably significant results. The pioneering work of the Personnel Research Laboratory (whose birthday we are celebrating) is particularly noteworthy. They have developed and used job inventories, consisting of lists of job tasks in a given career field, as a means of mass collection of job information directly from job incumbents. This technique, in combination with certain associated statistical procedures (which they have also developed), has made it possible to describe jobs in quantitative terms and to express job relationships in quantitative terms.

Although our primary focus in these meetings is that of personnel research, the title of this Symposium is Personnel Research and Systems Advancement. The "system" is the framework within which "personnel" perform their functions. The last couple of decades have been characterized in part by a frenzy of attention to the "systems approach" – both to its faddish aspects as well as to its real substance. An important facet of this "new look" at the context within which people work is the application of systematic efforts to design the physical equipment people use and the environments within which people work in terms of human considerations. Although these objectives are as old as man, the development of human factors engineering brought a systematic, scientific approach to this effort. The major thrust of this effort must be credited to the military services and, thus, must be chalked up as another very significant contribution of the military services. A very major part of the research in human factors engineering in the years since World War II has been carried out or sponsored by the military laboratories, such as the Air Force Behavioral Research Laboratories, the Naval Research Laboratory, and the Army Human Engineering Laboratory.

A logical evolutionary step developing out of this new look was the notion of the personnel subsystem, also a contribution of the military services. This concept has served as a catalyst to lend cohesion and integration to the two important facets of application of human effort to the achievement of stated objectives – namely, on the one hand, the personnel aspects and, on the other hand, the design of the facilities and equipment (i.e., the physical systems) that are provided for people to use. For too long these two sides of the coin have been treated independently. The personnel psychologist, for example, has acted as though the work situation (i.e., the equipment, the procedures, the work space, the environment) were God-given and not subject to change. Starting from this premise, his objectives were those of selecting personnel, and then training, supervising, and motivating personnel within this pre-ordained job situation. The early approaches to human factors engineering were equally lopsided in the other direction, being directed toward designing work facilities and equipment so anyone could perform the required tasks, thus brushing entirely aside the required need for consideration of individual differences, the thing most dear to the heart of the personnel psychologist. In its optimum form, the "systems approach" to the development of facilities and equipment does, indeed, take into account both sides of the coin. It is specifically the concept of the personnel subsystem, however, that tends to weld the two sides into (if you will pardon a mixed metaphor) a single, integrated ball of wax. In the development of this frame of reference, research psychologists and other personnel of the military services have played a major part.

In discussing this sort of "desegregated" or "integrated" way of looking at human beings in systems, I do not want to imply that this "new look" is now standard practice, or that completely adequate methods and techniques of analysis have been developed, or that substantive principles and guidelines have been developed to solve all of our problems. But the walls of Jericho–that have tended toward compartmentalized approaches to human problems in the military services and industry–these walls at least have some cracks in them, these cracks having been started by the military services.

Another major research effort that has major overtones relating to human performance is that of CDEC, the Combat Development Experimentation Center at Ft. Ord, California. This center is concerned with experimentation with military doctrine, tactics, organizations, etc. Their research relating to simulated military operations of various types is carried out in a grand scale over miles of terrain. This is the type of experimental opportunity that could cause a pragmatically oriented experimental psychologist to drool more than Pavlov's dog.

One could cite other examples of research contributions of the military services to the field of personnel and systems research that might be considered to be "significant," as in the study of human stress, leadership, or group performance, or in the development of criteria and statistical techniques. Various such efforts might be of the nature of landmarks or breakthroughs, or of the nature of interesting innovations, or of the nature of door-openers that shed a ray of light on a previously hidden nook or

cranny, or of the nature of yeast that stimulates research fermentation. But aside from these "significant" contributions, the military services have, of course, produced a massive volume of good, solid, substantive (though perhaps not earth-shaking) research. Actually, most of their research undertakings probably are of this nature. Obviously only a few people on the merry-go-round can catch the brass ring. Within the sub-world in which you and I live, there is need for un-glamorous, un-spectacular, but painstaking, thorough research to achieve certain objectives. And one must also give the military laboratories both an "A" for effort and an "A" for achievement in continually and diligently carrying out these unsung responsibilities. In addition, the military services must acknowledge responsibility for perpetrating their share of the (unfortunate) occasional examples of ill-conceived, poorly designed, ineptly executed research projects, and of those elegant, well-executed projects that deal with trivial topics.

The representatives of the various services here have reported on their on-going research programs. But perhaps these examples will illustrate my central point of the moment–that, indeed, the military services have collectively served as pioneers in the field of personnel research and systems analysis and, in addition, have produced over the years a very extensive, respectable body of pragmatically oriented research in these ballparks.

In connection with the research approach used in the various research programs, it should be added that, in the case of certain specific programs, the research efforts sort of move across the spectrum of basic and applied research, although the direction of this shift can be either way. For example, in a HumRRO study of tactile communications, the research approach started from the springboard of the basic work of Geldard and others and is moving to the actual development of a communication code and hardware to transmit electropulses to the skin. On the other hand, another HumRRO program, dealing with performance under stress, started at the empirical end of the spectrum and has culminated in the development of a theoretical formulation–namely, a conceptual model of behavior under stress.

The wide spectrum of substantive research of the military services is, of course, not a fortuity. Rather, it is the consequence of various features that are characteristic of the research laboratories. High on the list of such features are the personnel of the laboratories, including both the research directors and their staffs. Most of these individuals are persons of significant stature, ability, and dedication. Another feature of behavioral research in the military services is the ready availability of a never-ending supply of, shall I say, "volunteer" subjects–such as to cause other researchers to turn green with envy. Further, the subjects frequently can be induced to return again, the next day, or the next week, for retesting or other experimental procedures. With respect to subjects, the services over the years build up extensive longitudinal information about individuals, covering the span of recruiting, training, work experience of various types, through termination. These data frequently are in the form of automated personnel records, readily available for use. This career-span of data also makes it possible to have access to, or to develop, criterion information of an intermediate or ultimate nature that is usually difficult to come by in typical industrial or university situations. And in certain circumstances (unfortunately not all), the physical equipment available for research is unexcelled.

Earlier I said I would be discussing some of the "rocks" in the road. Let me mention a few of them now. Looking at the research activities of the military services back over the years, there have been periods of lack of administrative (i.e., financial) support; there have been times when an organization has worked with an axe poised over it–and times when the axe has actually fallen; there have been times when an organization has suffered under a new commanding officer with a new broom, or a commander who perceives the activities as a lot of folderol; there have been organizational changes and reorganizations and re-reorganizations, with agencies being split assunder and scattered to the four winds or shifted to some new (and occasionally very surprised) parent organization; there have been congressmen and newspaper reporters who have raised questions about certain research projects, thereby by implication raising questions about entire programs; and always the work is carried out in glass-walled buildings without

window shades, open to public view. I am not necessarily saying that the researchers in the military services have more problems than anybody, but they certainly have had their share, and I think it is doubly significant that the major accomplishments of the research laboratories have been achieved in the presence of some of these kinds of problems.

This, of course, leads to the question as to the factors that have nurtured such research activities in the midst of the occasional (or frequent) drought, hailstorms, changes in climate, etc. I think that this can be attributed to various factors, but I believe that the critical factor is the fact that the research programs basically have been focused on the practical human problems of the military services, and further, the fact that the products of such research have been perceived as actually reducing or even solving, at least some of these problems. Even the frequent rocking of the boat over the years--distressing as it may be, especially to those in the boat--does not detract significantly from the fact that personnel and systems research has effectively demonstrated to the services its virtually indispensible quality. At this point in time, I believe it can be said with reasonable confidence that the concept of research as an essential aspect of a personnel program or of a systems development program is indeed intrinsically ingrained within the fabric of military operations. It is this conviction that lends confidence to the prospects of generally continued support of such research by the military services.

This is not to say that there is universal and enthusiastic support for personnel research in all branches and at all echelons of the services. Most of us here would wish that such support would be more universal. But basically the services are now so firmly committed, and I think rightly so, to personnel research that it is inconceivable that there could be any significant backtracking. Most of us here probably would concur in the opinion of one general officer who has stated that there is greater payoff, dollar for dollar, from personnel and human factors research than from the other dollars.

Looking ahead, I would envision the need for even greater dependence upon behavioral research as we move toward more and more complex military systems.

Needless to say, then, we cannot relax in the face of the challenges of the future. Let me just suggest a few. For example, as I implied earlier, the prospect of using computers in personnel classification and assignment and in reassignment places an increased burden upon the personnel psychologists to come up with valid information to crank into the computers in carrying out these operations. Computers are much more efficient in perpetrating errors than are people, and if people feed them incorrect information to begin with, the computers can be very efficient in producing wrong answers. In addition, I would envision the development or elaboration of research requirements relevant to such problems as the development of job structures and career ladders; total career utilization of personnel, including reassignment policies; the development of organizational structures and military doctrine that tend to optimize the use of human talent; the development and design of systems that are optimum in terms of multiple criteria or considerations, such as system performance per se, human abilities, training costs, personnel career development objectives, manpower utilization considerations, etc.; and long-range manpower procurement policies, designed to minimize imbalances in the "flow" of manpower through the military systems. But Dr. Wallace, the anchor man of this Symposium, will tell us about the future, using his crystal ball. My point here is that what now appear, in a very hazy way, to be the future research requirements of the military services, are going to provide opportunities galore for the behavioral sciences. But the extent to which research objectives can be fulfilled depends, of course, on the constraints and boundaries within which the activities are carried out. Virtually every research effort has some sort of constraint imposed upon it; in some circumstances these are more annoying than impeding, but in other circumstances they can be seriously impeding. Within the existing research laboratories of the military services these constraints, in various situations, cover the gamut of poor physical facilities, inadequate experimental equipment, and unfortunate personnel practices. For example, there are certain artifactual aspects of the personnel classification system and restrictions on promotional possibilities that tend to impose lower "ceilings" on

research personnel than might otherwise be the case. The occasional changes in the prevailing temperature–that is, the blowing hot and the blowing cold of the budgeteers on high–of course, constitute a source of worry and concern. Obviously, one never achieves Nirvana in this world, and even the prima donnas of the research community should not expect to be relieved of all sources of worry, irritation, and difficulty in carrying out their functions. However, it would seem in order for those responsible for administering research undertakings at various levels to keep their antennas continually tuned in, in order to be able to sense those conditions and circumstances in research activities that might significantly impair their carrying out of their functions.

In discussing personnel and systems research, I have been talking primarily about the closed system of the military services as such. In general the primary focus of the services must be in terms of their own requirements. But the military services also have a vested interest in the world about them. Research, of course, knows no organizational bounds, and the interaction between research in and out of the military services contributes mutually to both sides of the fence row. This interaction is manifest in the significant effects of military research in personnel and systems research on the civilian community. This by-product effect in the behavioral sciences is something akin to the effects of military and space research in the physical sciences in the form of new and improved civilian products. And the military services, in turn, have benefited from the reverse flow. In this interchange, it is to the credit of the military services that they have, in overt ways, encouraged and promoted this interchange, such as in the research they have sponsored in universities and elsewhere (e.g., the program of the Office of Naval Research). Thus, much of the research carried out within the military establishment, although specifically addressed to the pragmatic problems of the military, will also contribute gradually to the better understanding of human behavior in its various manifestations and, thus, help to promote human welfare in general.

After these random musings I should somehow get myself back to the track I started on–namely, the opportunities for research in the military services. In contemplating this topic myself, I must confess to having viewed it through a one-way mirror, namely, that of viewing the personnel and systems research milieu from the point of view, for example, of prospective employees, such as new Ph.D.s who are considering various employment opportunities. From this point of view I have said, and will say again, that, considering both positive and negative valences, I think the opportunities for research within the military services are greater than those in industry, or even in universities.

But then, all of a sudden–as with a reversible illusion such as the classic staircase–it dawned on me that this same phrase, "the opportunities for research in the military services," could also be viewed from the point of view of the military services themselves. From such a point of view, one could pose the question in various ways: for example, "What are the opportunities within the military services for carrying out behavioral research that might contribute to the solution of some of their human problems?" or, "What are the human problems, now or in the future, about which we (speaking for the military services) have some opportunity to do something–via the behavioral research route? Have we (the military services) been pulling answers to human problems out of thin air, or making decisions on the basis of virtually zero information, when we could through research obtain some data that might increase significantly the possibility of making better decisions?" Obviously my "instant insight" is old hat to the military services, for they have of course asked themselves these questions in one form or another thousands of times, and have answered them themselves, usually in some affirmative way; otherwise, such organizations as the PRL would not have any birthdays to celebrate.

Although my reversible twist-of-a-phrase about "opportunities" for research really is not very clever, I would like to say, in conclusion, that I believe it does relate to a basic verity. In particular, I believe that the "establishment" (i.e., the military services) and the research personnel of the laboratories have the same resonant frequency–that there is sort of a coincidence or unity of mutual objectives that has been a source

of research vitality and productivity over many years. On the one hand, any research psychologist who gets his kicks out of dealing with the hard, knotty, difficult, practical problems of the real world can find the problems and the opportunities to work on them in the military services. And on the other hand, the military services can facilitate the solution to human problems by being alert to the opportunities to bring research approaches to bear upon them when such action will likely contribute information that is relevant to their decision-making processes.

XIV

RESEARCH INTO ACTION

Robert E. Murphy
Aerospace Medical Division
Department of the Air Force

Twenty-five years ago, over my objections and those of my Commander, I was transferred from the neat and orderly life of an infantry sergeant, where I at least felt as if I knew what I was doing most of the time, to San Antonio, Texas and the considerable confusion attendant to setting up a psychological testing unit. The basis of this transfer was that I, enroute to a degree in history, had taken a course or two in elementary psychology. Within a week after reporting for duty, I was given a one-hour on-the-job training course and assigned as the noncommissioned officer in charge of group testing. If you think I am exaggerating the confusion, let me tell you of just one incident. Each day, two hundred men were marched up to an open bay testing building where they were seated in alphabetical order and given a day-long session of group tests. This two hundred-man group had no designation–there were no rosters–just a bunch of boys chosen from the cadet "pool." One morning, after I had been in charge for about a week, nobody appeared for testing, and I called Major Robert Rock to report the missing men. He advised me that there was a little delay in getting the men out of the pool, but that my two hundred men would be along in a few minutes. About five minutes later, as predicted, about two hundred men, with a corporal in charge, came down the street toward my building. I stopped them in front of the building, went through the usual procedure of filing them into their places in alphabetical order, and then started the testing. About two hours later, a very bewildered Lt Colonel peered into the window, his face registering disbelief, then fury. He called me outside and, after bracing me properly, told me that the men being tested were not aviation cadet candidates at all. I had intercepted his group of enlisted men who were marching down to the railroad siding for shipment to Keesler Field. The train had departed with all their baggage and records. Incidentally, I have never found out if the men and their paraphernalia ever got together. It was a turning point in my life, however, for immediately I decided that it was an appropriate time for me to apply for Officer Candidate School.

In discussing the topic which was suggested for me, "Research Into Action," I do not intend to draw any fine lines between basic and applied research and the category of technical development. With regard to the Personnel Research Laboratory, it is often difficult to trace the Laboratory products back to the original ideas or individuals responsible. The payoffs to the Air Force have resulted from the ideas and efforts of many people over, sometimes, long periods of continuing refinements.

Many of you who have been associated with this Laboratory remember the annual program reviews at higher headquarters when the "show me" boys would ask, "How is what you propose to do in your research program going to help the Air Force do its job? What's the payoff in terms of specific action which can be taken?" These are good questions which required asking but were not always easy to answer at the time. Now let us cover a few years in retrospect and see if we can now answer some of the "payoff in action" questions which were difficult to answer a few years ago.

In the middle fifties, Joe Ward, Bob Buttenberg, John Merck, and Frank Ford were pushing a program of mathematical model development. The concepts were abstract and not of apparent relevance to problems which existed at the time in the management of the Air Force Personnel System. It was pointed out that such models to the extent that they simulated a personnel system, if only in part, might provide information useful to personnel or manpower planners. There was some difference of opinion at the time as to whether the effort represented something the Air Force should support or merely a "hobby shop" to satisfy the interests of a few scientists. In any event, the research continued to receive the support of those in management who recognized that instant usefulness has not always been the best criterion by which to judge a research project.

One of the interests at the time was in the Markov chain and its application to personnel movement from one state or category to another. This resulted in the Personnel Flow Model which appeared to have real potential if one were able to get good estimates of the transitional probabilities of movement from one state to others (1). At this point, it was clear that personnel records, though accurately reflecting the current state of individuals, were deficient in a longitudinal respect. Therefore, it was necessary to immediately start a file on Air Force personnel which could be continuously updated so as to provide a means by which transitional probabilities could be empirically determined. The usefulness of the model and the data file were first demonstrated in 1963, when a requirement was established by the Air Force to project or forecast the characteristics of the officer force through the 1967 time period, assuming, of course, that conditions and policies of 1963 remained stable. Based on only limited probability figures accumulated since 1961, the model was used and the projections were forwarded to Headquarters USAF. Followup studies indicated that the projections were very acceptable. It was then evident that this model, with some refinement and improved data files, had great potential for providing information to personnel planners.

The following year, late 1964, the Air Force was considering possible changes in the promotion policies to hasten airman promotions within established grade ceilings. Information was needed on the flow of airmen through the top six grades under current policies and projected into the future. "Project Top Flow" was initiated in the Personnel Research Laboratory to meet this need. Employing the Personnel Flow Model and an improved data base, the characteristics of the airman promotion force were projected over the following five years. From these projections, it was found that the promotional situation would improve under existing policies, and predictions were made as to when improvements would occur. As it turned out, this information was a major input into the decision not to change then existing airman promotion policies.

The usefulness of the flow model has been established now as a powerful instrument for use in personnel planning, and it has been used to accomplish a number of special studies for various Air Force agencies. Just recently, the Air Force has established a requirement for a study of career medical officer retention. In the requirement, it was specifically requested that the flow model be used to provide a means for understanding personnel input and outflow, as well as to permit a more empirical base than presently exists for planning purposes.

An important "peel-off" of this research has been the usefulness of the transitional probabilities themselves, and the Laboratory has responded to many requests for special studies for the analysis of retention rates, cross-training rates, retirement trends, and replacement rates for overseas areas. This research has also pointed up the need for better longitudinal files on Air Force personnel, and the characteristics of specifications for such files are now being studied.

In the early fifties, there was much interest on the part of a number of people in the Personnel Research Laboratory in the application of regression analysis to the solution of Air Force personnel problems. At the same time, there was a special interest in developing techniques for optimizing the

assignment of Air Force personnel. Out of these two interests grew two models: the "assignment model" and the "policy- capturing model" (2; 3; 4; 5). The assignment model requires a judgment as to how well an individual with a set of known characteristics fits a job with a set of known requirements. Then given a group of people and a number of jobs, it is possible to assign people to jobs so as to optimize the assignments of the total group. The usefulness of the model was limited by the number of judgments required in any practical situation. This limitation led to the development of the "policy-capturing model." This model is a special case of the regression model which, on a sampling basis, derives an equation which closely approximates the relative weights of variables used by a panel of judges in assigning people to jobs. With the policy of the judges thus "captured" and programmed for the computer, it appeared feasible to automate assignment procedures. The feasibility of automated assignments was tested last year, 1965, when on an experimental basis for a sizable sample of airmen, computer assignments were compared with the traditional assignment procedures. The results were encouraging and, after some refinements, a further series of tests was run last July and August when four groups of approximately 3,000 airmen at Lackland Air Force Base were computer- assigned to satisfy weekly requirements for technical schools and direct duty assignments. The weights used in the model were those obtained from a sample of 400 airmen actually assigned by career counselors of the Personnel Processing Squadron. The computer assignments were checked against actual assignments of the same airmen by career counselors and reviewed by experts in the Personnel Processing Squadron. The results were so encouraging that computer assignments have been requested by the Personnel Processing Squadron and are being provided on a weekly basis as a major input into assignment procedures. I should mention here that the computer assignments take into account the preferences expressed by the airmen for assignment weighted as the counselors weight the preferences, which by the way, is quite heavy. The model has, of course, limitations in assigning special cases and for special fields. Currently, the initial assignment procedures for airmen at Lackland are being modified so as to take advantage of the computer assistance which can be provided by the assignment model.

The policy-capturing model, which was originally developed to support or reduce the limitations of the assignment model, has become a most useful tool in permitting a new look at some old problems in the personnel and manpower fields.

In 1963, the Director of Manpower and Organization, Headquarters USAF, established a requirement for a study of the officer grade structure of the Air Force. He pointed out that grade distributions traditionally have been based on the judgments of operating officials, manpower officers, and the broad guidance contained in the Air Force manual which concerns manpower policies, procedures, and criteria. He requested the assistance of the Personnel Research Laboratory in an analysis of grade requirements based on actual job requirements. Dr. Raymond Christal and his associates, in considering various approaches, decided that the policy-capturing model would be an appropriate means of providing this new look at officer grade distributions (6). To implement the necessary procedures, detailed descriptions of 3,575 actual Air Force officer jobs for every career utilization field were prepared. A board of 22 colonels, chaired by a general officer, was assembled in Washington. The board's task was to assign to each job what it considered to be an appropriate grade. Not fewer than five board members independently rated each job with very high agreement. An equation was developed which reflected the policy of this board, predicting its judgments with a validity of .92. This equation was then applied to the assessment of some 130,000 officer jobs in the Air Force and an appropriate grade distribution determined, based on the policy of the criterion board. The Director of Manpower and Organization, Headquarters USAF, at the time stated in a letter, "In this particular era of United States National Defense, this project in my opinion, is of first importance."

The mathematical models built, refined, and tested in the "hobby shops" of the Laboratory during the last ten years have had substantial payoffs in terms of providing inputs to decision makers for specific actions. The payoffs or applications certainly were not apparent, nor could they be conceived of at the time the research was initiated. One of my colleagues, a Texan, puts it this way, "In research, the situation for

the scientist is much like that of drilling for oil. You don't know whether you are four feet away from a million dollars or a million feet away from four dollars. Only time will tell!"

In the late 1950's and early 1960's, considerable concern was expressed over the rather high incidence of airmen discharged for unsuitability. The problem was felt to be particularly acute in certain areas, for instance, the nuclear weapons and intelligence career fields. For some years before this time, Dr. Eli Flyer of the Personnel Research Laboratory had been carrying on research to identify variables predictive of unsuitability discharges (7). He found, for example, that high school graduates had much lower unsuitability discharge rates than did nongraduates at all levels of aptitude. These findings were taken into account in establishing airman enlistment standards. This research pointed up the need for further information in areas which was not available in the records and which could be obtained reliably at a reasonable cost. It was found that pre-service adjustment factors which were relevant to assignment to high-risk jobs could be identified from information supplied in pre-assignment self reports and from reports by parents, teachers, and employers. Procedures for obtaining these reports were developed and the feasibility of obtaining them on a systematic basis was established. A direct outgrowth of this effort is the Assessment Branch of the Personnel Processing Squadron on Lackland Air Force Base. Established in January 1966, this organization screens some 17,000 new recruits annually for assignment to sensitive career areas, implementing in a fully operational sense the procedures and techniques developed in the Personnel Research Laboratory.

As an Air Force officer who has spent most of his career in the administration or management of some aspect of the Air Force psychological research program, I have been constantly kept aware of the need to demonstrate the usefulness or application of our research products to the solution of real Air Force problems. While I recognize the requirement for continuing critical reviews of programs, and appreciate the need to place manpower and dollar resources in support of programs which have the greatest payoff for the Air Force, I must confess that there have been times in the past when I have felt as if I, and indeed the research program, had been placed on trial.

I feel that it is a tribute to the courage, the dedication, and just plain hard work of those now in the Laboratory and those of you who have been formerly associated with the Laboratory that the research program of the Personnel Research Laboratory is no longer on trial. Its contributions are recognized; its products are in demand; the support provided by higher headquarters is strong. In this healthy atmosphere, I look forward to the next twenty-five years of even greater productivity and sincerely hope that you will invite me to speak on the same subject on your Fiftieth Anniversary.

REFERENCES

1. Merck, J.W. *A Markovian model for projecting movements of personnel through a system.* Technical Report 65-6, Personnel Research Laboratory, Lackland Air Force Base, Texas, 1965.
2. Bottenberg, R.A. & Ward, J.H., Jr. *Applied multiple linear regression.* Technical Documentary Report 63-6, Personnel Research Labbratory, Lackland Air Force Base, Texas, 1963.
3. Votaw, D.F., Jr. *Functional tests of solutions of personnel assignment problems.* WADC Technical Report 59-358, Aerospace Medical Research Laboratories, August 1959.

4. **Ward, J.H., Jr. & Davis, Kathleen.** *Teaching a digital computer to assist in making decisions.* Technical Documentary Report 63-16, Personnel Research Laboratory, Lackland Air Force Base, Texas, 1963.

5. **Christal, R.E.** *JAN: A technique for analyzing group judgment.* Technical Documentary Report 63-3, Personnel Research Laboratory, Lackland Air Force Base, Texas, 1963.

6. **Christal, R.E.** *Officer grade requirements project: I. Overview.* Technical Report 65-15, Personnel Research Laboratory, Lackland Air Force Base, Texas, 1965.

7. **Flyer, E.S.** *Prediction of unsuitability among first-term airmen from aptitude indexes, high school reference data, and basic training evaluations.* Technical Documentary Report 63-17, Personnel Research Laboratory, Lackland Air Force Base, Texas, 1963.

XV

THE CONCEPT OF A PERSONNEL SYSTEM

Launor F. Carter
System Development Corporation

Introduction

I feel honored to have been asked to participate in this Twenty-fifth Anniversary Symposium on Personnel Research and Systems Advancement. The continuous activities in personnel research and operations of the Air Force Personnel Research Laboratory for a period of twenty-five years is a great tribute to the many people who have served so well at this important personnel research center. As Chairman of the Psychology and Social Sciences Panel of the Air Force Scientific Advisory Board, I have been asked by Dr. Guy Stever, the Chairman of the Board, to express his and the Board's congratulations on the accomplishments of the Personnel Research Laboratory over this long period of time and to wish it continuing success in its endeavor to serve the Air Force.

It is particularly appropriate that the preceding paper in this symposium was given by Dr. John Flanagan, since it was he, more than anyone else, whose great energy and devotion to the concepts of personnel research and the operational application of research findings led to the very valuable achievements of the World War II Aviation Psychology Program and to the solid base on which the further work of this laboratory was able to build. Although many of us who were associated with that early program have gone on to other activities in civilian life, I am sure we all look back on the period of our association in the Air Force Aviation Psychology Program as one of the most productive and useful periods of our lives. We owe much to John Flanagan for having so vigorously and foresightfully led this truly revolutionary application of psychological principles to problems in personnel management.

The committee organizing this symposium assigned me the title "The Concept of a Personnel System." In some ways this is a very fortunate title selection, since had I been left to my own devices, I might well have chosen a much narrower and more limited topic. I would probably have only talked about the possible application of advanced information processing in personnel research and operations, but these topics are being well handled by other speakers of the symposium, and the broader topic of personnel systems gives me the opportunity to range widely.

First, I would like to present some personal thoughts regarding problems of personnel management in the military service. I do this with some trepidation, since I wish to examine some deeply held beliefs regarding the proper philosophy of personnel management in the military services. I hope that as I do so you will appreciate that I raise these questions only because I believe that their consideration and evaluation would benefit the Air Force, both in terms of the efficiency of its operation and its ultimate capability as a fighting force.

My second topic is more closely germane to the narrower conception of a personnel system. Here, I will consider some implications of modern information handling technology for the management of an

individualized personnel system. My general thesis is that the new developments in information processing technology allow personnel management to resolve one of the long-standing problems regarding the centralization and decentralization of personnel management, and at last to make the benefits of both possible through the efficient management of personnel information.

Finally, I will turn to an examination of problems in utilizing research and advanced development in operating systems. I believe this problem to be of great general significance, not only in the personnel area, but throughout the military research and development establishment, as well as in industrial and university institutions.

Some General Considerations Regarding Personnel Management in the Military Services

In the last decade we have seen more and more emphasis on rationalizing management systems. This applies to weapon systems, to industrial organizations, to personnel systems, or, indeed, to any large organization. We are all familiar with the emphasis on system analysis, cost effectiveness, programs-plans and budgets, and the alleged decision making by computer. As individuals we would like to believe that we behave sensibly and that our decisions are made in the light of rational considerations. I want to suggest that in the military, and particularly in the military personnel system, because of complex competing forces many actions are required which appear at the individual level to be based on relatively irrational grounds. At least these actions often appear to be less than optimal when viewed from the point of view of the individual involved or even from the point of view of a given set of major objectives. I would contend that in the management of military personnel there are unusual factors which determine that the management of the personnel system cannot be completely rational. I refer to a number of imposed constraints which seem to be more prevalent in military personnel management than they are in many other situations. Most of us in small organizations try to operate our personnel system within relatively well-understood procedures. Since we are in organizations where there is considerable individual control over the practices of the organizations, it often appears that our personnel practices are relatively well thought out and executed. On the other hand, in a military personnel system there are competing forces such as the various laws which govern the way in which military personnel must be treated with regard to pay, promotion, retirement, and utilization; likewise, there are many organized groups which bring pressure directly on the military personnel management system and indirectly through Congress and other governing organizations. In addition to these factors, there are the shifting role requirements for military personnel. During most of their military careers, military personnel are engaged in activities directed toward training and managing the military establishment, but not in active fighting; and yet for certain periods during the last half century, many military men have been called upon to engage in actual fighting. This creates a very difficult personnel management situation if for no other reason than the relative age distribution of the personnel involved.

We are all familiar with the distortions that took place in the distribution in various grades and ages of military personnel as a result of the integration of large numbers of officers immediately following World War II. These humps in age result in a grade distribution that becomes distorted, and the ability of officers in various ranks and age categories no longer fit the requirements of the various organizational jobs within the Air Force. This situation is about to repeat itself again because of the very radical age distribution change that is taking place within the American population. In a penetrating analysis of this problem, Peter Drucker (1) has recently said, "We face in this country one of the most dramatic shifts in age structure of the population any country has ever gone through. The age group that is now moving into the 20- to 25-year bracket is about 60 per cent larger than the age group just ahead of it, which is about the biggest jump in any five-year group we have ever gone through. In the meantime the people in the 35- to 55-year group will get fewer, not only in total percentages of the work force and population, but in actual numbers.

For the next fifteen years we will have fewer people of what most of us consider the prime-of-life than we have today. Or to put it statistically, within a year or two, maybe three, one-half of the American population will be under 25 years of age. About 29 is the average today, and a four-year drop in three years is almost unheard of. By the mid-seventies at least two-thirds and perhaps three-quarters of the American population will be under 35. This is the youngest we have been since Andrew Jackson and, at the same time, let me say this is probably the most relevant thing a personnel man can know about the next ten years." Factors such as this will have a more profound influence on military personnel management than they will on civilian personnel management because of the relatively inflexible laws and regulations which govern basic military personnel policies.

Another factor which makes military personnel management extremely difficult is that the military is expected to play many national roles in its management of personnel. If the single problem facing the military personnel manager were the creation of a highly effective fighting force, he would certainly have an easier problem than he has when he is expected also to serve the role of one of the major educational institutions in the country. Recently Mr. McNamara, in announcing the program to induct 100,000 Category IV men,* stated that this action was taken as a service to the nation, to be accomplished by training underprivileged persons so that they could assume a more productive role on the national scene. While this goal is undoubtedly laudatory from a broad national point of view, from the narrower perspective of the military personnel manager, this decision clearly complicates his management problems. As was well demonstrated in some of the Army's Fighter Studies, with which I was associated, the better men, both from the point of view of their fighting ability and from the point of view of their ease of management, are those with higher degrees of education, intelligence, and personal adjustment. The induction of Category IV personnel runs counter to the choice that would be made from a purely military point of view and the efficient operation of the military personnel manpower pool. The main point of the above is to emphasize that the problems facing the military personnel manager tend to be more difficult than those of his counterpart in the civilian economy, and that criticism of his management is often less than justified when these additional constraints are taken into consideration.

I would like to turn next to some basic assumptions regarding the proper role and management of military personnel. In many ways, I hesitate to even discuss this subject, since I am sure most of you will disagree with the comments I am about to make. Yet, I believe their consideration is fundamental to the optimum rationalization of military personnel management. Even though current practices are based on long tradition and acceptance by most of those in the military service, I believe it would be useful to re-examine from time to time some of the most fundamentally held ideas regarding the proper role of military personnel. I do not suggest that these comments afford answers to the questions I will raise, but rather I would like to open up a dialogue regarding these subjects and to suggest that they be made subject to serious study and system analysis in an effort to improve the overall quality and ability of the Air Force as a fighting service.

The first concept which I would like to consider is the concept of the fungibility of military officers. I was first introduced to the concept of the fungibility of military personnel in 1941 when, as a civilian, I was employed by the Army's Adjutant General's office in the development of selection procedures. The underlying idea of fungibility is that a military officer is able to fill any role which is appropriate to an officer in the military service. The dictionary defines fungibility as "the designation of goods, as grain, any unit or part of which can replace another unit, as in discharging a debt; capable of being used in place of

* Category IV inductees are those who have scored at low aptitude levels on the Armed Forces Qualifying Test and/or have limited physical inadequacies.

another." As I have indicated, in 1941 the Army held as a fundamental tenet that any officer was fungible in that he could take on the role of any other officer. By and large, this concept has been becoming more and more untenable as the functions of officers in the Air Force, and in the Army, too, become more technical. It is obvious that a non-rated officer cannot take over the role of a pilot, but it is generally assumed that a pilot can take over almost any other role if he has had a small amount of training. I would argue that, as the various jobs being performed in the military become more and more specialized and technical, the concept of fungibility becomes more of an inhibiting philosophy than it does one which is useful as a personnel management concept.

Some will argue that this concept is no longer really followed in the military service, and yet I believe that it is, in fact, one of the fundamental concepts underlying much military personnel management. Just before his retirement, I was discussing this matter with General Schriever.* We were discussing the fact that a number of officer billets in the various Systems Command divisions were seriously undermanned and that officers who had spent many years in the Systems Command were being drained away to the Vietnamese effort. They were being used in ways which did not take advantage of their long experience in the Systems Command divisions and, for some, an extensive and expensive education in preparation for roles in system management and development. While General Schriever regretted the depleted manning of his divisions, he went on to argue that this was really appropriate since an officer always had to be prepared to join any active engagement and although the education and training of these officers was not being used in Vietnam, they had a fundamental obligation to fill their fighting role as a part of the military. One might find this argument more acceptable if, in fact, many of the officers leaving the Systems Command were still on active flying status and could be used as combat pilots; but I suspect that, in fact, most of them are leaving the Systems Command to fill billets where they will not be used in an active fighting capacity but, rather, in connection with the many support activities that must be undertaken in a military situation. I would raise the question as to whether or not it would have been better to recognize the highly specialized nature of the jobs being filled in the Systems Command and also the specialized nature of the jobs that would be required in the case of an operation such as that in Vietnam and to have planned the military careers of each individual officer in terms of the relatively specific requirements in these two different types of situations. This is a particularly difficult point to argue at a time when we are engaged in such an extensive military operation, and yet I feel it is during such periods that it is most profitable to examine the actual performance of personnel who have been brought up in a period where for the last twenty years there have previously been only about two years of actual military engagement. In other words, I am suggesting that there should now be an intensive personnel research effort to assess the ability of personnel who have been matured under a concept of fungibility and peacetime military management. Such an assessment should focus on the extent to which the personnel involved in active conflict are, indeed, able to shift from one type of military role to a different one as demanded by wartime operations.

The second major assumption I think should be examined is the general assumption that every member of the military is fundamentally a fighting man. Almost without exception when I have discussed such matters with senior military personnel, I have found them saying that every member of the military must be prepared to engage in active combat and that this is his fundamental duty. In some broad moral sense I do not disagree with this position, but as a practical matter, it does not seem to me to be an assumption which leads to effective management of a fighting organization. Studies have repeatedly shown that a relatively small number, not more than 5 or 10 per cent, of military personnel ever become involved in actual combat. It seems to me that this fact should be recognized and that personnel management concepts should be fashioned around the inevitable fact that most military personnel will be involved in

* General Bernard Schriever, then Commander of Air Force Systems Command.

duties which are in support of the actual combat fighter. I should immediately make it clear that I am not suggesting that the role of the military man as a fighter should be downgraded. Indeed, if anything, I think it should be upgraded and that very special treatment and special rewards should go to those who actually engage in true combat operations; but, at the same time, I would like to suggest that the military management of the 90 to 95 per cent of military personnel who will not engage in combat should be based on different assumptions and practices than apply to the small proportion who do, in fact, engage in combat. Again, I would suggest that during a time of conflict, the performance of non-combat personnel should be actively assessed and question should be raised as to whether their past training and job assignment have been such as to fit them for the duties which must be filled in the support of combat. Based on such studies the manpower, training, and assignment policies of the military service should be modified to optimize the effectiveness of the service during periods of combat.

Next, I would like to turn to the difficult subject of the proper relationship and management of civilian personnel working in the military establishment, an area which includes the interaction between their management and the management of military personnel. One of my fundamental assumptions is that most people working in an organization aspire to advance their status, either through pay or promotion, and would like to have the opportunity to rise from an entering position into the highest positions within the organization. It seems to me that the management policies for the civilian component of the military establishment confront the civilian employee with the inevitable fact that he cannot rise to a high decision-making level within the military institution. As an extreme example, one can examine the air staff at Headquarters USAF and observe that there is no civilian who has ever been a deputy chief of staff or has held the senior position as comptroller or judge advocate or any of the other major staff positions. Indeed, as one goes through the headquarters organization, one can find very few civilians who are heads of directorates, divisions, or even branches. At times a civilian will be found who is assistant to the head of a major organization, but by long tradition the civilian is denied an opportunity to reach the top. I would argue that this same phenomena can be observed in almost all major command headquarters, as well as echelons below major commands. In other words, the civilian who aspires to a career in the military establishment must decide that he will be content with a relatively limited career progression and that even though he rises to the top of the civilian progression he will always have a military officer as his immediate superior.

I have always been struck by the apparent paradox that at the very top of the military establishment is the civilian secretary and his associates. This small group of civilians accords with our long tradition of assuring the ultimate civilian policy-setting and control under our constitutional forms. Immediately underneath this top echelon of select civilians, however, one finds a complete military group, and this military group spreads down the chain of command and senior positions through a number of layers before any significant number of civilians appear in responsible positions.

It can be argued that the military establishment is primarily for military operations and that military personnel should be responsible for those operations which are directed toward combat operations. It will further be argued that in a military setting the military man should have the highest degree of career possibilities available to him, that is to say, to rise from the lowest ranks to the position of chief of staff. My basic problem is not in disagreement with this point in view. Indeed, I believe that the military man should be able to look forward to a career in which progression to the highest level is possible; but, at the same time, I believe the civilian should be able to look toward a career in the military establishment which also allows him to progress to the highest level, except for those operations which are particularly associated with military operations leading toward combat, that is to say, the chief of staff and the deputy chief of staff for operations and training and for plans and programs. It is not immediately apparent why the deputy chief of staff for personnel or the comptroller or the deputy chief of staff for research and development could not be a civilian, providing he had clearly outstanding qualifications. Lest my argument be misunderstood, I should state that my concern for the career pattern of the civilian is not because of any

particular preference for the civilian over the military officer but, rather, because of my belief that by the present strict limitations on civilian career opportunities, the better civilian personnel are almost forced to leave the military establishment in fulfillment of their career aspirations. If an able civilian enters the military service, he soon finds that his direction and career progression is circumscribed and that he will have difficulty in achieving a high level of authority and responsibility. Thus, in his own self-interests he tends to leave the service. This is a serious loss to the military establishment and results in less than optimum performance from civilian personnel. Parenthetically, I should say that there are some dedicated civilians who, because of their patriotism and general sympathy with the military organization, prefer to stay with the military service at considerable sacrifice to themselves, but the number of such people is limited, and no organization should want to base the quality of its personnel on such a degree of devotion. In self-interest the military establishment should want to make it possible for the civilian to progress, as well as for the military to progress.

Some of you will recognize that the above discussion is consistent with positions I have taken for some time. In 1960 the Air Force Scientific Advisory Board formed a Committee on In-house Laboratories, of which I was a member, and, indeed, the establishment of the Air Force Systems Command Research and Technology Division was suggested in that report. We argued that in the limited area involving research laboratories it was important to give the civilian engineer or scientist an opportunity to rise to the head of not only his laboratory but to the head of the research and advanced technology organization. This suggestion has been partially followed under General Demler's direction, and there is now one civilian who is head of an R&D laboratory, although it should be noted that even he rose in responsibility as a military officer. That attracting personnel is still a serious problem was emphasized recently by General Demler in "Missile/Space Daily" (September 29, 1966, p. 136).

"SYSTEMS COMMAND WAGES CAMPAIGN
TO RECRUIT MORE SCIENTISTS AND ENGINEERS

"That the Air Force is bent on luring more scientific and engineering talent into its ranks, either as military or civilian personnel, is quite evident from remarks by Maj.Gen. Martin C. Demler, Commander of the Research and Technology Division of Air Force Systems Command.

"Speaking before a Joint Engineering Management Conference in Washington, D.C., this week, Demler talked about service-wide efforts–that collectively, incidentally, amount to a full-scale campaign–to make the Air Force as attractive and worthwhile an employer as industry, if not more so....

"Various strategies have a common objective–to secure for the Air Force its share of the scientific and engineering crop as it becomes available and then to ensure that this force as a whole stays at the front of technological advancement.

"Systems Command alone has more than 10,000 scientists and engineers assigned today, nearly half of them with the Research and Technology Division, but as Demler pointed out, the present shortage of technically trained young people is such that there are not enough of them to replace yesterday's leaders, let alone today's."

While young scientists tend to be devoted to the laboratory bench, many of them, as they grow older, become involved at an introductory level in administration and policy direction; as such, they need the opportunity to be able to rise to a high level of responsibility or they will seek this opportunity outside of the military service. The pointing to a few exceptions is not a sufficient answer to counter the fact that many able people have left the laboratories and then have risen to positions of significant importance in industrial organizations. Again, I would say that as a matter of self-interest the military establishment should see to it that the career opportunities for civilians are equal to those for the military.

Finally, let me emphasize that the questions posed here are raised with the thought that they should be investigated carefully. It is not enough that they be debated in a conference or investigated by a committee. Rather, system analysis and personnel research experts should devote their time to investigating these questions. Such investigation should be sponsored at the highest level in the Air Force so that the investigators can be assured of the ability of making thorough, comprehensive investigations and also be assured that their results will have the attention of policy-makers at the highest level.

Some Implications of Information Handling for Individualizing Personnel Actions

One of the serious problems in large organizations is the antithesis between centralized policy-setting and management on the one hand, and on the other hand, the decentralization of individual actions. Although the desire to decentralize personnel action is often advocated, it is extremely difficult to do so unless there are easy and rapid communication channels between the policy and management echelons and the action-taking echelons. Because these communication channels are so frequently blocked or are unavailable, there is a tendency for the level of approval to creep upward to higher echelons with a loss of flexibility and adjustment to individual needs. I believe that the advent of modern information processing has made it possible to achieve a level of both centralization and decentralization, which was not possible previously.

I would like to introduce this subject by telling a true story involving the late Walter Van Dyke Bingham. Dr. Bingham was one of the early psychologists who served in World War I and helped develop the old Army Alpha and Beta tests. In World War II, Dr. Bingham was called back to the Adjutant General's Office as the chief civilian advisor on selection and classification. Many of us young fellows tended to look on Dr. Bingham as a somewhat elderly gentleman who, while undoubtedly wise in the lore of testing, was not really quite up on the latest techniques. In those days, Eleanor Roosevelt wrote a newspaper column called "My Day." Early in World War II, she published a column in which she talked about the poor practices of the Army in assigning and classifying its personnel. She cited the case of a brilliant young lawyer with whom she was acquainted, who had been inducted and, according to her, completely misassigned into a machinists' training organization for which she said he was totally unfit. In a refined way, she took the Army to task for its bad performance in assigning personnel. This article greatly distressed Dr. Bingham, and he felt that Mrs. Roosevelt was uninformed regarding the work of the Adjutant General's Office, and particularly the Personnel Research Section. Because Dr. Bingham had so many acquaintances from his World War I days, he was able to call on Frances Perkins, who at that time was Secretary of Labor, and suggest to her that Mrs. Roosevelt might like to know more about Army personnel practices. Mrs. Perkins assured Dr. Bingham that Mrs. Roosevelt certainly would, and she acted as an intermediary to invite Mrs. Roosevelt to the Adjutant General's Office. I remember one day when Mrs. Roosevelt arrived, and Dr. Bingham escorted her into our temporary building and explained what we were doing and the efforts being made to develop tests and procedures to assure the proper assignment and utilization of personnel. Generals hovered about, and Dr. Bingham was at his gentlemanly best. A few days later Mrs. Roosevelt wrote another "My Day" column, in which she praised the Army and talked about the excellent efforts being led by Dr. Bingham and his psychologists. Now, either fortunately or unfortunately, both of her articles, in my opinion, had a considerable degree of truth to them. While we psychologists were trying very hard to devise sound classification methods, the Army also had quotas for many different kinds of assignments and training schools. These assignments had to be made in accordance with schedules, and since induction stations were scattered throughout the country and training was likewise scattered many places, the best each local assignment officer could do was to fulfill the assignment he had at a particular time. As a result many inductees were malassigned. Today this is no longer necessary because we have abundant, high-speed communication and the possibility of transmitting information through netted computer-based personnel systems.

The Air Force should take great pride in its leadership in the development of a computer-based military personnel system. Within the last few years, at Randolph AFB there has been established the Military Personnel Center, where the records are kept in digital form and handled through computer-based systems. I believe this is such an important development that its nature should not go unnoticed at this conference, and therefore I take the liberty of describing the Personnel Data System operated by the Military Personnel Center.*

Personnel operations are conducted by three levels of the Air Force. These are USAF Headquarters, the major commands, of which there are 20, and the consolidated base personnel offices, which are located at some 270 bases throughout the world.

The Headquarters personnel operation is divided geographically and functionally into two groups. The Military Personnel Center (MPC), Randolph AFB, is responsible for the design and maintenance of the Personnel Data System and for fulfilling the personnel functions of assignments, promotions, separations, and retirements for officers and E-8 and E-9 enlisted personnel.

At the Pentagon the personnel functions of planning, budgeting, and predicting for Air Force-wide personnel resources are carried out. The Directorate of Personnel Planning is the principal user of Air Force personnel management information and is also the focal point for analyzing such information for use by Air Force and DoD agencies. The Assistant for Personnel Systems and his staff comprise the responsible agency in the Headquarters for long-range personnel system requirements and planning, and significantly augment the capability of the Headquarters in conducting analytic studies.

An office known as the Management Information Office is established at most major commands and is the interface point between the Personnel Data System and the functional area users. The Management Information Office is responsible for maintenance, processing, distribution, and analysis of data contained in the Personnel Data System for Officers and the Interim Personnel Data System for Airmen.

The personnel organization at the base level is standard throughout the Air Force, with the manning level at the Consolidated Base Personnel Office varying with the number of base personnel being serviced. It is responsible for maintenance, processing, and distribution of data in the personnel system. It also provides the interface between the Personnel Data System for officer and airmen personnel operations.

The traditional separation of military personnel functions has been reflected in the creation of a two-part automated data system, the Personnel Data System for Officers (PDSO) and the Personnel Data System for Airmen (PDSA). These systems receive data from the Military Personnel Center, the major commands, and the Consolidated Base Personnel Offices, the majority of data being introduced at the base level. The users of personnel data are also located at these three echelons and receive products from the system for use in accomplishing their management and personnel operations tasks. Almost the only use of personnel data by Air Force agencies in the Pentagon is for management purposes; at the other levels more use of the data is made for operational than for management purposes; but both are necessary.

The present officer personnel data system evolved from a model implemented in 1963. This prototype provided the operational experience and preliminary data base necessary for implementing an improved model. The present system uses computer processing at the Military Personnel Center and the

---

* Further information concerning the Personnel Data System is contained in the presentation by Captain George H. Walther on "Time-Sharing."

major commands and punched-card equipment at the bases. At each echelon a different kind of computer capability is used for personnel data processing. The Military Personnel Center has a soon-to-be-augmented Burroughs 5500 computer complex which currently includes eight tape drives, 25 disc files, a card reader, a high-speed printer, and six modules of memory, each of which contains 4096 "words." Ten of the major commands have Honeywell 800-200 computer systems, while Air Force Logistics Command uses an IBM 7080. The remaining nine major commands are cross-serviced by other major commands with Honeywell computers. The bases have an even greater diversity of equipment which includes punched-card accounting machines, IBM 870s, 1401s, 1410s, and Burroughs 263s.

The Personnel Data System for Officers was designed and implemented to process personnel data on some 135,000 Air Force officers. Some of the characteristics of the system are:

A larger data base with standardized data elements.
Use of disc files at the MPC.
A remote inquiry capability at the MPC.
Standardized computer programs for processing data within and between echelons.
Standardized computer programs for inquiry purposes.
Standardized procedures for processing data.
Standardized hardware at the Major Commands.

Due to the limited capability of the bases, the data for the system have been specified in terms of 80-column cards. The Uniform Officer Record includes ten basic cards and, in addition, the system requires six transaction cards and eight types of miscellaneous cards.

The information in the data base falls into the following three groups.

Uniform Officer Record Information	Transaction Information	Miscellaneous Information
Strength Data	PDS Availability Data	Grade Spread Summary
Organizational Data	Accession/Levy/Allocation	Name-Service No. Change
Service Data		
Assignment Data	Reassignment Request	MAC Add-on
Availability Data	Projected Promotion	Personnel Accounting Symbol
Education Data	Projected Separation	MPC Local Data Series
Dependent Data	Projected Integration	MAC Local Data
Aeronautical Data		CBPO Local Data
Rated Qualification Data		CBPO Suspense Data
Rated History Data		
Previous Duty Data		
Current/Previous Duty Title		

main purpose of PDSO is to assist the functional managers at the Pentagon, the Military Personnel Center, the major commands, and the bases in carrying out their specific missions. Included in these functional areas are officer assignments, retirements, promotions, separations, and, to a limited degree, personnel analyses. In addition, such items as the monitoring of flying status, regular appointments, and selection for in-service schools are considered part of the personnel missions. In general, the system has been designed to bring about improvements in (1) availability of timely, accurate, and compatible data, (2) assignment actions, (3) personnel accounting, (4) notification of individual moves, (5) the selection process, and (6) stability in the personnel force, so that the most suitable men for the available jobs may be selected at minimal cost.

The personnel data system described above is a very important step forward in the expeditious and efficient handling of individuals within general management policies, but there are other developments which promise to make such systems even more useful in the assignment and career management aspects of personnel actions. Recently at System Development Corporation, we have been working on an automated counseling system. While this work has been oriented toward the public school situation, it is easy to see that it could be easily adapted to the military counseling and assignment environment.

In analyzing the advisory and counseling behaviors of people engaged in these functions in public schools, it has been shown that a major portion of their time is spent in processing routine information and in giving out repetitive information. Research financed with the assistance of the U. S. Office of Education has led to the development of a computer-based counseling system that will, among other things, conduct an automated counseling interview. The University of Oregon and the Palo Alto School District cooperated with SDC in the initial research. A program was developed to record the pre-interview behavior and interview interactions of an experienced high school counselor with a number of ninth grade students. Analysis of the recordings indicated that between 70 and 90 per cent of the counselor's behavior could be automatically handled. A model of the counselor's decision rules, both in appraising a student's records and his behavior in the interview, were defined for simulation on a computer.

Two programs were written: one to review the student's record and provide the counselor with an appraisal automatically, and another to conduct an automated interview with the student. The automated portion of the interview system operates as follows:

The student sits at a computer-linked teletype and receives a logically developed sequence of interview questions. For example, after a number of interactions, the computer may print the following question on the student's teletype: DO YOU PLAN TO CONTINUE YOUR FORMAL EDUCATION BEYOND HIGH SCHOOL? If the student types "NO," the next question he will receive on this typewriter might be: WHICH OF THESE IS MOST LIKE WHAT YOU PLAN ON DOING AFTER HIGH SCHOOL: (1) JOIN A MILITARY SERVICE; (2) GO THROUGH VOCATIONAL JOB TRAINING; (3) ENTER AN APPRENTICESHIP; (4) GET A JOB; (5) DECIDE LATER. After the student selects one of the alternatives the computer then asks an appropriate question to continue high school schedule planning.

The selection of each question depends both on the student's previous answers and on the information about him that has been previously stored in the machine. After a number of questions and answers, the interview will terminate in a schedule of suggested courses that is uniquely tailored to that student. Thus, students will be aided in selecting their subjects by answering a series of questions presented by the computer. This system should relieve the counselor of much routine work and at the same time be sufficiently sensitive to extraordinary responses so that the student needing expert help will be directed to the counselor for individual attention.

To help the counselor in preparation for either an automated or live interview, programs have been written which use the same rules as an experienced counselor in analyzing the information in the student's cumulative folder. This program was designed to simulate a specific counselor who was asked to "think aloud" as he analyzed twenty student cumulative folders prior to counseling interviews. Analysis showed that most of the pre-interview logic, as well as the interview itself, is specifiable and capable of being programmed. The pre-interview program accepts data similar to that which is normally found in the cumulative folder, such as grades, aptitude test scores, and biographical information. By making various comparisons of these data, the program prints messages for the counselor which are unique for each student. Messages regarding different students might be:

1. Student's grades have gone down quite a bit. Ask about this in interview. Possibly there are personal problems.

2. This student should be watched closely. He will probably need remedial courses.

3. Student is a potential dropout.

4. Should be headed for college. Encourage student to explore widely in academic areas.

5. Low counseling priority. No problems apparent.

6. Student should improve verbal skills. If not, student may not be able to attain desired academic goals.

In continuing the application of information processing technology in vocational counseling under field conditions, our study team is now in the process of conducting an extensive survey of vocational education installations, including state employment agencies and private and municipal vocational guidance projects. When the survey is completed, a sample field site will be selected for a detailed analysis of counseling procedures, and then a computer-based man-machine counseling system will be developed by a team that will include the counselors at the selected field site.

The computer-based system will have a data base containing student information, with an input-output system for updating and retrieving information. Computer programs in the system will provide appraisals of student data, interviewing procedures, tracking of student performance, and identification of students who are experiencing difficulty (2).

The possibilities of adapting such an automated counseling and assignment program to the military establishment has been enhanced by recent advances in computer time-sharing systems and computer networking technology. It is only within the last two years that there have been large practical time-sharing systems, and even now the experience with time-shared computer networks is not as extensive as would be desired. Nevertheless, the technical capability is clearly existent for an Air Force-wide application of the kind of program we are working on for the school systems. It should be possible for a military counselor to sit at a console at Lackland, or at Travis, or at Westover, or at any other appropriate base and be able to query the data base at Randolph, either at the Military Personnel Center or at the Headquarters Air Training Command regarding possible assignments for a counselee.* For some ten years now, the air defense network has been based on computer-to-computer communications, with data received at remote radars being transmitted through digital lines to a central computer, being processed, reported to a higher echelon at the division, and from there reported to NORAD headquarters. Since it is technically possible to achieve this sort of coordination for air defense, it should certainly be both possible and desirable to do the same for the Air Force's most important resource, its personnel.

The Problem of Utilizing Personnel Research and Developments

There is a growing concern regarding the organizational and operational practices which will result in better throughput from research to development to use. President Johnson has recently asked Secretary Gardner (Health, Education, and Welfare) to examine the activities of the National Institutes of Health to see if there are adequate efforts to bring their research programs to fruition in medical practice. Similar concern is expressed regarding physical science programs.

---

* Such queries were made during the Symposium from the meeting place and prior to the Symposium from Lackland AFB. Both the MPC data and the data located at SDC were queried.

For some time I have been studying this problem as it relates to the behavioral sciences, and recently the Joint Economic Committee of Congress reprinted one of my studies in this area (3). While there is concern regarding throughput problems in hardware engineering, there should be double concern in social and behavioral engineering. Developments in the physical sciences are resisted because of expense, reliability, and technical compatibility factors, and also because of user resistance springing from habit and the threat of disruption. In trying to engineer changes in the behavioral and social area, the difficulties of innovation and resistance are much greater–and rightly so. Significant changes in such areas as education, personnel practices, and work organization are met with resistance for two major reasons. First, such changes intimately affect the lives of individuals, and most of us are very cautious in accepting some new way which may disrupt our usual life patterns. The second reason is related to the first: namely, that for many social and behavioral changes it is somewhat uncertain as to exactly what effect the proposed change will have. It is difficult to build the first experimental prototype to see if it will work in the population at large.

Thus, those of us in the behavioral and social sciences have a greater problem and corresponding responsibility to see that our research and advanced developments are, indeed, exploited in actual practice. In examining a number of studies I have concluded that the following points need to be observed in helping achieve a proper transition from research to development to use.

1. We have begun to realize the special importance of the innovator and leader in transitioning from research to application. There are many successful research people who develop theories and demonstrate their validity, but then do not carry their application forward. In these instances the fruitfulness and utility of the idea become lost until some person picks it up later in connection with some new project.

2. The organization which is willing to try something new must be supported by a climate of encouragement and reasonable permissiveness which is set at the top. This climate must make it possible for staff to challenge existing practices and must reward those who are willing to try innovation as an instrument of improvement.

3. In large organizations, there are frequently procedural and organizational difficulties relative to the transition from research to development and to application. Often these functions are assigned to different major divisions of an organization on the theory that ideas developed in research will be picked up by a different group of people who will transform these ideas into an advanced development that will in turn be applied in some other part of the organization. It appears that considerable management and organizational flexibility is required, along with much crossing of organizational lines and management hierarchy, to carry forth successful developments.

4. Similarly, with respect to funding, large organizations, and particularly the Government, are constrained to develop budgets and administer funds under fairly rigorous financial procedures. However, this tends to inhibit the needed flexibility for development of new research. Studies indicate that the funds used for various research and development activities often did not come from the logically expected budgetary category. Rather, the leaders of new developments tend to find their funds wherever they can and to have little regard for formal funding authority. While this is disruptive of both management responsibility and neat accounting practices, it may well be one of the prices to be paid for effective research and development activities.

5. Another area critical to the application of new knowledge has to do with communication. From the evidence we have it would appear that the formal publication of new findings does not by any means assure that the results will be expeditiously translated into a useful development. Rather the indications are that informal communication is by all odds the most important method or technique for transmitting ideas from one environment into a different one.

6. Perhaps more important, however, is the requirement that innovations must be given credible demonstrations in the sense that they must be demonstrations in the ordinary operational setting, carried out by regular personnel and not by specialists who come into the operational situation and then leave.

It is generally acknowledged that the Bell Telephone Laboratories are among the most successful laboratories in industry with respect to both basic research and in helping the Bell system with its technical communications problems. Jack Morton (4) has ably described their operational and organizational philosophy in an article titled "From Research to Technology." This article is replete with insights regarding industrial research management. I believe his most important points were:

1. First, each scientist and engineer must have some clear goals which are related to the organization's goals. Morton says, "A major corporate goal came down to a narrow area of relevant science. I have told this story to show the importance of tight integration in the system: Everybody must know what the overall goal is, so that within each man's area he can look for those solutions which are most relevant to the goal....But what is required? Take the rubber products engineer, What is the first thing he must know? He must know what the president of his company thinks is important: What are the goals? And why? And what does the vice president think is important, and so on down the line. Until he knows, he is in no position to choose which thing, from a range of things, he should do. This is the way to build a machine–an organization–which will allow information to flow and individual judgments to be made. The man must know, whatever his level, what the important goals are. And he must know not only that the boss said it was so, but why." It is the responsibility of management to set goals–goals which are neither so general as to be platitudinous nor so specific as to stifle initiative.

2. Physical space and organizational space are important determiners of the effectiveness of research throughput. If the research organization is both geographically and organizationally isolated from the using organization, throughput is very difficult. But it is also unwise to be too close. If there is a geographic separation there should be an organizational tie, and if there is organizational separation there should be a geographic tie. Morton says, "When we studied our machine as a system, one thing we began to understand was that it was bad to have both an organizational barrier and a space barrier between development and manufacturing. Together, these barriers make communication too difficult, and one has problems in transferring new scientific developments into new technology. So we built a space bond between the laboratory and manufacturing: We actually moved our development-and-design-for-manufacture group–a Bell Labs group–into a laboratory on the Western premises. Organizationally, they belong to the laboratory. But physically, they are linked to Western. And now we know we should never have a space barrier and an organizational barrier on top of one another. We use organizational and spatial links in complementary relations–wherever we have a space barrier we also have an organizational bond, and vice versa. This gives us an integrated design.

3. There must be an organization concerned with relating goals, research, development, and use. It is not enough that goals be clearly stated or that there be physical and organizational closeness. There must be a separate entity concerned with seeing that throughput occurs. In the Bell Telephone Laboratories there is a special system engineering group which cuts across basic research, applied research, development and design, into the actual manufacture in the Western Electric Company. Morton says:

"[There is] a separate laboratory called Systems Engineering. It is not a line function. About 500 people are involved here, under an executive vice president, Ken McKay. What does McKay do? What is the function of Systems Engineering? By training and practice he is a physicist. By further experience, he is an excellent engineer as well. In Systems Engineering, he is concerned with economics and sociology, as well as with physics and engineering. One of his jobs is to build a bridge between the Laboratories and AT&T headquarters and the operating companies, and to be alert to needs and problems.

"Because he is part of Bell Labs, McKay is also in touch with scientific-technical developments. From the synapse of these he can draw a plan for a proposed system. His study will say: If we succeed in doing the following, this new system will provide x millions of new revenue per year, or it will save x millions per year. His people will also take inventory of all the technology needed for success of the new system. Large-scale developments are not undertaken without knowledge of missing critical capabilities. Systems Engineering makes many such studies every year.

"But McKay does not tell Baker [the vice president for research] he must do specific research, nor does he tell me to give him a specific gadget. He does tell us what is needed and what the priorities are, because the purpose of Systems Engineering is to provide overall guidance for research and development in terms of corporate needs. Systems people are in constant touch with all groups and the realities of their needs and possibilities—operating company, long lines, research. This is the way we form our synapses. But the choice of exploratory projects in research is the responsibility of research people and their management, for they can make the best decision, given the fact that they know the needs. Each man, whatever his level, takes the overall goal, translates it into the goals for his range of science and technology, then makes his choice of relevance in those terms."

It would appear that the Air Force might profit from giving serious consideration to the three points mentioned above, as they relate to both personnel and training. It would be wise to consider whether or not the goals of personnel and training are spelled out in sufficient detail by higher echelons so that research, development, and operating organizations will all have the same goals in mind and be able to point toward the achievement of these objectives. For some time I have been concerned with regard to both the geographical separation and the organizational separation of Air Force research activities in personnel and training as they related to the operating commands. Both the Personnel Research Laboratory and the laboratory responsible for training research are a part of the Air Force Systems Command, which is organizationally separated from both the Military Personnel Center and the Air Force Training Command. This separation may be wise if there is close geographic proximity between the organizations doing research and the operating organizations. I have the impression that this close relationship has not generally existed; however, I speculate as to whether or not it would be desirable for training research and personnel research to be on the same geographic base with the using organizations; if geographic separation must be maintained, then closer organizational identity needs to be established. I believe it would be appropriate for an operations research group to study research throughput in this area, being sure that in such an analysis members of both the research components and the operating components make up a part of the study team.

Finally, it does not appear that the Air Force has appropriate organizational entities to perform the systems engineering function described in Morton's article. It may be that in hardware development, there are adequate organizations in terms of the various AFSC divisions and their related systems project office organizations; but in the area of personnel and training, there does not seem to be an organization whose major purpose is to see that goals are stated and are recognized by research and development organizations as well as by the operating user, and to act as an active feedback and liaison mechanism in this area. This is not a function which can be adequately performed by a small headquarters organization. It is significant that in the Bell Laboratories as many as 500 professional people are engaged in the systems engineering activity. It seems that in the Air Force a new organizational invention is required, and I would suggest to Headquarters USAF that this is a high priority item in their effort to achieve maximum utilization in the research-to-development-to-use cycle.

Conclusions

As I think back to 1941 when I had my first association with military personnel activities and contrast the philosophies of military personnel operations of that time with those of today, I believe that

great advances have been made. The resources available in the military services have been very great and have allowed them to exploit much of the latest technology and latest thinking in personnel management. However, there are still many problems which need to be solved. These problems range from basic examinations of the philosophy of personnel management to the implementation of some of the more advanced information processing technologies. I strongly suggest that the investigation of the problems I have outlined above deserves support by the highest levels of the Air Force and that the careful analysis and study of these problems by a permanent staff of operations researchers and system analysts should result in the formulation of new programs which would result in an even more effective Air Force.

REFERENCES

1. **Drucker, P.** Planning for big changes just ahead. *Personnel administrator.* March-April 1966.
2. **Cogswell, J.F., Donahoe, C.P., Estavon, D.P., & Rosenquist, B.A.** *The design of a man-machine counseling system.* SP-2576/001/00. Santa Monica, California: System Development Corporation, September 30, 1966.
3. **Carter, L.F.** From Research to development to use. *Technology in education.* Subcommittee on Economic Progress of the Joint Economic Committee. Congress of the U.S. Washington: U.S. Government Printing Office, 1966, 226-236.
4. **Morton, J.A.** From research to technology. *International science and technology.* May 1964.

XVI

IMPROVING EFFICIENCY IN THE USE OF MANPOWER RESOURCES

Jack W. Carlson*

Introduction

I appreciate the invitation to present a paper on the occasion of the Silver Anniversary of the Personnel Research Laboratory. Many of the alumni and current members of the Laboratory are attending this conference and must justifiably feel a sense of accomplishment for what the Laboratory has now become.

The challenge of the Laboratory today is to remain dynamic. Members of the scientific community are forecasting a doubling of presently accumulated technical knowledge within the next 25 years. While the Laboratory and participants in this Symposium celebrate the accomplishments of the past, they should realize that the world of manpower analysis is rapidly changing.

My purpose is to identify areas for analysis and improvement in the management of manpower resources in both the Department of Defense (DoD) and the Air Force. First, the objectives of the Department of Defense are considered in relationship to our national goals. Brief consideration is then given to the appropriate mix of machine-power and manpower followed by a discussion of some approaches for identifying the most efficient mix of manpower skills. Policies which influence the actual flow of personnel are evaluated in terms of their impact on attracting, retaining, and motivating each skill-type required. In the final section, I have commented on the nature and magnitude of manpower research in the Air Force.

One conclusion is clear: *the Department of Defense and the Air Force are only moderately efficient and effective in the use of their personnel. It is equally clear that significant improvements can be made.* From incomplete analysis, it is reasonable to assert that a few important improvements could reduce manpower expenditures by more than $2 billion a year for the same level of performance, or DoD could benefit from a marked increase in personnel performance with no reduction in cost.[1]

---

* Mr. Carlson is currently with the President's Council of Economic Advisers; during 1965-66, he was Assistant for Special Studies in the Office of the Secretary of the Air Force. The analysis and conclusions in this paper are attributable to the author and should not be interpreted as reflecting the views of any Federal agency.

[1] This doesn't include the savings to society of eliminating the draft. A partial analysis of this aspect of manpower wastage can be found in an unpublished paper by W. Lee Hansen and Burton A. Weisbrod, "Economics of the Military Draft," University of Wisconsin, October 28, 1966.

National Security and other National Goals

As an economist, I am naturally concerned with allocating resources among competing claims on a basis of least costs. Although there are a few resources that are relatively "free," such as the air we breathe and the water we drink, most resources are scarce. (Even those that were once free are becoming scarce, e.g., clean air.)

These resources are distributed on the basis of the willingness and ability of each individual to purchase them (demand) and the corresponding availability of each (supply). The actual matching of these two forces, demand and supply, is called a price, or the trade-off of one good for other goods. While prices are often automatically determined in the private sector of our economy, they are not always automatically available in the public sector. Instead, the Government must deliberately compute prices ("shadow prices"). The ingredient that is lacking is not usually the supply cost—although at times computed incorrectly—but rather, the demand. In the private sector, demand is expressed automatically by "dollar votes" of consumers willing and able to purchase goods. In the public sector, Government demand for goods and services comes from national goals which are given expression through the media of the ballot box, demonstrations, and legal proceedings.

Several attempts have been made to forecast the future goals and objectives of Americans. The National Planning Association has recently enumerated and costed one set of goals.[2] Their analysis shows, first, that the cost of achieving even a modest array of goals far exceeds the ability of the economy to pay for them. Therefore, trade-offs between highly desirable goals must be made. Second, the achievement of even a lower level of these goals means that we must follow policies that sustain the economy at a high level of economic growth. Third, there are a variety of public goods demanded by Americans in addition to national defense. Some of these are urban development, social welfare, improvements in the quality of our environment, and better health and education.

The Department of Defense fits into the achievement of these American goals by (a) providing the primary focal point for national security and (b) by way of a byproduct, contributing in the achievement of other national goals. Obviously, if the Department of Defense can help raise the skill level of disadvantaged Americans while fulfilling its primary goal of national security at little or no additional cost, then it should. In this way DoD can achieve foreign policy objectives while simultaneously developing skills for economic growth and reducing poverty. This is extremely important now because total Federal expenditures for nondefense purposes have had to be cut to reduce pressures on prices and provide more funds for the war in Vietnam. In fiscal year 1967, expenditures subject to immediate Presidential control are estimated to be $23 billion or less,[3] while DoD expenditures are estimated to be $70 billion or more,[4] and expenditures not subject to immediate Presidential control are estimated to be $31.5 billion or more (e.g., Veterans' pensions, farm price supports).[5] Under these conditions other worthwhile objectives suffer unless they can be achieved within the Department of Defense budget of over $70 billion.

---

[2] Leonard A. Lecht, *Goals, Priorities, and Dollars, The Next Decade*, New York: The Free Press, 1966.

[3] Testimony given by Bureau of the Budget Director Charles Schultze before the Committee on Ways and Means, House of Representatives, September 12, 1966.

[4] The Budget Message of the President, January, 1967.

[5] *Op. cit.*, Director Schultze's testimony.

However, we are not sure of the total impact of Defense expenditures on the Nation's manpower resources and thus on both national security and other goals. Nor do we know what the real increases in cost and lower levels of effectiveness are to Defense by including programs to simultaneously achieve Defense objectives and other national goals. Therefore, as a first step towards reducing our ignorance, a study should be initiated to measure more adequately the impact of Defense expenditures on the economy through withdrawing, disciplining, training, educating, and discharging men. Then other studies should follow to determine more intelligently the extent that DoD should assist further in achieving other national goals.

Manpower Costs and Requirements

Once the national objectives of the Department of Defense are determined, the Department must then combine men, equipment, and facilities at the least cost. Alternatives should be explored to accomplish various levels of achievement. Unfortunately, the Department has traditionally concentrated its planning activities on the basis of weapons systems, with less regard for manpower resources. This has occurred in spite of the fact that roughly $30 billion or one-half of the Defense Budget is spent for military and civilian compensation and training costs. With over one-half of Defense expenditures immediately dependent on effective use of manpower resources, DoD should give this area at least as much analysis as that given to the development of weapons systems.

After arriving at a rough approximation of the proper mix of manpower and machine-power, further adjustments must consider the "qualities" or characteristics of personnel in relationship to each current and future job. It makes a difference if military and civilian personnel have achieved differing "quality levels" based on higher plateaus of education, skills, and experience. It matters both in terms of performance and costs.

Each Service and the Office of the Secretary of Defense now separate the development of manpower requirements (demand) and the actual management of the flow of personnel (supply). In both cases, costs are usually excluded or understated.

Numerous information systems covering portions of manpower resources in each of the Services and the Office of the Secretary of Defense have never been effectively integrated. A data system which exists for costs is completely separated from the data for manpower requirements. In addition, there is another describing the characteristics of personnel actually on active duty. Furthermore, to add to the confusion, there is only a primitive Air Force or Defense-wide data system for civilian personnel.

One of the few crude bridges in this gap is the development of personnel "cost factors" which usually reflect only current budget costs. They do not include in-kind benefits such as dependent medical care or future cash payments such as retirement or "insurance" payments for disability. There is no costing of personnel as they flow through career patterns, including expenditures for training, education, orientation, and relative levels of performance. Consequently each Service has practically no information as to the cost of each personnel type. One reason for this situation is the difficulty of developing this kind of data in the present structure of the Planning-Programming-Budgeting System. Apparently no attempt has been made to develop data for a Personnel System.

The first attempt in the Department of Defense to develop even crude total personnel costs factors was conducted during 1966 in the "civilianization" studies (Project MIXFIX). Even this seemingly obvious step was incomplete, and only the Air Force attempted to relate costs to each type of skilled personnel affected by civilianization of the military force. In the future, both average costs and those costs associated with adding a few individuals to an on-going activity (marginal costs) should be developed.

Costs can be computed for each type of skilled resource based upon occupation (or specialty), educational level, major academic field (for recently graduated personnel), experience, mental capabilities (e.g., AFQT, AFOQT), physical attributes (e.g., eyesight), risk acceptance levels, and other useful characteristics (or fewer). For example, the cost for an M.B.A. statistician, educated at the expense of the Department of Defense, with 3 years in the field and 10 years on active duty, who is adept at quantitative analysis, and who seeks to avoid risk-taking is different from the cost of a B.S. physical education major with 1 year as a maintenance officer and 5 years active duty, who is low on verbal and quantitative capability, who has 20:20 vision, and who seeks risk and is a pilot. There is a difference between an experienced engineer in the Civil Service and an inexperienced officer returning from the Engineering School at the Air Force Institute of Technology. In addition, there are differences in both the effectiveness and substitutability of different individuals with different characteristics.

In setting manpower requirements, these cost and performance differences should not be ignored. Manpower requirements are now set with only the most indirect consideration of costs. Boards of officers or distinguished individuals decide what the educational level should be for entire military occupations, or each supervisor is asked to specify personnel qualities needed for each job under his jurisdiction. Skill levels are therefore not directly related to cost nor are they constrained by increases in costs. Boards of officers and supervisors quite naturally prefer more highly trained personnel, irrespective of the job. Nor are cost considerations forced on higher level decision-makers. At the highest levels, increases in Defense appropriations are easier to obtain; in addition, the draft guarantees a flow of skilled personnel who are paid at less than civilian market wages. Although the draft does not assist the recruitment of Civil Service personnel, the fact of easily obtainable Defense dollars creates a bias toward specifying high-skill levels within the civilian component of DoD. The average distribution of grades among military officers in 1951 was O-3.0 (Captain) and was O-3.1 (or essentially the same) in 1964. Enlisted grades crept from E-3.6 (between Corporal and Staff Sergeant) in 1951 to E-3.9 (nearly Staff Sergeant) in 1964. During the same period the average Civil Service grades rose rapidly from GS-5.3 (1951) to GS-7.4 (1964). While the Military Services rushed to increase the educational and skill levels of most jobs with the grade structures remaining the same, Civil Service escalated grade levels to accomplish the same thing.

An obvious example of over-investment is found in the requirement that all pilots have baccalaureate degrees or higher. There is no indication that officers with baccalaureate degrees do better in purely flying assignments than those without baccalaureate degrees. Currently (July 1, 1966), 183 pilots on flying status hold doctor's degrees, and 4,586 have earned master's degrees. Many other officers with advanced degrees and enlisted personnel with high school diplomas and college training are doing menial military tasks which are well below their capabilities.

The overstatement of requirements for each personnel type has resulted in each Service maintaining an unnecessarily large educational and training program. Yet, the graduates of the programs which develop the highest skills are not likely to stay in the military beyond the period of time that they are obligated to serve because the overstatement of requirements results in their being assigned jobs which to them are unchallenging. (In addition, current personnel policies fail to encourage either retention or motivation as discussed below). The Services consequently find themselves in the wasteful position of educating and training two or more men when only one may have been adequate, if the requirement had not been overstated. There are analytical methods which can reduce the likelihood of either understating or overstating manpower requirements.

Manpower requirements can be determined by treating each organization as a producer of services or products for other organizations. What is necessary is an identifiable service or product and the manpower inputs. Each skill type (broadly defined) has a different and measurable impact on the output of each

organization. The same is true for various combinations of skill types. Assuming the trade-off between machine-power and manpower is made,[6] then the output measure of each organization can be identified. For example, a wing which supports and operates B-52s or F-105s might be measured by Operational Readiness Indices such as number of sorties flown/ number of aircraft assigned, number of successful missions/number of aircraft committed, number of aircraft in commission each day/number of aircraft flying hours, and average time required for the turn-around of aircraft.

In some cases, qualitative measures will have to be used because some outputs do not lend themselves to easy quantification. In the case of a research organization such as the Personnel Research Laboratory, merely counting the number of reports or pages in reports each year compared with the number of professional staff members employed to produce them is inadequate. Some measure of the quality of the work is necessary.

All military organizations (e.g., squadrons, wings) with similar missions or outputs and with similar equipment can be evaluated as to variations in the mix of skill types; one mix of skill types may tend to achieve higher levels of performance--as for example measured by the Operational Readiness Indices. The variations in skill types could be made to identify personnel types (based on education, experience, training, aptitudes, etc.) which are associated with higher levels of performance. The cost (compensation, training, variations in number of working days, etc.) of each group of personnel or skill type could be attached to its apparent productivity. The direction of change in the skill-mix, for each type of organization can be revealed by this process. The decision to change is easy if the costs of two skill types are the same, but one typically performs better in relationship to the organization's mission. If the cost of the higher performer is also larger, however, the choice is more difficult and the decision-maker must judge if the increase in performance is worth the increase in cost.

In contrast to the organizational approach, requirements can be determined by studying each existing job and how it is related to other jobs. The elements making up each job can be identified. Each group of factors thus compiled for each job can then be identified with personnel types. Then the requirements of an entire organization can be determined by summing the personnel types for each job. Care must be exercised to recognize the interdependence of jobs within a military unit. This approach looks at the problem from the small or element of each job, and then adds up the elements for each job requirement, hence applying this to all the jobs in an organization.

The organizational approach identifies the more productive organizations and the personnel types making for the greater effectiveness. Then more of the productive types are substituted for the less effective personnel types under the cost constraint until the best mix is achieved; this becomes the organizational requirement including requirements for each job.

The process of determining requirements should be repeated regularly to be consistent with the dynamic changes in DoD. Both of the above processes will greatly assist in reducing the misstatement of manpower requirements and possible over-investment in personnel skills.

In addition, not only the demand for personnel types changes, but the supply of each personnel type changes. For example, the quality of instruction provided by schools and colleges has increased dramatically during the past decade. This is reflected in the 3.2 per cent (trend) increase in productivity per man-hour, of which some studies show seven-eights can be explained by education, training, and other

[6] This is a simplifying assumption. The relative cost and effectiveness of personnel types will impact on the deepening and broadening of machine power, but this is not included in the discussion that follows.

causes of technological change. Over 50 per cent of the growth of Gross National Product has been attributed to this source.[7] Under these growth conditions, the Department of Defense is found wanting. In many areas, jobs or organizational requirements have not been reevaluated for 10 years and longer. This means jobs that once required a college degree might now require a high school degree, but the former requirement still exists. The net result of lagging behind the dynamics of the real world is even greater and continually increasing the military's overstatement of requirements for each job. This naturally leads to an unsatisfied man performing in his job at only one-half or two-thirds of his capability. He becomes discouraged and dissatisfied, waiting for the day he can return to an environment in civilian life which will allow him to use his full capabilities and which will reward him for this.

Personnel Policies and Research

Personnel policies must be modified to attract, retain, and motivate the flow of personnel types required for the Defense mission. It is not enough to look at the mix this year or next, but the mix must be examined for several years. Therefore, recruitment, disciplining, training, educating, and discharging should be evaluated as an integrated personnel system with all the techniques associated with systems analysis. The objectives of this should be to match the flow of personnel types to the requirements of each job. Alternative personnel policies and practices should be evaluated to determine those which attract, retain, and motivate each personnel type to meet adequately each job requirement at the lowest costs. If DoD objectives include assisting disadvantaged young men (Category IVs), then the costs and effectiveness may vary accordingly.

During the past few years, each Service has generally attempted to retain any and all personnel who would stay. This is true almost without exception for servicemen who had completed more than their initial obligation. Fortunately, this policy has been successful in providing the total number of people needed. Prior to the buildup in Vietnam, June 30, 1964, the Air Force, for example, had stated requirements that 55 per cent of their total force should be career personnel. The Air Force achieved 54 per cent career personnel, or 98 per cent of its goal. But, the individuals who accepted career status were not usually the ones the Air Force needed to fill the more technical jobs identified for career servicemen. For example, during the last five years, 31 per cent of all eligible first-term Motor Vehicle Operators, Management Engineering Specialists, Packaging Experts, and Information Helpers agreed to accept career status in the Air Force. Only 6 per cent of the Defense Missile Guidance Mechanics, Ballistic Missile Launch Repairmen, and Instrument Repair Technicians accepted career status. A similar pattern appeared among officer personnel. Sixty-three per cent of the nonrated personnel in the Operations Occupational Group remained beyond the sixth year while only 28 per cent of the Civil Engineering Occupational Group stayed.

The loss must be measured by both reduced performance from inadequately trained or insufficiently experienced personnel and increased training and processing costs due to higher turnover of people. The training of a Defense Missile Guidance Mechanic costs about $10,000 and requires 255 training days, while the Motor Vehicle Operator costs only $1,000 and requires 75 days.[8] The loss of the Missile Guidance

---

[7] See Edward Denison, *The Sources of Economic Growth in the United States*, Committee for Economic Growth, 1962; J. Kendrick, *Productivity Trends in the United States*, National Bureau of Economic Research, 1961; and Robert Solow, "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, 1957.

[8] Secretary McNamara's testimony before the Committee on Armed Forces, House of Representatives, June 7, 1965.

Mechanic is ten times greater in money and over three times more costly in time. If the retention rate for both skill areas is desired to be the same at, say 18 per cent, then the Department of Defense could pay up to $6,000 in cash to each mechanic while maintaining the current pay of the vehicle operator and disallow reenlistment of excessive drivers, or pay the vehicle operators less so as to retain only 18 per cent of the drivers, and pay the mechanics something greater than $6,000 additional.

Shortcomings of the personnel system in the past need not plague the Services in the future. There are several personnel policies that can and do assist in providing the required mix of skilled personnel.

First, the draft is of tremendous assistance to each Service in providing personnel in both the numbers, and skills required, at lower pay levels than the same young men would receive in the private sector. The Department of Defense has indicated that it would cost an additional $5 to $15 billion to attract and retain the same type and number of young men without the draft.[9]

Originally, the draft was designed to distribute the burden of protecting our country upon all citizens; and, as a byproduct, it provided a desirable "civilian flavor" in each Military Service. The draft, however, is a great subsidy for DoD, and its cost unfortunately is not equally distributed among the Nation's young men. In fact, the cost tends to be borne by young men who are unable to avoid the draft through college or other types of deferments.

Second, the structure of compensation and incentives is very important in securing the mix of personnel types required. The experience of the private sector, where market forces operate more fully, does provide insights which may be helpful in improving the structure of compensation. The private sector pays for:

- Higher levels of skill
- Relative scarcity of each occupation
- Superior performance
- Higher levels of education
- Greater experience
- Greater responsibility
- Increases in hazard
- Disruptions to family life

The private sector pays over 88 per cent of its compensations in cash.

In contrast to the private sector, the military has recognized only half of the criteria above: specifically, increase in hazard, greater experience, disruption to family life, and recently, some increased payment for relative scarcity of occupations among enlisted personnel (e.g., Proficiency Pay and the Variable Reenlistment Bonus). Also increased pay is given to a relatively few officers for special qualifications: doctors of medicine and lawyers. Except for the above cases, increased pay is not usually given for higher levels of skill, higher levels of education, superior performance or relative scarcity of each

[9] Because the draft provides manpower to the Services at less than the prevailing wage rate for similarly skilled men in the civilian economy, the Services tend to favor military techniques and methods which use relatively large amounts of labor in general and skilled labor in particular. Also, total production for the entire economy is lower; manpower is used where it is not most productive. Therefore, the social cost of $5 to $15 billion, which is now borne by drafted young men would be moving towards a voluntary force.

occupation. Almost without exception, DoD has requested higher across-the-board compensation increases instead of differential pay so as to recognize more fully higher skill and educational levels. The salary reform studies of the Hook Commission in 1949, the Cordiner Committee in 1958, and the Randall panel in 1963 did not fully consider differential pay. And now, weapons systems and military planning are more sophisticated, but DoD is saddled with an archaic compensation structure. The net result is that DoD pays highly skilled personnel too little and less skilled personnel too much. With the economy at a high level of employment and with increases in growth dependent on the efficient use of manpower resources, the Federal Government and the Department of Defense cannot be satisfied with continually over-paying some of its personnel and under-paying others. Pay must be used, as it was intended, to attract, retain, and motivate the wide range of personnel skills that are required. No more or no less should be paid. There are other critical claims on the taxpayers' dollars that cannot go unrecognized. The current under-payment of some personnel leads to excessive turnover and the associated higher costs for additional training. The over-payment of others leads to higher than necessary compensation costs and training costs, which compounds the problem.

Military compensation distributes more non-cash benefits by providing more "free" present and future services than is the practice in the private sector. To add to the problem the proportion of non-cash and delayed elements of total compensation has increased since 1949.

Roughly 50 per cent of total compensation is now labeled basic pay and is paid in cash. Thirty per cent is paid for special purposes such as housing, clothing, and hazard allowances which may be paid in cash (e.g., hazard pay) or in kind (e.g., housing). Over 20 per cent is paid for strictly non-cash or future benefits for such services as dependent medical care and for future pays such as retirement and "insurance" for disability.

The more elderly use non-cash services more than younger personnel. Therefore, greater compensation accrues in this form to older and more senior personnel. Also, it is evident that free services are used more extensively than is necessary. If the Federal Government gave in cash to military personnel an amount equal to the total cost of providing free medical facilities to dependents and Service personnel for nonservice-connected medical care, less medical service would be demanded by Service personnel because of their higher preference for other goods and services. A move in this direction can help to alleviate over-crowding in military medical facilities.

Almost all servicemen prefer a higher amount of cash now rather than future payments. This is evident by the interest rate military personnel are willing to pay for consumer goods. Generally, younger people are willing to pay high rates of interest to borrow for the essentials of courting and setting up housekeeping. The time preference or interest rate for cash is typically 10 to 20 per cent for younger military personnel. This rate is significantly higher than either Government bond yields of 4½ per cent or lower Government accounting rates of 3 or 3½ per cent. Because of this difference, income promised 20 years in the future means little or nothing now to young servicemen, but it means a higher cost by Government computations. Characteristically, Government pay increases during the past five years have ignored this fact and have increased retirement pay at a faster rate than cash payments.

To an airman or officer in his nineteenth year of service, a basic pay increase at the guidepost rate of 3.2 per cent means 25 to 35 per cent increase in compensation that year. This arises because the retirement pay is a fixed percentage of terminal basic pay and provides a stream of retirement income at 50 per cent of

the basic pay increase for the remaining life of the serviceman.[10] In contrast, the airman or officer completing his first year of service receives practically nothing in present value terms from the increase in his retirement. (A similar but smaller impact occurs with Civil Service pay; retirement pay is based on a percentage of the average income of the highest five years.) Thus, pay increases given to military and Civil Service personnel during the last five years have given very small compensation increases to junior and middle career personnel and have given very large increases to personnel who have long since committed themselves to a career and who generally possess less education and are less open to innovation. This is another example of the Government's increasing the compensation of senior personnel much more rapidly than junior and middle career personnel who are generally in short supply. The additional cost of rapidly increasing compensation for senior personnel is in many cases unnecessarily burdening the taxpayer without any significant returns.

The above deficiencies in the military compensation system (and to a lesser extent, Civil Service) are further encouraging younger and more skilled personnel to return to the private sector as soon as possible, where compensation is not predominantly based on old-age or paid in services. The compensation system should be modified to provide more compensation in cash, greater differentials based on skill and educational levels, and less future pay if the Services wish to increase retention and motivation at much lower costs.

Third, promotion policies are successfully used in the private sector to retain and motivate skilled personnel. In contrast, in the Military Services and to some extent in the Civil Service, promotions are seldom given for above-average performance and for purposes of retention. Promotions are primarily based on waiting for the requisite age and then being promoted along with poorer performers. For example, in the December 1965 promotion to Major in the Air Force, only 1.3 per cent of all promotions were based in any way on above-average performance. The remainder, or 98.7 per cent of the officers promoted, were treated as if they were of equal performance and were promoted after roughly the same number of years of service.* Quite obviously, many officers among the 98.7 per cent had superior performance records and tremendous potential for further development, while in the same group many officers had below-average records and low potential for development. There is nothing so discouraging to young and highly trained men as to be told that they must wait and be promoted with "the herd." Yet this is precisely what current policies implicitly say to highly qualified young men. Naturally, they won't stay and they don't stay. For example, the retention rate of enlisted personnel after completing four years in 1964 was 13 per cent for those with the rank of E-3 (Airman First Class) and 33 per cent for those with the rank of E-4 (Staff Sergeant).[11] A similar pattern emerges when comparing officers with more advanced degrees in skill areas in short supply and officers with baccalaureate degrees in academic disciplines which were in ample supply.

---

[10] For example, an officer of 43 years of age, who is a Lt Colonel earning $10,000 base pay, who has a remaining life expectancy of 25 years, and who receives a base pay increase of 3.2 per cent, or $320, will receive a compensation increase of 25 per cent, or $2,527.75 = $320 + $\sum_{n=1}^{19} \frac{\$160}{(14.05)^n}$

(The above calculation considered only basic pay and retirement and thus did not include other benefits.)

[11] The causal relation may have been reversed: Promotions may have been given to enlisted personnel because they expressed a desire for a career and not because of higher achievement levels.

* **Ed. Note**—Although the statement here is generally correct, it is somewhat misleading since those passed over for promotion to major tended toward the lower end of performance distribution as measured by effectiveness reports.

When men who are highly trained and educated leave because of essentially no opportunity and no promotion, the Services must maintain a larger training and education program for less capable personnel. It is hard to estimate what tremendous losses are sustained by such personnel policies, but the combination of existing compensation and promotional policies must cause at least a 6 per cent wastage in DoD manpower resources. This would be equivalent to roughly $2 billion dollars. Perhaps it is closer to $5 billion.

Fourth, awards and special recognition motivate people to higher performance levels. There is no substitute for the pat on the back, public recognition, and medals for retaining and motivating personnel. Even the seemingly insignificant addition of lightning bolts for the visors of Lieutenant Colonels gave outward recognition of the officer's high status and probably increased his motivation. Commendably, each Service uses this policy tool very effectively.

Fifth, assignment policies can and do have an impact on attracting, retaining, and motivating skilled personnel. In the civilian economy, individuals specialize early in their careers and then broaden themselves for additional levels of responsibility. Among enlisted and Civil Service personnel, the Department of Defense does essentially the same thing. But this is not followed with officers. Officers are rotated at 2- to 4-year intervals into and out of specialties and occupation groups. Although each Service attempts to develop some career patterns, their efforts are either ineffective or are ignored in actual assignments. For example, an officer may fly for one tour, then be an academic instructor, then a management analyst, and then return to the cockpit. Each job requires on-the-job training and may mean the equivalent of a 1-year loss for each 3 years assigned to each new occupation. This approach is frequently rationalized because of the need to prepare all officers for general officer responsibilities in spite of the fact that only .03 per cent can be general officers at any one time. Or, a loose career ladder is maintained to allow older personnel no longer suitable for flying or combat jobs to move into a "desk job." Of course, the preparations and execution of a war require flexibility, but the current assignment policies appear unduly flexible and intolerably wasteful. The education and experience patterns of general officers could include earlier specialization with broadening responsibilities later in a career, as occurs with many top executives in industry. Flying personnel could be given a career pattern which includes only flying with the flexibility that pilots could leave with reduced retirement benefits or a lump sum payment after, say, 10 or 15 years. If the airlines are an example, pilots in nonfighter aircraft can have successful careers through at least 20 or more years. Fighter aircraft may require the faster reflexes of youth and, therefore, fighter pilots should be considered for transfer to bombers or cargo-carrying aircraft later in their careers.

The great advances in technology now affecting the Department of Defense make it folly to ignore specialization among all personnel. If ignored, then the substitution of civilian personnel in the place of military personnel should and must be considered as a desirable alternative in all but direct combat-type jobs.

While it is easy to identify problems, and obviously wasteful policies, it is much harder to correct them. I do not wish to minimize the difficulties. But it is not always clear that we are moving in the right direction.

Conceptually, it is possible to define a personnel type (e.g., a man who is an experienced comptroller with more than one year's experience, is a non-rated company-grade officer, and has a master's degree in business) and trace his movement in the personnel system, given current policies.[12] This can be followed by

[12] See the developments at the Personnel Research Laboratory. The Laboratory has successfully applied a useful flow simulation technique (modified Markov chain) to the Air Force officer force in some recent research efforts.

simulating the modification of personnel policies (e.g., compensation system based on additional pay for higher skill and educational levels and a promotion system based on performance) in order to calculate the likely changes in the actions of each personnel type. Then the cost stream for each personnel type under alternative personnel policies should be determined. Thereby, least-cost policy options can be isolated for the decision-maker.

To carry this approach further, changes in requirements for each personnel type can be simulated in relationship to the needs for manning weapons systems and other military jobs. Again, cost streams must be attached to alternative mixes of personnel types. Similarly, alternative requirements can be evaluated. In the final step, simulation of various personnel types can be made on the requirement side as well, with the ultimate matching of requirements (demand) and personnel flows (supply). The decision-maker can intuitively add factors not already considered and then make a selection of the mix of personnel policies and a statement of requirements at or near the lowest costs. This type of model, or more correctly, family of models offers a tremendous potential for greatly improving the effectiveness of personnel and manpower management, and at less cost.

The creation of these models and the development of other research projects discussed above will not be an easy task. Moreover, it will require improvement in the operation of manpower research and development activities within each Service.

To be most effective, applied research is best performed under special environmental conditions. These are: semi-isolation from operating organizations; latitude to initiate projects as well as to perform assigned research projects; and access to the highest levels in each Service (Deputy Chief and Chief of Staffs, Assistant Secretaries and Secretaries of each Service and DoD). Basic research is ideally conducted in an academic environment with its associated freedom and inquiring atmosphere. DoD should include grants to the academic community to foster this type of research.

Applied research can be performed in essentially three environments: academic community, consulting organizations, and in-house applied research centers. Each has its advantages and each should be used by DoD. There should be more contact with competent professors in civilian universities, and more contracts should be let to these agencies. Consulting groups, such as the Rand Corporation for the Air Force, provide a valuable and indispensable bridge between the academic community and each Service, as well as generating imaginative approaches to resolving old problems.

In-house research efforts are particularly important in developing new systems for personnel management and recommending policy changes. The conceptual development of the personnel management model and the development of occupational analysis at the Personnel Research Laboratory are two very good examples.

Although research efforts are not built overnight, now is the time to start. Competent researchers are hard to find, and, unfortunately, the Services have deluded themselves into thinking they can have a research effort in-house merely by reorganizing and labeling new jobs as research positions. It is time that the Services reduce the detrimental fragmentation of their efforts and further isolate research centers from the day-to-day demands of military operations.

Improvement in applied research and development will require the integration of diverse data sources. A fruitful starting point can be the integration of all costs, personnel, manpower, training, education, pre-accession, ROTC, retirement, and Reserve data files which already exist. Computer facilities set aside for manpower analysis may need to be expanded in each of the Services. As in the past, the Services will want to safeguard the privacy of each serviceman's personnel records.

In the Air Force, the Personnel Research Laboratory is the focal point of personnel and manpower applied research. It will need far more support than it is now receiving to match the challenge of the 1960's and 1970's. Also the Laboratory needs to increase its ties with all relevant disciplines, including economics, in the academic community (including the Air Force Academy and Air University). One approach is to reserve manpower spaces at the Laboratory for visiting researchers from academic life for part or all of a year. Also, consultants could be budgeted to work on special problems. In addition, it may be useful to have biannual symposia on Air Force and DoD Manpower research.

XVII

THE IMPACT OF COMPUTER APPLICATIONS ON PERSONNEL MANAGEMENT SYSTEMS

John M. Leiman
California State College at Los Angeles

When one looks at the voluminous literature, on the one hand in the area of computer and information management developments in the last ten years and, on the other hand in the area of personnel management systems development and use in the last 25 years it seems obvious that the impact of the first upon the second has been and will continue to be tremendous indeed. Most of you will agree with what I have said so far. You may not agree with the rest of what I have to say. I believe that computers can and will replace and do better and cheaper the current functions of most personnel staffs. The grim specter often raised of the management of men by machines, untouched by human hands, is not so grim because it would likely be an improvement over most personnel management systems now existent. If the only hope, or the only danger, is automation of current functions, then my advice to personnel managers is to find an interesting hobby, for machines will put them out of business. But I think that this replacement of man by machines in this respect has been thus far, and will continue to be, only a superficial impact, for it does not in itself change any of the basic concepts or solve any of the fundamental problems of personnel management.

It is also obvious to me that the use of computers and associated technology as tools in traditional personnel selection research and development, while of considerable use, will not have any basic impact on personnel management. I do not mean to disparage these two impacts since I think they are salutory, but on looking at some of the more fundamental objectives of personnel management I see, if anything, a worsening condition. I think at least one of the most important objectives of personnel management systems is that they contribute to the realization of the Great Society, and I was gratified to hear Dr. Dailey give us the Health, Education, and Welfare Secretary's definition. I am also gratified to know that the Secretary's definition agrees with the aim of governments set forth in a document of some years standing, the one I have taken as my text for this paper. I believe that the really important impact of computers on Manpower Utilization Systems can be the realization of the truths held to be self-evident in the Declaration of Independence. I would like to quote a few familiar, if often ignored, lines from that document.

> "We hold these truths to be self-evident, that all men are created equal, that they are endowed by their Creator with certain inalienable Rights, that among these are Life, Liberty and the pursuit of Happiness. That to secure these rights, Governments are instituted among Men, deriving their just powers from the consent of the governed."

Governments are social systems, and their primary purpose is to regulate the behavior of men. Personnel management systems also exist for the main purpose of regulating the behavior of men. The question is: Do current personnel management systems have as their main purpose the regulation of the behavior of men to the end that the inalienable rights of *all* men are secured? I believe that the answer to

this question, in this country today, for the overwhelming majority of systems, is *no*. I believe that the real impact of computers on personnel management systems is that they will make it possible to bring about the development and use of systems which will more surely guarantee to all men the right to a useful place in society--the prerequisite to the pursuit of happiness.

In anticipation of those who might want to question the assumption that the aim just quoted is indeed the main and rightful purpose of personnel management systems on whatever grounds, let me simply challenge them to argue publically, on moral or any other grounds, that there is a better fundamental aim of governments or of personnel management systems.

Let me also say that we know that men are not really created equal as is stated in the quotation, but I think it is a fair interpretation that the equality must refer to equality before the law or before the social systems that have been instituted among men; in other words, equality of opportunity to pursue happiness and to live at liberty.

Most of the evidence bearing on the question of equal opportunity is often interpreted as differences in the degree to which men exert themselves in exploitation of their opportunities. I will not argue that all men should be equally successful in the pursuit of their rights, for men do differ in their native abilities, but I will argue that the differences, that I will discuss below, are just too great to be the result of differences in energy of pursuit or native ability. They are rather the results of personnel management systems, and other social systems, that do not place the equality of the rights of all paramount in their aims.

Let me cite some statistics that illustrate the differences. The lowest one-tenth of the spending units (families and single persons living in separate households) in the United States receive 1 per cent of the total personal annual income, while the upper one-tenth receive 28 per cent of the total income. This means that some 6,000,000 families get 28 times as much, on the average, as 6,000,000 others; if one considers more extreme samples, the 3,000,000 in the top 5 per cent receive more than 40 times as much, on the average per year, as those in the lowest 10 per cent. And this is only in terms of annual dollar income. The discrepancy is even greater when one considers the degree to which men of these extreme classes have succeeded in the pursuit of total material wealth (1).

The recent minimum wage law raises the number of those who are covered to about 30,000,000 workers. If one remembers that the labor force in the United States is now about 75,000,000, one can ask why are less than 50 per cent of the labor force covered by minimum wage laws? Why not everyone? Does this look like equality of opportunity?

Of the some 75,000,000 members in the labor force, fewer than 30,000,000 are members of labor unions, and you may be sure that few, if any, of these work at or below the minimum legal wage. This is not because 45,000,000 employed persons are professional, managerial, or proprietory workers, for only some 25 per cent of the labor force is so classified. It is difficult to know how many workers do in fact work for less than the minimum wage, but a little arithmetic with the numbers given in the Statistical Abstract of the United States indicates that the number is in the neighborhood of twenty to twenty-five million. One may ask: Why is this? Are some jobs just not worth any more, or is it because our social systems are not really dedicated to the principle that men are entitled to any equal opportunity?

I would like to cite just one more piece of evidence before I go on to a discussion of what might be done about the rather grim picture I have drawn. One need only spend some small amount of time in the vast slums of our great cities, or in the itinerant worker camps of some of the most productive agricultural areas of the nation, or in the deep south or the southwest to realize the abject lack of hope for a decent life that is the lot of about 50,000,000 of our citizens--one- fourth of our people (2).

We live in a very affluent society. Our total personal income approaches $600,000,000,000 annually, and this represents an average income of about $3,000 per year for each of the some 200,000,000 people. Yet the 50,000,000 who are estimated to live in poverty receive an average of only about $700 per year. By contrast the 50,000,000 people who are most affluent receive an average of about $6,000 per year. Is it really necessary that there be this wide difference? Does the difference represent just differences in native ability or the energy with which people pursue their opportunities? Can it not be the result of our personnel management systems and other social systems that do in fact operate to prevent equal opportunities for all?

I would like to turn now to a discussion of what, if anything, may be done about the existent inequities in the degree to which men have been successful in the pursuit of the "goodies" of life. I do not want to argue for a strictly equalitarian society for, because men are different in their capabilities, some must be more successful than others in the pursuit of those things which bring satisfaction and fulfillment to our lives. But the differences between the high and the low are just too great. At first glance the simplest solution might seem to be that we take from the rich and successful and give to the poor and the unsuccessful. That is just the solution that we have been trying for the last several years in our federal, state, and local welfare programs. Yet these programs, while they probably prevent starvation and stave off revolution, are generally conceded to be failures in the sense that they provide nothing more than a stop-gap solution. Indeed we know that there is growing among many of our poor a culture of poverty. We are in the third generation of this in many of our city slums. Welfare and charity are doubly defeating, for they do not provide the means for the recipients to improve their state, and they create resentment into the bargain.

It has been suggested that we institute a guaranteed minimum income for every one, or that we provide for a "negative income tax." I think such an arrangement would be far better than our current welfare programs and might be even less expensive, but I do not think it would solve the problem, for the one necessary thing to lift people above the state of poverty is that they be made to feel that they do belong to the society, and that they do contribute to it. This means, in our culture, that they have a job--a real job, not one that is welfare or charity in disguise. I believe that nothing short of a society in which suitable jobs are available for all will serve to secure the rights of all men. This situation should be of concern to us all in our roles as citizens and human beings. It should be of particular interest to those of us who are concerned with the development and use of personnel management systems. I think the kind of systems we now use are failures, and I think they are failures because we have given blind allegiance to a doctrine of selection rather than a doctrine of classification. By a doctrine of selection I simply mean that we determine how many jobs a particular system needs and then line up the available men in an order we believe relevant to their ability to perform the job. We then simply take from the top of the list and throw away those we do not need. Of course, the process is repeated many times, but the same people tend to be at the bottom of every list. They are the breadwinners of the 50,000,000 poverty-stricken people of the nation. For the most part they are not very capable people. We all know they are in general, poorly educated, poorly adjusted, unskilled, old, and unhealthy, or the offspring of such people, for these are the characteristics of the poor. But they do have some capabilities, and we must use their capabilities even though it may reduce the efficiency of our production systems.

Personnel management systems dedicated to the selection doctrine might not be so bad if the basis of selection were a consistent one. All too often, however, men are placed at the bottom of the lists because of such immaterial things as color of skin, religion, sex, or just the whim of some administrator. Like success, nothing fails like failure, and one slip often means the end. And there is no recourse for most men, since not to employ is the prerogative solely of the employer. In short, the selection doctrine means that men are to be engineered to fit jobs.

There is an alternative doctrine–what I referred to earlier as a doctrine of classification. Under such a doctrine the supply of men is classified according to their capabilities and potential capabilities. Jobs are engineered to fit the supply of men. The system is made to serve men rather than men made to serve the system.

There are some difficulties in making a change from personnel management systems dedicated to selection to systems dedicated to classification. All changes in doctrines or customs of long standing have two fundamental kinds of difficulties, and personnel management systems are no exception to this principle. One of the problems is a technical one or simply the lack of knowledge of how to devise better solutions. The second kind of problem is how to find the resolve to implement the improved solution once it is known. The distinction often grows a little fuzzy, for the problem of how to find financial and other kinds of support for the search for a better technology is more like the problem of implementation than it is like the problem of the search for new knowledge.

In order to change to a doctrine of classification, we need capability in four areas. (1) The ability to classify men according to their potential for future job performance; (2) The ability to engineer new jobs and re-engineer old jobs so that they fit the capabilities of men; (3) The ability to match men to jobs so that no men are left over and no jobs are unfilled; (4) The ability to control the creation of new jobs and the discarding of old ones so that the supply of men and jobs does not get out of balance. This last may also imply the ability and responsibility to train and retrain men before new jobs are created or old jobs discarded. The first three are concerned with technology; the last is a problem of implementation.

The first area is comprised of an old familiar set of problems. They are the problems of predicting future job performance. For formal criteria such as success in educational and training programs, we have achieved validity coefficients in the sixties pretty consistently since World War II, and we have not, in the last 25 years, improved this very much despite a very considerable effort. Factor analysis and other sophisticated statistical techniques have been employed. Some very powerful computational procedures have been used. We have expanded the scope of psychometric devices; that is, tests are available for a much larger set of jobs than was the case a few years ago. Valiant efforts to develop non-cognitive type tests have been made. We have automated the test scoring process and have even begun to automate some of our administrative procedures. And yet, validity coefficients have not changed significantly. I would not argue that we should stop working for improvements in this area, but rather that the change to a classification doctrine does not depend on such improvement. We should use more extensively the capability that we have, for it should be remembered that the large majority of the working force is not selected on the basis of psychometric tests. The technology in this area may not be as good as we would like it to be, but we do now classify men in regard to future job performance. The current technology is good enough for a classification doctrine if we just use it.

The second problem area, job engineering, is also an old familiar one. Charles Babbage who invented the forerunner of the electronic computer in the early 1800's wrote a paper on job analysis as far back as 1832 (3). In fact, Babbage's discussion of an efficient way to manufacture pins sounds much like some of the work in job analysis done today. It has been only recently, however, that human engineering has had much of an impact on the design of equipment with the capabilities and welfare of operators in mind.

There remain a great many unsolved problems in job engineering despite very great progress in the last ten or fifteen years. But like the prediction of job performance, we engineer jobs everyday; the fact that we do not do it as well as we know how, or as well as we would like to, need be no deterrent to the adoption of a classification doctrine in the use of men.

The third problem area, that of matching men to jobs, is also an old familiar one. However, very little work of a scientific nature was attempted until just a few years ago. This is probably because the computation problems involved are of such a nature that one could not hope to solve such problems

formally for more than just a few men and a few jobs without the aid of a computer. In fact, direct solutions for more than about one thousand men and one thousand jobs tax modern-day computers. Fortunately, good approximation procedures are available that are well within the capability of small- and medium-sized present-day machines. I do not know how feasible it would be to develop a reasonably good solution for the working force and jobs in the United States today, but a modest effort should produce such a solution.

The fourth problem area, the ability to control the creation and discarding of jobs and to train and retrain men, so that a man-job match on a local as well as a national scale is feasible, is a far more difficult task. There are two main difficulties, and both involve customs and doctrines and even legal codes which have become outmoded by circumstances. These difficulties do not require any technical breakthroughs for their resolution unless one considers that technical developments are required to show us how to make changes in long established customs.

The first difficulty is in the tacit assumption, embedded in the selection doctrine of manpower utilization, that men can and will perform, given the proper motivation, any job that is created, and that they can and will learn to perform a new job if their old job is discarded by technical progress or for any other reason. It is the doctrine that men will change themselves to fit the society. The existence of the 50,000,000 people who live in poverty and the enormous welfare programs in this country, however, are testimonies to the fact that many men have not acquired the skills required by the jobs that the society has created. One may argue that it is simply a matter of motivation, that the men and women who do not have jobs or a place of dignity in the society are in such a state simply because they do not try hard enough. I do not believe this is true. I believe that the great majority of unemployed and underemployed people would much prefer to work and earn a place of dignity in the society if a reasonable opportunity to do so were provided them. We have in this nation our share of geniuses, but we also have our share of the other end of the scale, and I see no practical way to change this. I see no way to secure the right of all men to an equal opportunity for a life of dignity if we continue to engineer jobs without regard to the entire range of abilities that exists and will continue to exist in the pool of men available to fill the jobs.

The second difficulty is just the problem of how to control the creation and elimination of jobs so that the total fits reasonably well the available pool of the abilities of men. We know enough to describe the totality of work with respect to both quality and quantity that can be reasonably expected from the manpower pool. We know enough to be able to project this for several years into the future. We have the technology to design systems such that the work of men associated with the systems will result in efficient operation without undue stress upon the worker. We also have the technology to match men to jobs so that few jobs need be unfilled. I think we have the technology to do all these things without an undue strain on the efficiency of our production systems.

The question is, Why do we not do it? Why do we have 50,000,000 people who do not share in the affluency of the nation? I think the answer is that we do not make use of the technology that we have, and in particular, we do not have the means to control either men or jobs when there is a surplus of either–be the surplus one of quality or quantity. In the past, the problem of imbalance between men and jobs has seemed to have been solved by the law of supply and demand. Perhaps a job surplus has been solved in this manner, for unfilled jobs are soon modified or else the incentives to fill them are modified. The result is that if the jobs are critical they do get filled. The other surplus, that of more men than jobs, is not so easy to remedy. The result, if the latter condition persists, is that the unemployed people are reduced to a state of poverty. It is often argued that unemployment is structural in nature; that is, that there is not a shortage of jobs in the aggregate, just that the men who do not have jobs do not have the skills to perform the jobs that are available. On this assumption, we have in the last few years, at all levels of government, instituted a wide variety of programs that will upgrade the skills of the unemployed so that they can perform the unfilled jobs. These programs will not solve the problem of the 50,000,000 people who live in poverty. For the most part, they are not very trainable; even if we do succeed in training them, the new jobs they can

aspire to are already filled by marginal people. The result may be some mobility within the ranks of the poverty-stricken without much change in the aggregate number.

A basic philosophy of the selection doctrine in manpower utilization is that men who are at the bottom of everyone's list will somehow raise themselves in relation to the scale of ability to perform. The lists are crowded at the bottom. They are also crowded near the bottom.

We have only become interested in those on the bottom of the lists in the last twenty or thirty years. This is probably because the selection doctrine includes rather small concern for a man's welfare if we cannot use him in one of our established jobs. Not only do we throw people away, we largely put them out of our minds and off our conscience as well. But we do exploit them. Almost all the necessities of life cost more in an absolute as well as a relative sense for the people who live in poverty. Rents are higher. Food is higher. Similarly, the public services provided are of lesser quality. In fact, the very things that poverty-stricken people need most–their right to an opportunity to share in the affluency of the society–are denied them. This is all, I think, a part as well as a consequence of the selection doctrine in the use of men.

A classification doctrine would guarantee to all men an opportunity to work at a job with dignity. It would not prevent efforts to upgrade the skills of men, but if that failed, as it has failed in the past, then it would control the jobs so that all men would have an opportunity to perform one. I do not know how we can make the necessary changes in customs and beliefs or how we can circumvent the special interests that would be necessary to institute the new doctrine. I am convinced that we have the technology, and that we can afford the cost in efficiency, if there be a cost, to bring about the change.

Such a change in doctrine in the utilization of manpower will require an enormous capability to manage information. Computer systems will provide this capability. I think that the real impact of computers on personnel management is not that they will be of great service in furthering technical developments in our ability to predict future job performance, to do a better job in job engineering or even in the process of matching men to jobs, although they have added and will continue to add a great deal to our capability in this respect. Rather, I think that the really important impact of computers will be that they provide a necessary capability for us to implement a change from a selection doctrine in manpower utilization to a classification doctrine. Computer technology and use is not a sufficient capability in itself; we need in addition, the courage and the resolve to overcome parochial interests and to change customs and practices of long standing.

It is never very easy to bring about change in any form, and it is particularly difficult to effect changes in social systems. We have solutions of a sort to almost all of our social problems. But these solutions become outmoded, and while some basic changes do come by the long evolutionary process, most have occurred relatively quickly and by the relatively violent revolutionary process.

I think that we may have the capability to implement by evolutionary means a system which will more surely secure the rights of all men. If we subscribe to stated goals, and really use machines, we might someday arrive at the human use of human beings. If we do not get on with it, I fear it will come about despite us and by violent means.

REFERENCES

1. Bureau of the Census. *Statistical abstract of the U.S: 1965.*
2. Lewis, O. The culture of poverty. *Scientific American,* 1966, 215, 19-25.
3. Merrill, H.F. (Ed.) Classics in management. *American Management Association,* 1960, 29-47.

XVIII

THE IMPORTANCE OF PERSONNEL RESEARCH TO THE AIR FORCE

William B. Kieffer
Assistant Deputy Chief of Staff, Personnel
Department of the Air Force

When I was invited to speak before this group, I accepted very quickly. You gentlemen are working in an area in which I have a very great interest--an interest engendered from three basic causes:

First, I am a user of your product. Every day in my business I am faced with the need to make decisions and to give advice on problems and situations where all of the pertinent basic factors which should be taken into account are not clear and where basic factors which would be helpful to the consideration at hand are just not available. As I hope to bring out later, our effectiveness in managing our people hinges on techniques and procedures which have been developed as a result of personnel research accomplished in the past. We users need your help.

Second, I am a sponsor of personnel research. I am conscious of the large resources in manpower and money which the Air Force devotes to this activity, and I feel the need to do my part to assure that these resources are expended wisely and to ends which are useful in the real world.

Third, I am a human being and have a natural interest and curiosity about my own kind which heightens my interest in your work.

For these three reasons, then, I welcome the opportunity to express some of my feelings, beliefs, and attitudes toward personnel research and to make my observations to a group who are knowledgeable and will understand.

Several weeks ago as I began to organize my thoughts in preparation for this talk, I found that they were difficult for me to articulate. The subject is complex. I found myself with intuitions which were difficult to rationalize. My view of your complex profession is necessarily limited; my knowledge, incomplete; and my judgment, fallible. In short, I am conscious that I am only one of the blind men describing the elephant.

One of the definitions of research is that it is a critical and exhaustive investigation or experimentation having for its aim discovery of facts and their correct interpretation. In my experience, it is much harder to find a man with the facts than it is to find a man eager to interpret such facts as may be available. I continually find in myself and others a tendency to interpret partial facts and jump to questionable conclusions.

My first message to you, then, is that I strongly feel the need for many intelligent, patient, thorough, and hard working fact-gathers in the personnel research community.

Frair Mendel was a man willing to spend his life in long, tedious experimentation and fact-gathering over a period of many years. Our basic understanding of the laws of genetics grew from his painstaking efforts, and he greatly enriched our understanding of the nature of life. But this result only came about because he was willing to spend the bulk of his days performing painstaking, repetitive tasks. We need people like Mendel.

What is the specific importance of personnel research in the Air Force? It is trite to say that people are our most important resource. But, nevertheless, it is true. Let me talk a little about the Air Force and some of its characteristics as seen through the eyes of an Air Force personnel officer.

First, contrast this most modern of fighting services, the United States Air Force, with the fighting services of the past. In the armies of the past--from Alexander's through Grant's and Pershing's--virtually 100 per cent of the members were fighting men. Only a few basic skills were required. The emphasis was on leadership and courage in the face of the enemy. Every cook and wagon driver was prepared to drop the tools of his support trade, pick up his weapons, and participate in the fighting. Logistics problems were straightforward.

Contrast these organizations with the United States Air Force. Here is the most powerful military force the world has ever seen--and yet, of its 1.2 million people only 2 or 3 per cent can properly be characterized as fighting men. These few are the combat aircrew members who man our fighter and bomber aircraft. These aircrew members would be helpless, however, without the large numbers of support personnel who set the stage for their actions.

The multiplicity of tasks performed by this great force of people--and the variety of skills which are required of them--are such as to pose to the Air Force problems in personnel planning, complications in personnel training, and requirements in personnel management which are unprecedented in any present or past military organization.

We need jet pilots and master bakers, helicopter mechanics and policemen, bulldozer operators and teletype repairmen, personnel clerks and photo interpreters, ordnance disposal technicians and research and development specialists, language teachers and electronic technicians, and so it goes.

I have painted a picture of an Air Force consisting of a vast number of airmen, officers, and civilian employees with a great variety of different skills.

Let me describe a second characteristic of the Air Force. Its requirement for people in terms of numbers and skills and locations changes radically and on very short notice. These large-scale changes result from changes in force structure and in deployments, which in turn result from a rapidly flowering technology and from changes in the world's international political situation and in our Government's evaluation of the threat.

B-47s give way to B-58s and B-52s; Snarks and Atlases give way to Minutemen; bombers and interceptor squadrons give way to new fighter-bomber squadrons; forces shift from one part of the globe to another; bases close and others open; old communications systems and command and control systems give way to new. DeGaulle invites us out of France, a crisis requires our presence in Lebanon, in the Berlin Corridor, in Korea, over Cuba, in the Congo, and in Southeast Asia.

It is people who must respond to these ever-changing demands. Man is a tough and flexible creature, capable of adjusting to remarkable degrees of changes and of responding to remarkable challenges. Nevertheless, there are limits. Some lead time is required for training to new equipment and new tasks, for travel, for rest and recreation, for family care.

Where lead time is inadequate, we pay in terms of effectiveness. It becomes an important task of management to forecast and make provisions for changes and to make personnel plans which will permit us to accommodate them effectively with minimum disruption and hardship among our people.

Now a third and last characteristic before I get to the subject of personnel research. Our personnel force is characterized by rapid turnover. Three out of four airmen who join the Air Force make their contribution and depart within four years. More than half of the officers do the same. By the end of fiscal year '67 more than 31 per cent of our airmen and 25 per cent of our officers will have less than two years of service. During the course of the current fiscal year, our roughly 850,000 military personnel will experience almost 700,000 permanent changes of station.

All of this, the size and variety of skills, the changes in force structure and deployments, and the turnover within our force means that we must have the best of leadership–civilian, officer, and NCO leadership–wise, understanding, knowledgeable, stable, flexible, strong, and capable leaders. Secondly, this leadership must understand the human material with which we work–men and women–and employ the best possible techniques in recruiting, classifying, training, assigning, disciplining, promoting, and otherwise managing their activities.

Here is where personnel research can and does play an important part–a part which is more than ever needed–and where improvement in performance promises a big payoff.

What do we need?

First, we need improvement in recruiting techniques to permit us better to identify the young men most likely to be productive in the Air Force and better to motivate them toward careers in the Air Force.

Second, we need better methods for screening applicants for service as airmen, officers, and civilians of the Air Force in order to identify and reject those who do not have a reasonable chance of success in the Air Force. In this connection, of course, we have developed through the years our qualification and aptitude tests, and we have clearly identified and verified their validity in indicating the percentage chance for success of an individual. Additionally, we have data which makes it clear that if a young man is a high school graduate his chance for success as an enlisted man in the Air Force is much greater than if he is a dropout. Correspondingly, we know that a young man's chance of success as an Air Force officer is better if he is a college graduate than if he is not. Nevertheless, to date our tests are imperfect. We know that we accept many who fail and reject others who would indeed be successful, if given the chance. The problem of more refined selection tests and procedures is with us, and much progress has yet to be made.

At this point I would like to digress for a moment to talk about the impact on our work of our program to increase the numbers of individuals of lower mental categories who are accepted into military service. I am sure that none of us would regard with favor a program which results inevitably in enlistment into the Air Force large numbers of individuals who are doomed to failure. A young man's chance for success in life would likely be reduced rather than increased by a short experience in the Air Force, characterized by frustration and ending in separation as unsuitable or unfit. The challenge for us is to find means of identifying those individuals in the lower mental categories who can be made successful in the Air Force and, once they are in the Air Force, to so classify, train, and manage these people as to give each one the greatest possible opportunity in fact to succeed. This will be no easy task.

Third, we need to build on the past and improve our methods for classifying our new people and directing them into those areas of Air Force activities where they are likely to be most productive. The old "stanine" tests, devised for aircrew selection a quarter of a century ago, convinced us all that it is possible to realize great benefit from the use of aptitude tests. We have gone far since those days, but I am sure that no one thinks we are near the end of the road.

Fourth, we need research and experimentation leading to better techniques for training and cross-training of our people. Programmed learning techniques and the use of television and computers in education promise to improve our effectiveness in training.

Fifth, we need better techniques and procedures for evaluating our people so as to identify those most worthy for increased responsibility and promotion and to identify individual areas of weakness which can be addressed and corrected.

Sixth, we need better methods of forecasting our requirements for people. Some work has been started in this area.* Computers are being used to study force structure models which promise to help us in our forecasting and personnel planning tasks. We are, of course, only in our infancy in the use of this technique, but certainly, personnel planners of the future will profit greatly from it. A difficulty here, of course, is that of quantifying human emotions and predictive attributes for simulation and/or costing on a computer.

Seventh, better systems are required to match the man in the Air Force to the job in the Air Force. This implies that we must have detailed knowledge of the capabilities, experience, skills, and desires of each individual in the Air Force. Supplemental to this we must have better and more knowledge of the actual skills required.

Of obvious value is the use of punch card accounting systems (PCAM) and electronic data processing systems (EDP) to keep track of these complex data and to present them in ways useful to the personnel manager. Just a few miles from here at Randolph Air Force Base is to be found the heart of the new data system for the managing of Air Force personnel. The changeover of this data system from inefficient time-consuming hand operations coupled with manual reports to computer tapes and EDP is virtually complete. Much, however, remains to be done in the refinement of the system. Our next major milestone will come in 1968 when we hope to replace our PCAM equipment with computers at base-level. As we become more expert in the use of these modern tools we shall be able to use our personnel resources in ways more effective to the Air Force and more satisfying to the individual.

Eighth, to go along with improvements in our management system, we need to better our understanding of human psychology. We must manage the careers of our airmen, officers, and civilian employees in such a way as to give each individual the maximum possible opportunity to develop his capabilities to contribute to the Air Force and to experience a satisfying life for himself. Each should have an opportunity to serve in the optimum manner consistent with his capabilities and his reasonable aspirations. Here we are getting into an area where I must sound a note of caution.

It seems to me that we suffer from a plethora of what people call "attitudinal surveys." I am one who believes that you can tell more about a man by observing what he does than you can from listening to what he says, particularly if he says it in the context of an anonymous, and often times poorly conceived and uninteresting questionnaire. It has been my lot to review many questionnaires circulated among segments of our Air Force population by many different groups of people ranging from individual researchers circulating questionnaires at their own expense, to individual sections of the air staff, and to contractors to the Air Force. Frequently, these questionnaires contain questions with built-in answers, questions which are offensive, questions which can be readily interpreted as an invasion of privacy, questions which appear useless, questions the answers to which are readily available elsewhere, and so forth.

Behavioral science is, of course, a science, but it is a science in which it is extremely difficult to separate the wheat of real scientific value from the chaff of shallow observation.

---

* The start of this program is referred to in Colonel Ritter's review of Air Force personnel research.

Let me quote from a statement to the House Appropriations Committee by Dr. Robert L. Sproull, former Director of the Advanced Research Projects Agency: "I might add this is a very difficult field in which to be sure one is sponsoring only high quality work....It is a field in which it is very difficult indeed to tell the articulate linguist from the really promising scientist."*

Again, from the same report of a recent hearing of the House Appropriations Committee: "Some of the areas of study being pursued in behavioral sciences appear not to offer any real promise of providing useful information. Other studies appear to be concerned with trivial matters on which intelligent people should not require studies in order to be informed...."

All of us need to know more about ourselves, our people, and their motivations--more about why we behave like human beings and just how we can be expected to behave in varying situations. The behavioral scientist promises to give us answers to these difficult questions and to give us insights which will help us in our problems of morale, motivation, procurement, and, most importantly, retention.

We think we know those factors which motivate the Air Force individual the most: pride in himself and the organization; satisfaction in doing something which he knows to be worthwhile and more valuable than merely earning a living; the prospects of a good career; fair treatment and faith in an organization which appreciates him as an individual--these I believe are the great motivators of Air Force people. A fair break in pay, housing, recreational facilities, commissary, medical services, and so forth, all play their part, but the overriding factors are more moral than physical.

We do need from the behavioral scientist, however, help in showing how we can increase the capacity of the Air Force to inspire and to appeal to people in these moral areas. I am confident that this help will come, particularly as the communications improve between the behavioral scientist and the users of his developing knowledge of man.

In my job I am always hurting for lack of facts about our forces today and the history of our forces in the past.

I get back to Mendel--we need a large collection of facts. For example, how many officers with engineering degrees came into the Air Force in 1956? From what sources did they come? What were their aptitude scores? What did they study in college? How many became pilots and navigators? To what career fields were the others assigned? How many were offered regular commissions? How many accepted? How many were given Air Force Institute of Technology (AFIT) training? What grades have they reached? What efficiency records have they gained? What has been their year-by-year attrition from the Air Force? What have been the causes of this attrition? And so forth.

We need to know more and more facts like these. Happily, we have computers to help us, and these promise to help us even more in the years ahead. Unhappily, however, the work of collecting these facts is more laborious and in some ways less interesting than some other aspects of the work of personnel research. Nevertheless, from where I sit, this collection of facts is our most crying need in the personnel research area at this time.

Let me touch now a final aspect of our personnel mission to which properly conceived research can contribute.

---

* House of Representatives, Report No. 528 to accompany H.R. 9221, DoD Appropriation Bill, 1966.

In the development of all of our weapons and support systems, we must assure that our hardware developers keep clearly in mind that people will have to maintain and operate the equipment that is developed. This means not only in relatively simple matters, such as that the canopy should be designed in a way that the pilot can hold his head up without hitting it, and the instrumentation should be designed, mounted, and displayed in a way suitable to the human operator, but also that the long-term costs of maintenance and operation of the system should be considered against various options in the design stage. Sometimes more expensive equipment in design and procurement is less costly in the long run because it will take fewer people of lesser qualifications to maintain and operate it.*

As is well known, we face great problems of communication in many aspects of our lives. This problem is most significant in the vital coordination between personnel researchers and the users of their products. This communication, of course, is a two-way street. I am sure you will agree when I say that we users have not communicated our needs to the research community as loudly and clearly as we should. The researcher on the other hand has a great problem in expressing the results of his research in such ways that the user can understand it and employ it. All of us must dedicate ourselves toward the improvement of our mutual communication.

In closing let me state this fact: Every dollar that is appropriated for and spent by the Air Force goes into someone's pocket. More than $5 billion each year goes into the pockets of our military people. Another $2 billion goes to our civilian employees. More than $14 billion goes into the pockets of the people who sell us the goods and services we need to operate. People are expensive.

The personnel research community has contributed heavily in the past to the improvement of our understanding of this complex, precious, and indispensible resource. This research has helped us to use our people in ways more effective for the Air Force and more satisfying to our airmen, officers, and civilian employees. I know that the personnel research community will contribute even more heavily in the unknown but challenging future.

* Human engineering studies in the Air Force emanate chiefly from the Behavioral Sciences Laboratory at Wright-Patterson AFB, Ohio.

XIX

A LOOK AHEAD IN USAF PERSONNEL RESEARCH

Eugene T. Ferraro
Deputy Under Secretary of the Air Force for Manpower
Department of the Air Force

I am gratified to have been invited to share in marking this twenty-fifth milestone of performance by a very significant resource organization of the Air Force.

While I am a relative newcomer to the service and to this field of endeavor, I have come to know many of your fine officers and professionals of the Personnel Research Laboratory from my previous visits. I have gained a high regard for their work and an appreciation of their potential for vital contributions at a critical time.

This is a time of crisis–a crisis which is, as yet, clear only to a limited audience. But it is an audience which has direct responsibility. We face today a critical deficiency in skilled manpower, particularly in certain age brackets. This deficiency has an impact on the whole national scene and threatens our capacity to meet national objectives–industrial and social as well as military. It is especially critical in the Air Force.

The Psychological Research Unit, the predecessor of the Personnel Research Laboratory of today, was born during a period of crisis. Its pioneer effort in personnel research helped the United States to meet the tremendous challenge of World War II and pointed the way for more effective use of our training resources to meet the vast aircrew requirement of that era.

In July, 1941–only six weeks after the first research project was begun–testing was started to collect data on pilot aptitude. By December of 1941, these tests were being used as the basis for a classification system for pilots and other aircrew members as well. These Stanine tests, developed through personnel research, have been called one of the major contributions of the Army Air Forces in World War II. And they were only one of the many innovations which have resulted from personnel research during the last 25 years. You know better than I the many areas in which the Laboratory has pioneered.

Our past concern has been largely with selection and classification test development, that is, efforts to predict training performance and select the people with the best aptitude for the skills we required. Today, we face a new challenge–the challenge of *creating* aptitude and of molding the available manpower resource into a usable and useful resource. It is a challenge that will demand our most imaginative efforts.

It is no longer enough for us in the Air Force to determine our needs and go out looking for the manpower to meet them. The manpower is simply not there in the types and quantities to meet all the requirements of all users. There is no shortage of what some of my military associates call "warm bodies." But "warm bodies" are of little use unless they have or can be taught the skills necessary to meet our requirements. It is the resource of trained or trainable manpower that is limited. And we must compete with industry, the other branches of the military, and the many levels of government–all of whom have need for the same high caliber people.

Futhermore, all too often, the people whom we have trained succumb to the enticements of higher paying jobs with our competitors in the manpower market. This increases our training costs by forcing us to train replacements on a continuing basis, and this is especially true in the critical complex-skill areas.

The cost of training itself is ever rising. In 1952, it cost us $800 to give training in aircraft maintenance fundamentals. Today the same course of training costs us $2,520. The cost of training in Electronic Maintenance Fundamentals has risen from a 1952 price of $1,040 to $3,420. Part of this stems from the general rise in cost of living, but a large part of the increase is due to the increasing complexity of the hardware we use in the Air Force.

This complexity factor also requires the training of more individuals. The total number of specialties in the Air Force manpower inventory has risen from 285 in 1945 to 818 today. In 1945, one in three of these skills was electronics, mechanical, or technical. Today, the proportion is one in two. As an example of how skill requirements have increased, it took eight men to keep a P-47 flying during World War II. The F-111 will require about 25 men–more than three times as many.

A further complication is that skills become obsolete at an ever increasing pace. For example, no sooner did we get the first Atlas in its silo than we were training for the Minuteman. Our principal fighter at the moment is the F-4, but we are concerned with training to meet the skill requirements of the F-111. We are currently operating C-124s while we complete the training of our airlift crews in the C-141. This never ending cycle–the constant demand for new skills and for the refinement of old skills–creates a need for continual training and retraining of the manpower that is available. New methodology for quick, effective retraining is a must.

The limited manpower resource, rising costs, our increasing requirements, the growing complexity of weaponry, and the speed with which change breeds change–all these factors impel us to plan ahead. We must anticipate our needs so that the manpower resource will be trained to meet the requirements of the future. Our efforts must take advantage not only of our in-house capability and resources, but also of the spirit and fact of cooperation which exists now throughout the Department of Defense, throughout the Government, and throughout the nation. All segments of our national institutions must share their knowledge and pool their abilities to solve these national problems.

We must move forward from the base of past accomplishments to rebuild and energize our research efforts so that we can anticipate our requirements and develop programs which will meet them. More than ever before, we need research which will help man work with man, help machines work for man, and help man live with himself.

The Air Force has consistently been a leader in the field of personnel research. Regardless of this fact, however, I do not believe that our concern with personnel research has been commensurate with the need for concern. Our Personnel Research Laboratory should and will receive greater attention. This capability must be used to its fullest advantage for the benefit of the Air Force and the nation.

I visualize this Laboratory as a central focus which will unify our Air Force research efforts in the manpower and personnel area. I believe that we must strengthen the position of the Laboratory as an essential program element within the Air Force and the Department of Defense, that we must continue to enhance the stature of the Laboratory in the eyes of the professional world as a whole.

I intend to take an active interest in the activities of the Personnel Research Laboratory. I further intend to assist you in obtaining the support and recognition you need to do the job which the Air Force requires–support in funds and other resources. And, I will expect results from you in return.

I will support your efforts to insure that a certain part of the total effort will go into basic research. I will also expect the Laboratory to furnish consultant services as necessary to meet the problems of the current day.

In order to accomplish our objectives, we must keep the channels of communication open and flowing between the Laboratory and all elements of the Air Force and the professional community. This is particularly vital as it pertains to communication with my office where policies, developed as a result of your research, must ultimately be reviewed and decisions made on their implementation.

The manpower challenge of the present and the future demands that our research effort give special attention to motivation, especially as it relates to education and training. Only part of this problem is the responsibility of the Personnel Research Laboratory, but it is a vital part.

Our research effort must include a focus on imaginative concepts in the transmittal of ideas, facts, and knowledge among people. And our research workers must become increasingly interested in the interface between research and application. Researchers must visualize and understand the problems and practical considerations of the real world of the managers and operators who apply the findings of research.

Research must be a part of--not apart from--the effort to improve the training and utilization of the manpower resource. This is particularly true when it comes to training. The program teams, the managers in the Air Training Command must have access to the reports of current research.

I am looking to this Laboratory for some real breakthroughs as a result of research analysis of the valuable data the Air Force has collected. We have traced the careers of our people with minute attention to detail. Vast amounts of these data are available. But the data have no value while stored away. They have value only if we use them, make them more usable if possible, re-evaluate our methods of collection and storage to enhance their value. The solution to many of our problems may be waiting for us right here. Let's not get behind on this one.

While I hurl these challenges so freely, I am aware of the efforts that are currently in progress. Right now, we have many projects underway or in development:

A methodology to identify what people actually do on the job, a methodology which can be used to restructure technical training for optimum results.

Utilization of pre-service adjustment and performance information in selecting airmen for high-risk jobs.

Analysis of the force structure, evaluation of the impact of personnel policy changes on procurement, motivation, and retention.

Use of assignment and classification models to maximize utilization of our available personnel resources.

Identification and selection of low-aptitude groups with the greatest potential to develop skills in service.

In conclusion, I want to emphasize the growing importance of the research function as a part of Air Force manpower and personnel management. Research has been a basic element of the systems engineering approach we have applied so successfully in other Air Force areas such as research and development, procurement, and logistics. Now as we intensify our concern with manpower matters, as we optimize our operations in this vital area, we must apply the same systems engineering approach.

We must determine the methods that will fit the manpower resource to the specific needs of aerospace power. We must also determine how military service can serve society by improving the national manpower resource through education, training, work attitudes, skills, personal adjustment, etc. The pattern of the new look in the manpower area is just beginning to emerge. It is not my design alone but is a part of the larger effort of all the armed services. Careful planning and months of consultation have preceded the steps which are now being taken.

My office, that of the Deputy Under Secretary for Manpower, is just beginning to emerge its metamorphic state and take on a new being. We will provide the focal point for all matters relating to the manpower and personnel affairs of the Air Force. This includes military, Government civilians, not-for-profit employees of the Air Force, and defense contractors.

In a very real sense, we provide the principal interface for the Air Force on problems of manning for its tasks--the interface between the resources available and the techniques of utilization. With resources limited, the demands upon research to improve utilization become more intense. With your cooperation and support through productive research, we will open the door to new ideas, especially imaginative ones that could lead to new breakthroughs in more effective use of people.

The 25 years of productive research effort which we recognize today represents a record of which you and all of us in the Air Force are proud. Let us work together so that the coming years of maturity shall be the creative years that the Air Force needs.

XX

A FUTURE LOOK AT PERSONNEL RESEARCH

S. Rains Wallace
American Institute for Research

Ladies and gentlemen. I would like to point out to you that I have a certain problem. I would suspect that approximately 25 per cent of the people in this room have said to me, "Rains, it's been a long meeting, keep it light, give us some anecdotes, give us a little nostalgia. You know, do a little buck and wing." The other 75 per cent of you have said more directly and more to the point, "Rains, what time does your plane leave?" It reminds me of the story having to do with a young subaltern recently out of Sandhurst who was assigned the task of drilling a platoon in close-order drill; he started out rather reluctantly but soon began to gain confidence. There was a precipice at one end of the drill field, and here was the platoon marching toward it. He froze absolutely, couldn't think of a thing to say, while the beautifully drilled platoon continued toward the edge. Just before they reached the precipice, one file leader turned his head just a little bit and said, "Do say something, Sir, even if it's only goodbye."

You know, it is interesting that that story should have come to mind because, of course, it is a military story and I have been a little bit distressed ever since Launor Carter–I am sorry he is not still here to defend himself–talked about the difference between the military and the civilians. I want to tell you that we who were involved in forming the Laboratory are just about as military as any group of people can be. In fact, none of you has ever seen such military as I saw as a first lieutenant walking one pace behind and one pace to the left of Captain Meredith Crawford, Captain Eddie Ghiselli and Captain Philip DuBois, all in sun helmets. Military was military. As a matter of fact, they tell me that I was rather military when I first joined the Laboratory, and that's not surprising because, while I don't often admit this, I am a graduate of Morgan Park Military Academy for incorrigible boys. I am also a living example of the fact that Lloyd Humphreys hasn't looked at the right data. There are reliable predictors, and I am an outstanding example. I had gone to a military academy, and so when I received the telegram that said "You will proceed," I thought, "When I report to Kelly Field, I'll do it right–uniform and all." I went down and bought a uniform and, upon arrival at Kelly, reported to the post adjutant–which surprised him. I also asked for side arms because, in my day, you reported under side arms for duty. That surprised a lot of people. Finally, I found the Laboratory, walked in like a troop of cavalry, and reported to Major Robert D. Rock who has never been able to speak to me since. I was, in fact, the very mirror-image of a military officer. I say mirror-image because, as Bob Murphy gently pointed out to me about two or three days later, all my insignia were reversed.

But let us get back to our Chairman, Dr. Eddie Kemp. I was asked to do a little reminiscing and tell a few anecdotes. I think there's one anecdote that must be told about young Major Kemp. There was a time when Major Kemp was transferred to Midland Air Force Base which had been, among other problems, confronted by a Pentagon decision about what to do with bombardiers who had completed their tour with

the Eighth Air Force. Send them to Midland Air Force Base. The bombardiers did not appreciate this nor, I'm sure, did Major Kemp. But, there they were, and they were rather unhappy as Major Kemp discovered when he was made Officer of the Day. At approximately two o'clock in the morning, he got a call from–shall we say–a nightclub They had a young bombardier lieutenant who–shall we say–had become somewhat intoxicated and had also found that things were rather stuffy and warm. So he had removed all of his clothing and was prancing around in the nude. Well, the only thing the owner could do, of course, was to call the Officer of the Day. Major Kemp decided that he had better go out and try to solve this problem. Major Kemp went to the rescue. But, on his way back, driving in the jeep with the young lieutenant, he thought to himself, "Well, what am I going to do here? I've got to prefer some kind of charges against this chap because he has done the ...." At the same time, as I say, Major Kemp was a wise and understanding, participative sort of–well, anyway, what to do? You have to put this man on charges, how do it without getting him into real trouble. He solved the problem with his usual intrinsic brilliance. That young lieutenant was charged with being out of uniform.

You know, I recalled when Bob Murphy was telling about some of his problems in his early days that as a result of the numerous losses of platoons somebody got the idea of putting name cards on them. Just about that same time, there was a really brilliant idea–actually, I think the Navy first fomented it or supported it, and it seemed to come somewhere between the wheel and the internal combustion engine. I mean, it was a great thing. It was called the Renshaw method of identification of aircraft. For those of you who don't know about it, let me say this. The idea was that you just sat these kids down, flashed a picture of an airplane for one-sixtieth of a second, and asked them to identify it. So they would sit there, and the trainer of the group would say, "Ready? Now!" The picture would flash on for one-sixtieth of a second, and they were required to write down what kind of aircraft it was. Strangely enough, some of the aviation cadets didn't enjoy it. In fact, there was an incident, one of the great moments I think in the Air Force program, and I like to think that the trainer was Bob Murphy. Anyway, here's a young aviation cadet and there is Bob. He is saying, "Ready? Now!" Flash it on for one-sixtieth of a second, write it down. Finally the cadet can stand it no longer. He takes his pencil, throws it across the room and shouts, "Nuts to it, I'll be no more part of it." At which point Sgt. Murphy goes to him and says, "Mister, what's your name?" The kid looks at him, covers his badge, and says, "Ready? Now!"

Well, I'm supposed to say a few serious things. There are some serious things I guess I would like to say. "Ready? Now!"

I have been fascinated in listening to this collection of, I use the word advisedly, old pros in the personnel research field talking about their interests, their conceptualizations, their approaches to our various problems. Let me use just one example. There's been a great deal of talk about computers, understandably enough, throughout this program. But it fascinates me how differently our speakers have talked in terms of what we're going to do with them. Brokaw emphasizes their aid in moving us from an atomistic to a holistic approach. Kossack is interested in their application to personnel utilization models. particularly their employment in advance weapon systems. Leiman is interested in personnel utilization but after hearing his speech, I've decided what he really wants to do is to replace Sir Thomas More with computer. Launor Carter sees the computer as giving us potentials that are inherent in the individualizatio of personnel actions. Fields likes to talk about the estimation of qualitative, quantitative requirements, early research, and the development cycle of weapon systems. Our host Laboratory is sold on policy-capturing and occupational structuring and, in all honesty, the day I had to write an abstract to qualify to get here I got kind of fascinated with what we could do with computers in terms of analyzing the work situation as an important variable in the things we are trying to do. So it seems apparent to me that we may be looking and groping at the same elephant, but we are certainly seeing different things, and since what we see, even as we review the past and survey the present, has immediate implications for the future, I suggest to you that it follows that one man's prophecies are another's inconsequentia.

You know, the program you h: eard here does a much better job of looking into the future than I could possibly do for you even if I e: given unlimited time–and my plane does leave fairly soon–so I'm not moved to take whatever rema might make from now on very seriously. I would urge you to use similar restraint. I am going to try throw out some points very rapidly, both in the interest of time and to increase your difficulty in tearing them apart. Perhaps I can serve to remind you of some of the aspects of the future which your own thinking may have neglected, or which you may have missed from some of the speakers that we have had in this outstanding program.

I do want to say one other thing before I dash into my crystal ball look. I'm going to address myself from here on to general questions of personnel research or of psychological research because the history of this Laboratory is not only a history of Air Force research but of psychological research generally. I don't believe anyone really wants this Laboratory to address itself only to specific Air Force requirements because its history has been one of contribution to the general field, the general science of psychology. We have here with us today, Dr. John Flanagan, who started this program and had this in mind throughout. While a great deal has been said and should be said about Dr. Flanagan's contributions to this program, let me point out that in the process of contributing to it he also found and contributed to the training of most of the top research psychologists in the United States today. There are some notable exceptions which I shall not mention. So, I am going to talk more generally because I think it is in the tradition of this Laboratory, thanks to the vision that John Flanagan had back in those days, and I hope that it is a tradition that we shall not forget.

So what is the future? Where are we going?

I'm going to suggest to you that one of the things we are going to do with the computer is to find out more about the situations in which people work, and Lloyd Humphreys has pointed out that probably this set of variables is more important in the prediction of human performance than any of the sets of variables that we are looking at now.

The second thing that I would suggest to you is that we are going to have some new ideas about job-structuring. We started out in a rather elementary and, I think, a rather naive way. I think we are getting a little more sophisticated about it. I think that the next step in job-structuring is going to be the recognition that job structures and interpersonal relationships are in and of themselves together, and we must look at both sides of that coin as we structure jobs.

I'm going to suggest, thirdly, that it is time that selection and training research got together again. They used to be together. They have been separated now for some years. Somehow or other we have to get the individual differences boys and the experimental psychology boys together again to look at the problem of selection, training, and classification in joint context. I think that we are going to start talking about validity again in relationship to training, and it's about time. We used to do this, and we have forgotten it somehow. I think we are going to have to spend a lot more time, as this Laboratory is doing, in the development of better performance taxonomies. In fact, what I am saying here is what Fleishman has been saying very well, and it is time that we listened. I agree with Lloyd Humphreys that we must find out more about our test and predictive theory and stop worrying too much about the personalities we are testing. Lloyd frightened me a little bit with his talk because I thought maybe he was saying, "Let's just stop thinking about theory," but I discovered that indeed we are on the same track. Let us think about the theory of validity and what we mean by it, rather than theories of personality and what we do not mean by them. I think we are going to have to be more and more interested in problems having to do with the selection, the training, and the assignment of second-level management, both in the Armed Forces and in business and industry.

We talk a great deal about retention in the Armed Forces. In the old days, we used to talk a great deal about morale and motivation, and we had many studies of these problems. I think we know a lot more about this kind of problem and about new ways to approach it than we knew then, but somehow or other we have put it under the rug. I suggest to you that the future will demand much more attention to the old problems of morale, of motivation, of what keeps people interested in jobs, of how you keep them in jobs, and of how you compensate them. You know, money is still an important part of our society. I think we need a return to this kind of interest.

Finally, we must recognize that there are problems confronting personnel research that we have never seen before. There are new kinds of people that we must be looking at, new kinds of problems that we must be thinking about. We must be thinking about the transfer, for example, of personnel techniques across both national and international cultures. This is becoming a more and more urgent problem for consideration. We must understand better the relationship between what we do in selection and classification and how this is related to what we do in training. We are going to have to accept the fact that somebody is going to have to select and train people to work with many, many different populations in many, many different ways. Perhaps Dr. Leiman is right, but if he is, we have got a really difficult and vast job to do. It's nice to be here and to know that there are the kinds of people represented here who can face up to these problems.

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) Personnel Research Laboratory Lackland AFB, Texas 78236	2a. REPORT SECURITY CLASSIFICATION 2b. GROUP

3. REPORT TITLE
PERSONNEL RESEARCH AND SYSTEMS ADVANCEMENT

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

5. AUTHOR(S) (First name, middle initial, last name)
A.L. Fortuna (Editor)

6. REPORT DATE December 1967	7a. TOTAL NO. OF PAGES 182	7b. NO. OF REFS
8a. CONTRACT OR GRANT NO. AF 41(609)-3173 DO NOT USE b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) PRL-TR-67-13 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	12. SPONSORING MILITARY ACTIVITY Personnel Research Laboratory Lackland AFB, Texas 78236

13. ABSTRACT

Papers included in these Proceedings were presented at the Symposium on Personnel Research and Systems Advancement sponsored by the USAF Personnel Research Laboratory. Topics covered by the speakers include reviews of activities underway in the field of personnel research, as well as projections into future activities of researchers, planners, and managers in both academic and operational settings. The papers reflect a wide range of experience and interests of speakers representing the military services, governmental agencies, and private industry.

DD FORM 1 NOV 68 1473

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT