

VOL. 51, No. 4
JULY, 1954

Psychological Bulletin

THE CRITICAL INCIDENT TECHNIQUE

JOHN C. FLANAGAN

American Institute for Research and University of Pittsburgh

During the past ten years the writer and various collaborators have been engaged in developing and utilizing a method that has been named the "critical incident technique." It is the purpose of this article to describe the development of this methodology, its fundamental principles, and its present status. In addition, the findings of a considerable number of studies making use of the critical incident technique will be briefly reviewed and certain possible further uses of the technique will be indicated.

The critical incident technique consists of a set of procedures for collecting direct observations of human behavior in such a way as to facilitate their potential usefulness in solving practical problems and developing broad psychological principles. The critical incident technique outlines procedures for collecting observed incidents having special significance and meeting systematically defined criteria.

By an incident is meant any observable human activity that is sufficiently complete in itself to permit inferences and predictions to be made about the person performing the act. To be critical, an incident must occur in a situation where the purpose or intent of the act seems fairly clear to the observer and where its consequences are sufficiently definite to leave little doubt concerning its effects.

Certainly in its broad outlines and basic approach the critical incident technique has very little which is new about it. People have been making observations on other people for centuries. The work of many of the great writers of the past indicates that they were keen observers of their fellow men. Some of these writers must have relied on detailed notes made from their observations. Others may have had unusual abilities to reconstruct memory images in vivid detail. Some may have even made a series of relatively systematic observations on many instances of a particular type of behavior. Perhaps what is most conspicuously needed to supplement these activities is a set of procedures for analyzing and synthesizing such observations into a number of relationships that can be tested by making additional observations under more carefully controlled conditions.

BACKGROUND AND EARLY DEVELOPMENTS

The roots of the present procedures can be traced back directly to the studies of Sir Francis Galton nearly 70 years ago, and to later developments such as time sampling studies of recreational activities, controlled observation tests, and anecdotal records. The critical incident technique as such, however, can best be regarded as an outgrowth of studies in the Aviation Psychology Program of the United States Army Air Forces in World War II. The Aviation Psychology Program was established in the summer of 1941 to develop procedures for the selection and classification of aircrews.

One of the first studies (40) carried out in this program was the analysis of the specific reasons for failure in learning to fly that were reported for 1,000 pilot candidates eliminated from flight training schools in the summer and early fall of 1941. The basic source used in this analysis was the proceedings of the elimination boards. In these proceedings the pilot instructors and check pilots reported their reasons for eliminating the particular pilot. It was found that many of the reasons given were clichés and stereotypes such as "lack of inherent flying ability" and "inadequate sense of sustentation," or generalizations such as "unsuitable temperament," "poor judgment," or "insufficient progress." However, along with these a number of specific observations of particular behaviors were reported. This study provided the basis for the research program on selecting pilots. Although it was found very useful, it also indicated very clearly the need for better procedures for obtaining a representative sample of factual incidents regarding pilot performance.

A second study (13), which emphasized the importance of factual reports on performance made by competent observers, was carried out in the winter of 1943-1944 in the 8th, 9th, 12th, and 15th Air Forces. This study collected the reasons for the failures of bombing missions as reported in the Group Mission Reports.

Although in the preparation of these reports much greater emphasis was given to determining the precise facts in the case, it was apparent that in many instances the official reports did not provide a complete record of all the important events. Even with these limitations, the information given was found to be of considerable value, and the systematic tabulations that were prepared provided the basis for a series of recommendations that resulted in important changes in Air Force selection and training procedures.

In the summer of 1944 a series of studies (74) was planned on the problem of combat leadership in the United States Army Air Forces. These represent the first large-scale, systematic effort to gather specific incidents of effective or ineffective behavior with respect to a designated activity. The instructions asked the combat veterans to report incidents observed by them that involved behavior which was especially helpful or inadequate in accomplishing the assigned mission. The statement finished with the request, "Describe the officer's action. What did he do?" Several thousand incidents were collected in this way and analyzed to provide a relatively objective and factual definition of effective combat leadership. The resulting set of descriptive categories was called the "critical requirements" of combat leadership.

Another study (74) conducted in the Aviation Psychology Program involved a survey of disorientation while flying.¹ Disorientation in this study was defined to include any experience denoting uncertainty as to one's spatial position in relation to the vertical. In this study pilots returning from combat were asked "to think of some occasion during combat flying in which you personally experienced feelings of acute disorientation or strong vertigo." They were then asked to describe what they "saw, heard, or felt that brought on the experience." This study led to a number of recommendations regarding changes in cockpit and instrument panel design and in training in order to overcome and prevent vertigo while flying.

In a project carried out in the Aviation Psychology Program in 1946, Fitts and Jones (12) collected descriptions of specific experiences from pilots in taking off, flying on instruments, landing, using controls, and using instruments. These interviews with pilots were electrically recorded. They provided many factual incidents that were used as a basis for planning research on the design of instruments and controls and the arrangement of these within the cockpit.

In addition to the collection of specific incidents and the formulation of critical requirements, as outlined above, the summary volume (13) for the Aviation Psychology Program Research Reports contained a discussion of the theoretical basis of procedures for obtaining the critical requirements of a particular activity. Perhaps the best method of describing the status of these procedures at the close of the war is to quote from the discussion in this summary volume, which was written in the late spring of 1946. In the section on techniques for defining job requirements, the present author wrote as follows:

The principal objective of job analysis procedures should be the determination of critical requirements. These requirements include those which have been demonstrated to have made the difference between success and failure in carrying out an important part of the job assigned in a significant number of instances. Too often, statements regarding job requirements are merely lists of all the desirable traits of human beings. These are practically no help in selecting, classifying, or training individuals for specific jobs. To obtain valid information regarding the truly critical requirements for success in a specific assignment, procedures were developed in the Aviation Psychology Program for making systematic analyses of causes of good and poor performance.

Essentially, the procedure was to obtain first-hand reports, or reports from objective records, of satisfactory and unsatisfactory execution of the task assigned. The cooperating individual described a situation in which success or failure was determined by specific reported causes.

This procedure was found very effective in obtaining information from individuals concerning their own errors, from subordinates concerning errors of their superiors, from supervisors with respect to their subordinates, and also from participants with respect to co-participants (13, pp. 273-274).

DEVELOPMENTAL STUDIES AT THE AMERICAN INSTITUTE FOR RESEARCH

¹ This study was planned by Paul M. Fitts, Jr., who also contributed to the previously mentioned USAAF studies and planned and carried out the interview study with pilots described below on the design of instruments, controls, and arrangements.

At the close of World War II some of the psychologists who had participated in the USAAF Aviation Psychology Program established the American Institute for Research, a nonprofit scientific and educational organization. The aim of this organization is the systematic study of human behavior through a coordinated program of scientific research that follows the same general principles developed in the Aviation Psychology Program. It was in connection with the first two studies undertaken by the Institute in the spring of 1947 that the critical incident technique was more formally developed and given its present name.

These studies were natural extensions of the previous research in the Aviation Psychology Program. The study reported by Preston (52) dealt with the determination of the critical requirements for the work of an officer in the United States Air Force. In this study, many of the procedural problems were first subjected to systematic tryout and evaluation. Six hundred and forty officers were interviewed, and a total of 3,029 critical incidents were obtained. This led to the development of a set of 58 critical requirements classified into six major areas. The second study, reported by Gordon (27, 28), was carried out to determine the critical requirements of a commercial airline pilot. In this study, several different sources were used to establish the critical requirements of the airline pilot. These included training records, flight check records including the specific comments of check pilots, critical pilot behaviors reported in accident records, and critical incidents reported anonymously in interviews by the pilots themselves. From this study, 733 critical pilot behaviors were classified into 24 critical requirements of the airline pilot's job. These were used to develop selection tests to measure the aptitudes and other personality characteristics found critical for success in the job. They also provided the basic data for the formulation of an objective flight check to determine the eligibility of applicants for the airline transport rating.

The third application of the critical incident technique by the staff of the American Institute for Research was in obtaining the critical requirements for research personnel on a project sponsored by the Psychological Sciences Division of the Office of Naval Research. In this study (20), about 500 scientists in 20 research laboratories were interviewed. These scientists reported more than 2,500 critical incidents. The critical behaviors were used to formulate inductively a set of 36 categories, which constitutes the critical requirements for the effective performance of the duties of research personnel in the physical sciences. This initial study provided the basis for the development of selection tests, proficiency measures, and procedures for evaluating both job performance and the research report.

Another project undertaken by the American Institute for Research in the spring of 1948 provided valuable experience with the critical incident technique. This study, reported by Nagay (48), was done for the Civil Aeronautics Administration under the sponsorship of the Committee on Aviation Psychology of the National Research Council. It was concerned with the air route traffic controller's job. One of the innovations in this study was the use of personnel of the Civil Aeronautics Administration who had no previous psychological training in collecting critical incidents by means of personal interviews. In previous studies all such interviewing had been conducted by psychologists with extensive training in such procedures. In this study, aeronautical specialists from each of the seven regions conducted the interviews in their regions after a brief training period. An interesting finding from this study was the clear reflection of seasonal

variations in flying conditions in the types of incidents reported. The study also demonstrated the selective recall of dramatic or other special types of incidents. This bias was especially noticeable in the incidents reported several months after their occurrence. The incidents obtained in this study were used to develop procedures for evaluating the proficiency of air route traffic controllers and also for developing a battery of selection tests for this type of personnel.

In the spring of 1949 the American Institute for Research undertook a study to determine the critical job requirements for the hourly wage employees in the Delco-Remy Division of the General Motors Corporation. This study, reported by Miller and Flanagan (46), was the first application of these techniques in an industrial situation. Foremen who were members of a committee appointed to develop employee evaluation procedures collected 2,500 critical incidents in interviews with the other foremen in the plants. On the basis of these data a form was prepared for collecting incidents on a day-to-day basis as a continuous record of job performance.

Using this form, the Performance Record for Hourly Wage Employees (21), three groups of foremen kept records on the performance of their employees for a two-week period. A group of 24 foremen recorded incidents daily; another group of 24 foremen reported incidents at the end of each week; and a third group containing the same number of foremen reported incidents only at the end of the two-week period. The three groups of foremen represented comparable conditions of work and supervision. The foremen reporting daily reported 315 critical incidents; the foremen reporting weekly, 155 incidents; and the foremen reporting only once at the end of two weeks reported 63 incidents. Thus, foremen who reported only at the end of the week had forgotten approximately one half of the incidents they would have reported under a daily reporting plan. The foremen who reported only at the end of the two-week period appeared to have forgotten 80 per cent of the incidents observed. Although it is possible that the findings may be partially attributed to the fact that the foremen making daily records actually observed more critical incidents because of the daily reminder at the time of recording, it is clear that much better results can be expected when daily recording is used.

Another analysis based on data collected at the Delco-Remy Division compared the number of critical incidents of various types obtained from interviews with those recorded daily by the foremen on the performance record. Although there were some differences in the relative frequencies for specific categories, the general patterns appeared to be quite similar. These results suggest that critical incidents obtained from interviews can be relied on to provide a relatively accurate account of job performance if suitable precautions are taken to prevent systematic bias.

In addition to the development of the performance record described above, the critical incidents collected in this study were used as the basis for constructing selection tests covering both aptitude (18) and attitude (2) factors.

STUDIES CARRIED OUT AT THE UNIVERSITY OF PITTSBURGH

A substantial number of studies have been carried out in the department of psychology at the University of Pittsburgh by students working for advanced degrees under the author's direction. Most of these studies had as their objective the determination

of the critical requirements for a specific occupational group or activity. Many of them also included contributions to technique. In 1949 Wagner (66) completed a dissertation on the critical requirements for dentists. In this study, critical incidents were obtained from three sources: patients, dentists, and dental school instructors. The incidents were classified into four main aspects of the dentist's job: (a) demonstrating technical proficiency; (b) handling patient relationships; (c) accepting professional responsibility; and (d) accepting personal responsibility. As might be expected, the patients did not report as large a proportion of incidents for demonstrating technical proficiency or accepting professional responsibility as did the other two groups, and the instructors reported only a relatively small proportion of their incidents in the area of handling patient relationships.

On the basis of the findings from this study, a battery of selection tests was developed for use by the University of Pittsburgh School of Dentistry. A number of proficiency tests for measuring ability with respect to certain of the critical requirements were also developed using these results as a basis.

Another dissertation completed in 1949 was Finkle's (11) study of the critical requirements of industrial foremen. This study was conducted in the East Pittsburgh plant of the Westinghouse Electric Corporation. Critical incidents were obtained from foremen, general foremen, and staff personnel. A number of points pertaining to technique were studied.

One finding was in reference to the effect on the types of incidents obtained of the degree of importance or exceptionalness set up as a criterion for reporting or ignoring incidents. The incidents obtained from the use of questions that asked for incidents only slightly removed from the norm were compared with incidents obtained from questions intended to elicit more definitely effective or ineffective behaviors. Some examples of these questions are:

1. Think of a time when a foreman has done something that you felt should be encouraged because it seemed to be in your opinion an example of good foremanship. (Effective—slight deviation from norm.)
2. Think of a time when a foreman did something that you thought was not up to par. (Ineffective—slight deviation from norm.)
3. Think of a time when a foreman has, in your opinion, shown definitely good foremanship—the type of action that points out the superior foreman. (Effective—substantial deviation from the norm.)
4. Think of a time when a foreman has, in your opinion, shown poor foremanship—the sort of action which if repeated would indicate that the man was not an effective foreman. (Ineffective—substantial deviation from norm.)

The frequencies of incidents obtained in each of the 40 categories into which the effective behaviors were classified were compared for the questions requesting slight and substantial deviations from the norm, and the significance of the differences was tested by means of the chi-square test. Two of the differences were significant at the 1 per cent level and one at the 5 per cent level. Comparisons of the frequencies in each of the 40 categories for ineffective incidents failed to reveal any chi squares significant at either the 5 per cent or the 1 per cent level.

The questions involving only a slight deviation from the norm resulted in more effective incidents concerned with gaining the respect and loyalty of the workers and also in more incidents that involved making, encouraging, and accepting suggestions. They

produced significantly fewer incidents regarding fitting men to jobs. The small number of significant differences—only three in 80 comparisons—suggests that the types of incidents obtained are not very greatly changed by variations in wording of the questions comparable to those shown above. It seems likely that this is at least partially due to the fact that the persons interviewed report only incidents that represent a fairly substantial deviation from the norm regardless of the precise wording of the question asked.

Another comparison made in this study related to the influence of asking for an effective or an ineffective incident first. About 10 per cent more incidents were obtained from booklets requesting effective incidents first than from booklets requesting ineffective incidents first. This difference was sufficiently small so that it could reasonably be attributed to chance sampling fluctuations.

The incidents collected in this study were used, along with other data, in the preparation of a Performance Record for Foremen and Supervisors (23).

A study was conducted by Nevins (50) on the critical requirements of bookkeepers in sales companies. She collected incidents relating to applicants for bookkeeping positions as well as for employees working in this capacity.

For the collection of the information about the practicing bookkeepers, a modification in the critical incident technique was made. This was done because, in the bookkeeping profession, success and failure are usually defined in terms of persistent behavioral patterns. Occasional mistakes in adding and balancing accounts are expected, but repeated errors are considered serious. Instead of the single incident, therefore, many of the items included represented either a pattern of behaviors or a series of similar behaviors.

Weislogel (72) determined the critical requirements for life insurance agency heads. A principal feature of his study related to the comparison of two types of agency heads—managers and general agents. It was believed that the critical behaviors for one type of agency head might provide a different pattern than that obtained for the other. This hypothesis was not confirmed by the analysis of the obtained incidents. The patterns of critical requirements were found to be quite similar for the two types of administrators.

Smit (58) carried out a study to determine the critical requirements for instructors of general psychology courses. Perhaps the finding of most general importance in this study was the existence of substantial differences between the patterns of critical incidents reported by students and faculty. The faculty reported a significantly larger percentage of effective behaviors in the following areas: giving demonstrations or experiments, using discussion group techniques, encouraging and ascertaining students' ideas and opinions.

The students, on the other hand, contributed a larger percentage of behaviors in the following areas: reviewing examinations, distributing grades, and explaining grades; using lecture aids such as drawings, charts, movies, models, and apparatus; using project techniques; giving test questions on assigned material; helping students after class and during class recess; the manner of the instructor.

The faculty reported a larger percentage of ineffective behaviors concerning maintaining order. The ineffective behaviors that were reported in a larger percentage by students involved these areas: presenting requirements of the course, using effective methods of expression, dealing with students' questions, pointing out fallacies, reviewing and summarizing basic facts and principles, using project techniques, using verbal diag-

nostic teaching techniques, achievement testing students on assigned material, objective type achievement testing, using humor.

This is a good illustration of the problem of the competence of various types of available observers to evaluate the contribution to the general aim of the activity of a specific action. Examination of the reports from students indicated a somewhat limited sphere of competence. Apparently one of the principal reasons for this was the lack of perspective on the part of the students and their inability to keep the general aim of the instructor clearly in mind because of its divergence from their own immediate aims. In many cases, this latter aim seemed to be directed toward achieving a satisfactory grade in the course.

Eilbert (7) developed a functional description of emotional immaturity. The contributors of critical incidents included psychiatrists, psychologists, psychiatric social workers, occupational therapists, nurses, and corps-men from a military hospital, plus 13 psychologists in nonmilitary organizations. The subjects of the incidents were primarily patients under psychiatric care.

The contributors were given a form that oriented them to the concept "emotional immaturity" by suggesting that it was revealed generally by childlike modes of behavior. The questions used to elicit incidents were: Have you recently thought of someone as being emotionally immature (regardless of diagnosis)? What specifically happened that gave you this impression? What would have been a more mature reaction to the same situation?

Because of the indefinite nature of the concept, it was felt that a check should be made on the contributor's understanding of his task. Twenty of the participating persons were asked to summarize briefly their interpretation of what they had been asked to do. This appeared to be very useful in developing the phrasing of the questions so that they were uniformly interpreted by the various observers.

The author of the study classified all the immaturities on the basis of a classification system developed from preliminary categorizations prepared by six of the contributors. This classification was submitted to 14 psychiatrists for review. They were asked to indicate which of the categories they were willing to accept as a type of immaturity as the term had been defined in an official document. More than half the categories were accepted by at least 13 of the 14 judges, and none was rejected by more than 50 per cent of the judges. It was felt then that the system was acceptable.

This study illustrates the application of the critical incident technique to the study of personality. It is believed that this study provides an excellent example of the possibilities for developing more specific behavioral descriptions.

Folley (24) reported on the critical requirements of sales clerks in department stores. The behaviors were abstracted from narrative records of individual shopping incidents written by shoppers who were relatively inexperienced in evaluating sales personnel. For various reasons, including the competence of the observers, their training, and their limited point of view, the resulting description must be regarded as only partial.

In the past few years, many other individuals and groups have made use of the techniques described above, or modifications of them, in a wide variety of studies. Some of these studies on which reports are being published will be reviewed briefly in the section on applications.

THE PROCEDURE IN ITS PRESENT FORM

From the foregoing discussion, it is clear that the critical incident technique is essentially a procedure for gathering certain important facts concerning behavior in defined situations. It should be emphasized that the critical incident technique does not consist of a single rigid set of rules governing such data collection. Rather it should be thought of as a flexible set of principles which must be modified and adapted to meet the specific situation at hand.

The essence of the technique is that only simple types of judgments are required of the observer, reports from only qualified observers are included, and all observations are evaluated by the observer in terms of an agreed upon statement of the purpose of the activity. Of course, simplicity of judgments is a relative matter. The extent to which a reported observation can be accepted as a fact depends primarily on the objectivity of this observation. By objectivity is meant the tendency for a number of independent observers to make the same report. Judgments that two things have the same effect or that one has more or less effect than the other with respect to some defined purpose or goal represent the simplest types of judgments that can be made. The accuracy and therefore the objectivity of the judgments depend on the precision with which the characteristic has been defined and the competence of the observer in interpreting this definition with relation to the incident observed. In this latter process, certain more difficult types of judgments are required regarding the relevance of various conditions and actions on the observed success in attaining the defined purpose for this activity.

It is believed that a fair degree of success has been achieved in developing procedures that will be of assistance in gathering facts in a rather objective fashion with only a minimum of inferences and interpretations of a more subjective nature. With respect to two other steps that are essential if these incidents are to be of value a comparable degree of objectivity has not yet been obtained. In both instances, the subjective factors seem clearly due to current deficiencies in psychological knowledge.

The first of these two other steps consists of the classification of the critical incidents. In the absence of an adequate theory of human behavior, this step is usually an inductive one and is relatively subjective. Once a classification system has been developed for any given type of critical incidents, a fairly satisfactory degree of objectivity can be achieved in placing the incidents in the defined categories.

The second step refers to inferences regarding practical procedures for improving performance based on the observed incidents. Again, in our present stage of psychological knowledge, we are rarely able to deduce or predict with a high degree of confidence the effects of specific selection, training, or operating procedures on future behaviors of the type observed. The incidents must be studied in the light of relevant established principles of human behavior and of the known facts regarding background factors and conditions operating in the specific situation. From this total picture hypotheses are formulated. In only a few types of activities are there both sufficient established principles and sufficient information regarding the effective factors in the situation to provide a high degree of confidence in the resulting hypotheses regarding specific procedures for improving the effectiveness of the results.

In the sections which follow, the five main steps included in the present form of the procedures will be described briefly. In order to provide the worker with maximum flexi-

bility at the present stage, in addition to examples of present best practice, the underlying principles for the step will be discussed and also the chief limitations with, wherever possible, suggestions for studies that may result in future improvements in the methods.

1. General Aims

A basic condition necessary for any work on the formulation of a functional description of an activity is a fundamental orientation in terms of the general aims of the activity. No planning and no evaluation of specific behaviors are possible without a general statement of objectives. The trend in the scientific field toward operational statements has led a number of writers to try to describe activities or functions in terms of the acts or operations performed, the materials acted on, the situations involved, the results or products, and the relative importance of various acts and results. These analyses have been helpful in emphasizing the need for more specific and detailed descriptions of the requirements of activities. Typically, however, such discussions have failed to emphasize the dominant role of the general aim in formulating a description of successful behavior or adjustment in a particular situation.

In its simplest form, the functional description of an activity specifies precisely what it is necessary to do and not to do if participation in the activity is to be judged successful or effective. It is clearly impossible to report that a person has been either effective or ineffective in a particular activity by performing a specific act unless we know what he is expected to accomplish. For example, a supervisor's action in releasing a key worker for a half a day to participate in a recreational activity might be evaluated as very effective if the general aim of the foreman was to get along well with the employees under him. On the other hand, this same action might be evaluated as ineffective if the primary general aim is the immediate production of materials or services.

In the case of the usual vocational activities the supervisors can be expected to supply this orientation. In certain other types of activities, such as civic, social, and recreational activities, there frequently is no supervisor. The objectives of participation in the activity must then be determined from the participants themselves. In some instances, these may not be verbalized to a sufficient extent to make it possible to obtain them directly.

Unfortunately, in most situations there is no one general aim which is the correct one. Similarly, there is rarely one person or group of persons who constitute an absolute, authoritative source on the general aim of the activity. In a typical manufacturing organization the foreman, the plant manager, the president, and the stockholders might define the general aim of the workers in a particular section somewhat differently. It is not possible to say that one of these groups knows the correct general aim and the others are wrong. This does not mean that one general aim is as good as another and that it is unimportant how we define the purpose of the activity. It does mean that we cannot hope to get a completely objective and acceptable general aim for a specific activity. The principal criterion in formulating procedures for establishing the general aim of the activity should be the proposed use of the functional description of the activity which is being formulated. Unless the general aim used is acceptable to the potential users of the detailed statement of requirements, the whole effort in formulating this statement will have been wasted.

The most useful statements of aims seem to center around some simple phrase or catchword which is slogan-like in character. Such words provide a maximum of communication with only a minimum of possible misinterpretation. Such words as "appreciation," "efficiency," "development," "production," and "service" are likely to be prominent in statements of general aims. For example, the general aim of a teacher in elementary school art classes might be the development of an appreciation of various visual art forms on the part of the students. The general aim of the good citizen might be taken as effective participation in the development and application of the rules and procedures by which individuals and groups are assisted in achieving their various goals.

With the aid of a form of the type shown in Fig. 1, the ideas of a number of well-qualified authorities can be collected. It is expected that in response to the question on the

**OUTLINE FOR INTERVIEW TO
ESTABLISH THE GENERAL
AIM FOR AN ACTIVITY**

1. *Introductory statement:* We are making a study of (specify activity). We believe you are especially well qualified to tell us about (specify activity).
2. *Request for general aim:* What would you say is the primary purpose of (specify activity)?
3. *Request for summary:* In a few words, how would you summarize the general aim of (specify activity)?

FIG. 1. SAMPLE FORM FOR USE IN OBTAINING GENERAL AIM

primary purpose of the activity many persons will give a fairly lengthy and detailed statement. The request to summarize is expected to get them to condense this into a brief usable statement. These should be pooled and a trial form of the statement of general aim developed. This statement should be referred either to these authorities or to others to obtain a final statement of the general aim that is acceptable to them. Necessary revisions should be made as indicated by these discussions. Usually considerable effort is required to avoid defeating the purpose of the general aim by cluttering up the statement with specific details and qualifying conditions.

In summary, the general aim of an activity should be a brief statement obtained from the authorities in the field which expresses in simple terms those objectives to which most people would agree. Unless a brief, simple statement has been obtained, it will be difficult to get agreement among the authorities. Also it will be much harder to convey a uniform idea to the participants. This latter group will get an over-all impression and this should be as close to the desired general aim as possible.

2. Plans and Specifications

To focus attention on those aspects of behavior which are believed to be crucial in formulating a functional description of the activity, precise instructions must be given to the observers. It is necessary that these instructions be as specific as possible with respect to the standards to be used in evaluation and classification. The group to be studied also needs to be specified.

One practical device for obtaining specific data is to obtain records of "critical incidents" observed by the reporting personnel. Such incidents are defined as extreme

behavior, either outstandingly effective or ineffective with respect to attaining the general aims of the activity. The procedure has considerable efficiency because of the use of only the extremes of behavior. It is well known that extreme incidents can be more accurately identified than behavior which is more nearly average in character.

One of the primary aims of scientific techniques is to insure objectivity for the observations being made and reported. Such agreement by independent observers can only be attained if they are all following the same set of rules. It is essential that these rules be clear and specific. In most situations the following specifications will need to be established and made explicit prior to collecting the data:

a. The situations observed. The first necessary specification is a delimitation of the situations to be observed. This specification must include information about the place, the persons, the conditions, and the activities. Such specifications are rather easily defined in many instances. For example, such brief specifications as observations of "the behavior in classrooms of regularly employed teachers in a specified high school while instructing students during class periods," constitute a fairly adequate definition of a situation of this type.

In complex situations it is probably essential not only that the specifications with respect to the situation be relatively complete and specific, but also that practical examples be provided to assist the observer in deciding in an objective fashion whether or not a specific behavior should be observed and recorded.

b. Relevance to the general aim. After the decision has been made that a particular situation is an appropriate one for making observations, the next step is to decide whether or not a specific behavior which is observed is relevant to the general aim of the activity as defined in the section above. For example, if the general aim of the activity was defined as sustained high quality and quantity of production, it might be difficult to decide whether or not to include an action such as encouraging an unusually effective subordinate to get training that would assist him in developing his ability in an avocational or recreational activity not related to his work. In this case, it might be specified that any action which either directly or indirectly could be expected over a long period of time to have a significant effect on the general aim should be included. If it could not be predicted with some confidence whether this effect would be good or bad, it should probably not be considered.

The extent of detail required to obtain objectivity with respect to this type of decision depends to a considerable degree on the background and experiences of the observers with respect to this activity. For example, supervisors with substantial experience in a particular company can be expected to agree on whether or not a particular behavior is relevant to the attainment of the general aim. On the other hand, if outside observers were to be used, it would probably be necessary to specify in considerable detail the activities that can be expected to have an effect on the general aim.

c. Extent of effect on the general aim. The remaining decision that the observer must make is how important an effect the observed incident has on the general aim. It is necessary to specify two points on the scale of importance: (a) a level of positive contributions to the general aim in specific terms, preferably including a concrete example, and (b) the corresponding level of negative effect on the general aim expressed in similar terms.

A definition which has been found useful is that an incident is critical if it makes a "significant" contribution, either positively or negatively, to the general aim of the activity. The definition of "significant" will depend on the nature of the activity. If the general aim of the activity is in terms of production, a significant contribution might be one which caused, or might have caused, an appreciable change in the daily production of the department either in the form of an increase or a decrease. In certain specific situations, it might be desirable and possible to set up a quantitative criterion such as saving or wasting 15 minutes of an average worker's production. In some situations, a definition of significance might be set up in terms of dollars saved or lost both directly and indirectly.

Actions which influence the attitudes of others are more difficult to evaluate objectively. Perhaps the best we might be able to do is to state it in terms of a probability estimate. For

example, one such criterion might be that the minimum critical level would be an action that would have an influence such that at least one person in ten might change his vote on an issue of importance to the company.

d. Persons to make the observations. One additional set of specifications refers to the selection and training of the observers who are to make and report the judgments outlined in the steps above.

Wherever possible, the observers should be selected on the basis of their familiarity with the activity. Special consideration should be given to observers who have made numerous observations on persons engaged in the activity. Thus, for most jobs, by far the best observers are supervisors whose responsibility it is to see that the particular job being studied is done. However, in some cases very useful observations can be contributed by consumers of the products and services of the activity. For example, for a study of effective sales activities, the customers may have valuable data to contribute. For a study of effective parental activity, the children may be able to make valuable contributions.

In addition to careful selection of the persons to make observations, attention should be given to their training. Minimal training should include a review of the nature of the general aim of the activity and a study of the specifications and definitions for the judgments they will be required to make. Where the situation is complex or the observers are not thoroughly familiar with the activity, supervised practice in applying these definitions should be provided. This can be done by preparing descriptions of observations and asking the observers to make judgments about these materials. Their judgments can be immediately confirmed or corrected during such supervised practice periods.

In Fig. 2 is shown a form for use in developing specifications regarding observations. The use of this form in making plans for the collection of critical incidents or other types of observational data should aid in objectifying these specifications.

3. Collecting the Data

If proper plans and specifications are developed, the data collection phase is greatly simplified. A necessary condition for this phase is that the behaviors or results observed be evaluated, classified, and recorded while the facts are still fresh in the mind of the observer. It would be desirable for these operations to be

<p><i>Specifications Regarding Observations</i></p> <ol style="list-style-type: none"> 1. Persons to make the observations. <ol style="list-style-type: none"> a. Knowledge concerning the activity. b. Relation to those observed. c. Training requirements. 2. Groups to be observed. <ol style="list-style-type: none"> a. General description. b. Location. c. Persons. d. Times. e. Conditions. 3. Behaviors to be observed. <ol style="list-style-type: none"> a. General type of activity. b. Specific behaviors. c. Criteria of relevance to general aim. d. Criteria of importance to general aim (critical points)
--

FIG. 2. FORM FOR DEVELOPING SPECIFICATIONS REGARDING OBSERVATIONS

performed at the time of observation so that all requisite facts could be determined and checked. Memory is improved if it is known in advance that the behavior to be observed is to be remembered. It is greatly improved if the specific aspects of what is to be observed are defined and if the operations to be performed with respect to evaluation and classification are clearly specified.

The critical incident technique is frequently used to collect data on observations previously made which are reported from memory. This is usually satisfactory when the incidents reported are fairly recent and the observers were motivated to make detailed observations and evaluations at the time the incident occurred.

The importance of obtaining recent incidents to insure that the incidents are representative of actual happenings was demonstrated in the study on air route traffic controllers by Nagay (48) reported above. However, as also discussed in that study, in some situations adequate coverage cannot be obtained if only very recent incidents are included.

Evidence regarding the accuracy of reporting is usually contained in the incidents themselves. If full and precise details are given, it can usually be assumed that this information is accurate. Vague reports suggest that the incident is not well remembered and that some of the data may be incorrect. In several situations there has been an opportunity to compare the types of incidents reported under two conditions (*a*) from memory and without a list of the types of incidents anticipated, and (*b*) those reported when daily observations were being made in a routine work situation, and the evaluations and classifications were made and recorded on a prepared check list within 24 hours of the time of observation. The results of one such comparison were discussed briefly above in connection with the American Institute for Research study of factory employees.

During the observational period a negligible number of incidents were reported by the foremen as not fitting into the general headings included on the list. Although the proportions of incidents for the various items on the list are not identical, they are reason-

ably close for most of the items. Items on such matters as meeting production requirements and accepting changes in jobs are higher in terms of the recorded than the recalled incidents. The fact that items such as wasting time and assisting on problems are lower for the recalled incidents suggests that part of this discrepancy lies in the interpretations of the category definitions. The classifying of recorded incidents was done by the foremen, while the classification of the recalled incidents was done by the research workers. In fairness, it should also be noted that the definitions used by the research workers were rewritten before they were incorporated in the foremen's manuals.

On the whole, it seems reasonable to assume that, if suitable precautions are taken, recalled incidents can be relied on to provide adequate data for a fairly satisfactory first approximation to a statement of the requirements of the activity. Direct observations are to be preferred, but the efficiency, immediacy, and minimum demands on cooperating personnel which are achieved by using recalled incident data frequently make their use the more practical procedure.

Another practical problem in collecting the data for describing an activity refers to the problem of how it should be obtained from the observers. This applies especially to the problem of collecting recalled data in the form of critical incidents. Four procedures have been used and will be discussed briefly below:

a. Interviews. The use of trained personnel to explain to observers precisely what data are desired and to record the incidents, making sure that all necessary details are supplied, is probably the most satisfactory data collection procedure. This type of interview is somewhat different from other sorts of interviews and a brief summary of the principal factors involved will be given.

(i) *Sponsorship of the study.* If a stranger to the observers is collecting the data, it is ordinarily desirable to indicate on what authority the interview is being held. This part should be as brief as possible to avoid any use of time for a prolonged discussion of a topic irrelevant to the purpose of the interview. In many instances all that needs to be said is that someone known and respected by the observer has suggested the interview.

(ii) *Purpose of the study.* This should also be brief and ordinarily would merely involve a statement that a study was being made to describe the requirements of the activity. This would usually be cast in some such informal form as, "We wish to find out what makes a good citizen," or, "We are trying to learn in detail just what successful work as a nurse includes." In cases where there is some hesitation about cooperating or a little more explanation seems desirable, a statement can be added concerning the value and probable uses of the results. This frequently takes the form of improving selection and training procedures. In some instances, it would involve improving the results of the activity. For example, the interviewer might say, "In order to get better sales clerks we need to know just what they do that makes them especially effective or ineffective," or, "If parents are to be more effective, we need to be able to tell them the things they do that are effective and ineffective."

(iii) *The group being interviewed.* If there is any likelihood of a person feeling, "But, why ask me?" it is desirable to forestall this by pointing out that he is a member of a group which is in an unusually good position to observe and report on this activity. The special qualifications of members of this group as observers can be mentioned briefly, as, "Supervisors such as yourself are constantly observing and evaluating the work of switchboard girls," or, "Students are in an unusually good position to observe the effectiveness of their teachers in a number of ways."

(iv) *The anonymity of the data.* Especially for the collection of information about ineffective behavior, one of the principal problems is to convince the observer that his report cannot harm the person reported on in any way. Usually he also needs to be convinced that the person reported on will never know that he has reported the incident. Assurances are not nearly so effective in this situation as actual descriptions of techniques to be used in handling the

data, which enable the observer to judge for himself how well the anonymity of the data will be guarded. Under no circumstances should the confidences of the reportees be violated in any way. The use of sealed envelopes, avoidance of identifying information, the mailing of data immediately to a distant point for analysis, and similar techniques are helpful in establishing the good faith of the interviewer in taking all possible precautions to safeguard the incidents reported.

(v) *The question.* The most crucial aspect of the data collection procedure is the questions asked the observers. Many studies have shown that a slight change in wording may produce a substantial change in the incidents reported. For example, in one study the last part of one of the specific questions asked was, "Tell just how this employee behaved which caused a noticeable decrease in production." This question resulted in almost all incidents reported having to do with personality and attitude behaviors. This part of the question was changed to, "Tell just what this employee did which caused a noticeable decrease in production." This second question produced a much broader range of incidents. To the question writer "how he behaved" and "what he did" seemed like about the same thing. To the foremen who were reporting incidents "how he behaved" sounded as if personality and attitudes were being studied. The subtle biases involved in the wording of questions are not always so easily found. Questions should always be tried out with a small group of typical observers before being put into general use in a study.

The question should usually refer briefly to the general aim of the activity. This aim might be discussed more fully in a preliminary sentence. It should usually state that an incident, actual behavior, or what the person did is desired. It should briefly specify the type of behavior which is relevant and the level of importance which it must reach to be reported. It should also tie down the selection of the incidents to be reported by the observer in some way, such as asking for the most recent observation, in order to prevent the giving of only the more dramatic or vivid incidents, or some other selected group, such as those which fit the observer's stereotypes.

An effective procedure for insuring that the interpretation of the persons being interviewed is close to that intended is to request a sample of persons typical of those to be interviewed to state in their own words what they understand they have been asked to do. These persons should be selected so as to represent all types who will be interviewed. From a study of their interpretations, necessary revisions can be made to insure that all interviewees will be in agreement as to the nature of the incidents they are to provide.

(vi) *The conversation.* The interviewer should avoid asking leading questions after the main question has been stated. His remarks should be neutral and permissive and should show that he accepts the observer as the expert. By indicating that he understands what is being said and permitting the observer to do most of the talking, the interviewer can usually get unbiased incidents. If the question does not seem to be understood, it can be repeated with some reference to clarifying just what is meant by it. If the observer has given what seems like only part of the story, he should be encouraged by restating the essence of his remarks. This usually tends to encourage him to continue and may result in his bringing out many relevant details that the interviewer did not know the situation well enough to ask for. In some cases, it is desirable to have the interviews recorded electrically and transcribed. This increases the work load substantially, and trained interviewers can usually get satisfactory reports at the time or by editing their notes shortly after the interview.

Usually the interviewer should apply certain criteria to the incidents while they are being collected. Some of the more important criteria are: (a) is the actual behavior reported; (b) was it observed by the reporter; (c) were all relevant factors in the situation given; (d) has the observer made a definite judgment regarding the criticalness of the behavior; (e) has the observer made it clear just why he believes the behavior was critical.

In Fig. 3 is shown a sample of the type of form used by interviewers to collect critical incidents. Of course the form must be adapted to the needs of the specific situation.

b. Group interviews. Because of the cost in time and personnel of the individual interview, a group interview technique has been developed. This retains the advantages of the individual interview in regard to the personal contact, explanation, and availability of the interviewer to answer questions. To some extent it also provides for a check on the data supplied by the

interviewees. Its other advantages are that the language of the actual observer is precisely reproduced and the time for editing the interviews is virtually eliminated.

The method consists of having the interviewer give his introductory remarks to a group very much as he would do in an individual interview. There is an opportunity for questions and clarification. Then each person is asked to write incidents in answer to specific questions contained on a specially prepared form. The size of the group which can be handled effectively will vary with the situation. If the group is fairly small, it is usually possible for the interviewer to read the responses of each member of the group to the first question and make sure that he understands what is wanted. There seems to be a certain amount of social facilitation, and the results in most situations have been excellent. In the report of the first use of this procedure by Wagner (65), the amount of interviewer time required per usable incident was 4.3 minutes for the group interview procedure as compared with 15.7 minutes for individual interviews. The quality of these incidents, obtained from officers in the United States Air Force, appeared to be about the same for the two situations.

“Think of the last time you saw one of your subordinates do something that was very helpful to your group in meeting their production schedule.” (Pause till he indicates he has such an incident in mind.) “Did his action result in increase in production of as much as one per cent for that day?—or some similar period?”

(If the answer is “no,” say) “I wonder if you could think of the last time that someone did something that did have this much of an effect in increasing production.” (When he indicates he has such a situation in mind, say) “What were the general circumstances leading up to this incident?” _____

“Tell me exactly what this person did that was so helpful at that time.”

“Why was this so helpful in getting your group’s job done?” _____

“When did this incident happen?”

“What was this person’s job?” _____

“How long has he been on this job?”

“How old is he?” _____

FIG. 3. SAMPLE OF A FORM FOR USE BY AN INTERVIEWER IN COLLECTING EFFECTIVE CRITICAL INCIDENTS

c. Questionnaires. If the group becomes large, the group interview procedure is more in the nature of a questionnaire procedure. There are, of course, all types of combinations of procedures that can be used. The one that is most different from those discussed is the mailed questionnaire. In situations where the observers are motivated to read the instructions carefully and answer conscientiously, this technique seems to give results which are not essentially different from those obtained by the interview method. Except for the addition of introductory remarks, the forms used in collecting critical incidents by means of mailed questionnaires are about the same as those used in group interviews.

d. Record forms. One other procedure for collecting data is by means of written records. There are two varieties of recording: one is to record details of incidents as they happen. This situation is very similar to that described in connection with obtaining incidents by interviews above, except that the observation and giving of incidents are delayed following the introductory remarks and the presentation of the questions until an incident is observed to happen.

A variation of this procedure is to record such incidents on forms which describe most of the possible types of incidents by placing a check or tally in the appropriate place.

As additional information becomes available on the nature of the components which make up activities, observers may thus collect data more efficiently by using forms for recording and classifying observations. In the meantime, because of the inadequacy of the information currently available regarding these components, it seems desirable to ask observers to report their observations in greater detail and have the classification done by specially trained personnel.

Size of sample. A general problem which overlaps the phases of collecting the incidents and analyzing the data relates to the number of incidents required. There does not appear to be a simple answer to this question. If the activity or job being defined is relatively simple, it may be satisfactory to collect only 50 or 100 incidents. On the other hand, some types of complex activity appear to require several thousand incidents for an adequate statement of requirements.

The most useful procedure for determining whether or not additional incidents are needed is to keep a running count on the number of new critical behaviors added to the classification system with each additional 100 incidents. For most purposes, it can be considered that adequate coverage has been achieved when the addition of 100 critical incidents to the sample adds only two or three critical behaviors. For jobs of a supervisory nature, it appears that between 2,000 and 4,000 critical incidents are required to establish a comprehensive statement of requirements that includes nearly all of the different types of critical behaviors. For semiskilled and skilled jobs between 1,000 and 2,000 incidents seem to be adequate to cover the critical behaviors.

Coverage of all or nearly all of the various critical behaviors is not the only criterion as to whether or not a sufficient number of critical incidents has been collected. If a relatively precise definition of each critical behavior category is required, it may be necessary to get at least three or four examples of each critical behavior. Similarly, if the critical incidents are to be used as a basis for developing selection tests, training materials, and proficiency measures, more incidents may be required to provide a sufficient supply of usable ideas for the development of these materials.

In summary, although there is no simple formula for determining the number of critical incidents that will be required, this is a very important consideration in the plan of the study; checks should be made both on the first hundred or so incidents and again after approximately half of the number of incidents believed to be required have been obtained in order to make it possible to revise the preliminary estimates, if necessary, with a minimum loss in effort and time.

4. Analyzing the Data

The collection of a large sample of incidents that fulfill the various conditions outlined above provides a functional description of the activity in terms of specific

behaviors. If the sample is representative, the judges well qualified, the types of judgments appropriate and well defined, and the procedures for observing and reporting such that incidents are reported accurately, the stated requirements can be expected to be comprehensive, detailed, and valid in this form. There is only one reason for going further and that is practical utility. The purpose of the data analysis stage is to summarize and describe the data in an efficient manner so that it can be effectively used for many practical purposes.

In the discussion which follows, it should be kept in mind that the process of description has been completed. The specific procedures to be discussed are not concerned with improving on the comprehensiveness, specificity of detail, or validity of the statement of the requirements of the activity. Rather, they are concerned with making it easier to report these requirements, to draw inferences from them, and to compare the activity with other activities.

The aim is to increase the usefulness of the data while sacrificing as little as possible of their comprehensiveness, specificity, and validity. It appears that there are three primary problems involved: (a) the selection of the general frame of reference that will be most useful for describing the incidents; (b) the inductive development of a set of major area and subarea headings; and (c) the selection of one or more levels along the specificity-generality continuum to use in reporting the requirements. Each of these problems will be discussed below:

a. Frame of reference. There are countless ways in which a given set of incidents can be classified. In selecting the general nature of the classification, the principal consideration should usually be that of the uses to be made of the data. The preferred categories will be those believed to be most valuable in using the statement of requirements. Other considerations are ease and accuracy of classifying the data, relation to previously developed definitions or classification systems, and considerations of interpretation and reporting, which will be discussed in a later section.

For job activities, the choice of a frame of reference is usually dominated by considerations of whether the principal use of the requirements will be in relation to selection, training, measurement of proficiency, or the development of procedures for evaluating on-the-job effectiveness. For selection purposes, the most appropriate classification system is a psychological one. The main headings have to do with types of psychological traits that are utilized in the selection process. For training uses, the best classification system follows a set of headings that is easily related to training courses or broad training aims. For proficiency measurement, the headings tend to be similar to those for training except that there is less attention to possible course organization and aims and greater attention to the components of the job as it is actually performed. For the development of procedures for evaluating on-the-job effectiveness to establish a criterion of success, the classification system is necessarily directed at presenting the on-the-job behaviors under headings that represent either well-marked phases of the job or provide a simple framework for classifying on-the-job activities that is either familiar to or easily learned by supervisors.

Similarly, in nonvocational activities the frame of reference depends on the uses planned for the findings. For example, if a study is being made to define immaturity reactions in military personnel, the frame of reference would depend somewhat on whether the functional description is to be used primarily to identify personnel showing this type of maladjustment or whether the principal use will be to try to prepare specifications for types of situations in which immaturity reactions would not lead to serious difficulties.

b. Category formulation. The induction of categories from the basic data in the form of incidents is a task requiring insight, experience, and judgment. Unfortunately, this procedure is, in the present stage of psychological knowledge, more subjective than objective. No simple rules are available, and the quality and usability of the final product are largely dependent on the skill and sophistication of the formulator. One rule is to submit the tentative categories to

others for review. Although there is no guarantee that results agreed on by several workers will be more useful than those obtained from a single worker, the confirmation of judgments by a number of persons is usually reassuring. The usual procedure is to sort a relatively small sample of incidents into piles that are related to the frame of reference selected. After these tentative categories have been established, brief definitions of them are made, and additional incidents are classified into them. During this process, needs for redefinition and for the development of new categories are noted. The tentative categories are modified as indicated and the process continued until all the incidents have been classified.

The larger categories are subdivided into smaller groups and the incidents that describe very nearly the same type of behavior are placed together. The definitions for all the categories and major headings should then be re-examined in terms of the actual incidents classified under each.

c. General behaviors. The last step is to determine the most appropriate level of specificity-generality to use in reporting the data. This is the problem of weighing the advantages of the specificity achieved in specific incidents against the simplicity of a relatively small number of headings. The level chosen might be only a dozen very general behaviors or it might be several hundred rather specific behaviors. Practical considerations in the immediate situation usually determine the optimal level of generality to be used.

Several considerations should be kept in mind in establishing headings for major areas and in stating critical requirements at the selected level of generality. These are listed below:

(i) The headings and requirements should indicate a clear-cut and logical organization. They should have a discernible and easily remembered structure.

(ii) The titles should convey meanings in themselves without the necessity of detailed definition, explanation, or differentiation. This does not mean that they should not be defined and explained. It does mean that these titles, without the detailed explanation, should still be meaningful to the reader.

(iii) The list of statements should be homogeneous; i.e., the headings for either areas or requirements should be parallel in content and structure. Headings for major areas should be neutral, not defining either unsatisfactory or outstanding behaviors. Critical requirements should ordinarily be stated in positive terms.

(iv) The headings of a given type should all be of the same general magnitude or level of importance. Known biases in the data causing one area or one requirement to have a disproportionate number of incidents should not be reflected in the headings.

(v) The headings used for classification and reporting of the data should be such that findings in terms of them will be easily applied and maximally useful.

(vi) The list of headings should be comprehensive and cover all incidents having significant frequencies.

5. *Interpreting and Reporting*

It is never possible in practice to obtain an ideal solution for each of the practical problems involved in obtaining a functional description of an activity. Therefore, the statement of requirements as obtained needs interpretation if it is to be used properly. In many cases, the real errors are made not in the collection and analysis of the data but in the failure to interpret them properly. Each of the four preceding steps, (*a*) the determination of the general aim, (*b*) the specification of observers, groups to be observed, and observations to be made, (*c*) the data collection, and (*d*) the data analysis, must be studied to see what biases have been introduced by the procedures adopted. If there is a division of opinion as to the general aim and one of the competing aims is selected, this should be made very clear in the report. If the groups on whom the observations are made are not representative of the relevant groups involved, they must be described as precisely as possible. The aim of the study is usually not a functional description of the activity as carried on by this sample but rather a statement relating to all

groups of this type. In order to avoid faulty inferences and generalizations, the limitations imposed by the group must be brought into clear focus. Similarly, the nature of judgments made in collecting and analyzing the data must be carefully reviewed.

While the limitations need to be clearly reported, the value of the results should also be emphasized. Too often the research worker shirks his responsibility for rendering a judgment concerning the degree of credibility which should be attached to his findings. It is a difficult task, but if the results are to be used, someone will have to make such a judgment, and the original investigator is best prepared to make the necessary evaluations either for the general case or for certain typical specific examples.

USES OF THE CRITICAL INCIDENT TECHNIQUE

The variety of situations in which the collection of critical incidents will prove of value has only been partially explored. In the approximately eight years since the writer and his colleagues began a systematic formulation of principles and procedures to be followed in collecting this type of data, a fairly large number of applications has been made. The applications will be discussed under the following nine headings: (a) Measures of typical performance (criteria); (b) measures of proficiency (standard samples); (c) training; (d) selection and classification; (e) job design and purification; (f) operating procedures; (g) equipment design; (h) motivation and leadership (attitudes); (i) counseling and psychotherapy.

Space is not available here to describe these various applications in detail. However, a brief description of the types of application that have been made, along with brief illustrative examples and references, will be presented. Some of the studies involve several of the types of applications to be discussed. The presentation is not intended to be complete, but rather to give the reader interested in further study some orientation and guidance.

Measures of typical performance (criteria). The simplest and most natural application of a systematically collected set of critical incidents is in terms of the preparation of a statement of critical requirements and a check list or some similar type of procedure for evaluating the typical performance of persons engaged in this activity. If an observational check list that includes all of the important behaviors for the activity is available, the performance of the individual can be objectively evaluated and recorded by merely making a single tally mark for each observation. Such records provide the essential basis for criterion data which are sufficiently detailed and specific for special purposes but at the same time can be combined into a single over-all evaluation when this is desirable. Such a procedure was first suggested and tried out in connection with developmental studies of the American Institute for Research. These included: Preston's study of officers for the United States Air Force (52); Nagay's study on air route traffic controllers for the Civil Aeronautics Administration (49); and M. H. Weislogel's study on research personnel for the Office of Naval Research (69). Another American Institute for Research study was reported by R. B. Miller and the present author (21). This was a performance record form for hourly wage employees developed in cooperation with personnel of the Delco-Remy Division of the General Motors Corporation, the Employment Practices Division of that corporation, and the Industrial Relations Center of the University of Chicago. The same authors have developed similar

performance records for salaried employees, and foremen and supervisors (22, 23). The principles and procedures underlying this type of evaluation of performance have been published elsewhere (14, 15, 17).

A number of important contributions to the development of functional descriptions and standards of performance have been made by other groups using the critical incident technique. One of the most notable of these is the development by Hobbs et al. (3, 31), of Ethical Standards of Psychologists. More than 1,000 critical incidents involving ethical problems of psychologists were contributed by the members of the American Psychological Association. It is believed that this represents the first attempt to use empirical methods to establish ethical standards. Because of the importance of this study, and the generality of some of the problems involved, certain of the conclusions reported by the Committee on Ethical Standards for Psychology in their introductory statement will be quoted here.

First, it is clear that psychologists believe that ethics are important; over two thousand psychologists were sufficiently concerned with the ethical obligations of their profession to contribute substantially to the formulation of these ethical standards. Second, psychologists believe that the ethics of a profession cannot be prescribed by a committee; ethical standards must emerge from the day-by-day value commitments made by psychologists in the practice of their profession. Third, psychologists share a conviction that the problems of men, even those involving values, can be studied objectively; this document summarizes the results of an effort to apply some of the techniques of social science to the study of ethical behavior of psychologists. Fourth, psychologists are aware that a good code of ethics must be more than a description of the current status of ethics in the profession; a code must embody the ethical aspirations of psychologists and encourage changes in behavior, bringing performance ever closer to aspiration. Fifth, psychologists appreciate that process is often more important than product in influencing human behavior; the four years of widely-shared work in developing this code are counted on to be more influential in changing ethical practices of psychologists than will be the publication of this product of their work. Finally, psychologists recognize that the process of studying ethical standards must be a continuing one; occasional publications such as this statement mark no point of conclusion in the ongoing process of defining ethical standards—they are a means of sharing the more essential discipline of examining professional experience, forming hypotheses about professional conduct, and testing these hypotheses by reference to the welfare of the people affected by them (3, p. v).

In addition to the study by Smit mentioned in a previous section (58), several other studies on the use of the critical incident procedures as a basis for evaluating teaching effectiveness have been reported. One of these was a study conducted under the joint sponsorship of the Educational Research Corporation and the Harvard University Graduate School of Education with funds provided by the New England School Development Council and the George F. Milton Fund. This was an exploratory study of teacher competence reported by Domas (6). Approximately 1,000 critical incidents were collected from teachers, principals, and other supervisors. Although this was an exploratory study, it was felt that it made an important contribution to the general problem of relating salary to teacher competence.

The second of these studies was conducted as part of the teacher characteristics study sponsored by the American Council on Education and subsidized by the Grant Foundation. This study is reported by Jensen (32). Teachers, administrators, and teachers in training in the Los Angeles area contributed more than 1,500 critical incidents of teacher behavior. The incidents were classified under personal, professional, and social qualities. The category formulation indicated that there were about 20 distinct critical requirements.

These were recommended as a basis for teacher evaluation and as an aid to the in-service growth of teachers.

Another study was that of Smith and Staudohar (59), which determined the critical requirements for basic training of tactical instructors in the United States Air Force. From 130 training supervisors, 555 tactical instructors, and 3,082 basic trainees, a total of 6,615 usable incidents were obtained. The authors comment that:

The training supervisors report a predominance of ineffective incidents in the major areas of: Sets a good example and maintains effective personal relations. The tactical instructors report more ineffective incidents in the area of Makes his expectations clear. Basic trainees show a predominance of ineffective incidents in three areas: Sets a good example, Considers trainee's needs, and Maintains effective personal relations (59, p. 5).

Another study on the evaluation of instructor effectiveness was carried out by Konigsburg (33). This study involved the development of an instructor check list for college instructors based on the critical incident technique and a comparison of techniques for recording observations. Its principal findings were the very low correlation coefficients between the total scores from the Purdue Rating Scale for Instruction and the instructor check list. When these two instruments were each given to half the class on the same day, the average correlation coefficient was found to be .29. The other principal finding is that the planned performances of a total of 46 predetermined behaviors were better reflected by the results obtained on the instructor check list than by the results on the Purdue Rating Scale.

A somewhat related study has been reported by Barnhart (4). This study collected a large number of critical incidents for the purpose of establishing the critical requirements for school board membership. The author applied his findings to the problem of evaluating the effectiveness of school board members.

Another type of application of the critical incident technique to the development of bases for evaluating behavior is the previously mentioned study of Eilbert (7). His list of 51 types of immature reaction based on a collection of several hundred critical incidents describing manifestations of emotional immaturity is believed to provide a useful guide to further investigation and appraisal of persons with behavior problems. It is believed that the results of this study provide substantial encouragement to the application of the critical incident technique to similar problems in the field of clinical diagnosis and evaluation.

Measures of proficiency (standard samples). A closely related use of critical incidents is to provide a basis for evaluating the performance of persons by use of standard samples of behavior involving important aspects of the activity. Such evaluations are called proficiency measures and are differentiated from the evaluation of typical performance on the job primarily on the basis that a test situation rather than a real job situation is used. Measures of this sort are especially useful at the end of training courses as checks on the maintenance of proficiency, and when the tasks assigned to participants vary a great deal in difficulty or are not directly observed by the supervisors.

One of the first applications of critical incidents to the development of proficiency measures was Gordon's study on the development of a standard flight check for the airline transport rating (28, 29). This study was done by the American Institute for Research under the sponsorship of the National Research Council Committee on Aviation

Psychology with funds provided by the Civil Aeronautics Administration. In this study data from analyses of airline accidents were combined with critical incidents reported by airline pilots to provide the basis for developing an objective measure of pilot proficiency. The flight check consisted of the presentation of situations providing uniformly standardized opportunities to perform the critical aspects of the airline pilot's job as indicated from the study of the accidents and critical incidents reported. The new check was found to yield 88 per cent agreement on the decision to pass or fail a particular pilot when examined on flights on successive days by different check pilots. The previous flight check when used on the same flights gave only 63 per cent agreement, which was little better than chance under the conditions of the study.

Similar studies on the development of flight checks at the American Institute for Research have been carried out by Marley (36, 37), G. S. Miller (39), and Ericksen (9). These studies, sponsored by the United States Air Force and the Civil Aeronautics Administration, were concerned respectively with objective flight checks for B-29 bombing crew members, B-36 bombing crew members, and private pilots flying light civilian aircraft. Ericksen also developed a light plane proficiency check to predict military flying success (10) on a similar project sponsored by the United States Air Force Human Resources Research Center.

A similar set of proficiency measures was developed by Krumm for Air Force pilot instructors (34, 35), also under the sponsorship of the Human Resources Research Center. These measures were based on more than 4,000 critical incidents collected from student pilots, flight instructors, and supervisors. The critical incidents were classified under three main headings: (a) proficiency as a pilot; (b) proficiency as a teacher; and (c) proficiency in maintaining effective personnel relations. The proficiency measures developed in connection with this study included paper-and-pencil tests presenting critical situations and requiring the instructor to select one of several proposed solutions.

Another development of this type carried on at the American Institute for Research was the construction of tests for evaluating research proficiency in physics and chemistry for the Office of Naval Research by M. H. Weislogel (71). This study was based on the critical incidents for research personnel (20) discussed in a previous section. The items for these proficiency measures were based on detailed rationales. The items described a practical research situation in considerable detail and outlined five specific choices concerning such matters as the best thing to do next, suggestions for improving the procedure as reported, etc. The critical behaviors tested in the items were taken directly from the critical incidents. The method of developing tests through the use of comprehensive rationales has been discussed generally in another paper (16).

Three studies have been reported by the American Institute for Research in which critical incidents were used as a basis for developing situational performance tests for measuring certain aspects of the proficiency of military personnel. These included the study of Sivy and Lange on the development of an objective form of the Leaders Reaction Test for the Personnel Research Branch, Department of the Army (57). This test included 20 situational problems based on the critical requirements of the noncommissioned combat infantry leader as determined on the basis of critical incidents collected in military maneuvers and during combat operations at the front in Korea. A second proficiency measure of a somewhat similar sort was developed for other types of personnel by R. L. Weislogel (73). The third study of this type was carried out by Suttell

(61) for the Human Resources Research Center. This study was based on critical incidents collected in previous studies of the American Institute for Research and reported the development and preliminary evaluation of the Officer Situations Test. This test was designed to measure nonintellectual aspects of officer performance through the use of 16 situational problems requiring about six hours of testing time.

Because of the great difficulty in obtaining valid and reliable measures of typical performance, accurate measures of proficiency are essential for many types of activities. It is apparent that a comprehensive set of critical incidents can be of great value in constructing such measures.

Training. Many of the applications of the critical incident technique to training problems have been carried out for the military in special situations so that the reports are classified security information. In addition to work by Preston, Glaser, and R. L. Weislogel, R. B. Miller and Folley have utilized critical incidents in establishing training requirements for specific types of maintenance mechanics (47) in a study for the Human Resources Research Center.

Similarly, Ronan has used critical incidents as a basis for developing a program of training for emergency procedures in multi-engine aircraft (54) in a study for the United States Air Force Human Factors Operations Research Laboratory. On the basis of several thousand incidents reported by aircrew personnel regarding emergencies, three evaluation devices were prepared. These involved a conventional type multiple-choice test; a special multiple-choice test designed to measure the individual's information concerning the important cues in the emergency situation, the appropriate actions to be taken, and the basic troubles or causes of the emergency; and a "flight check" to be used in evaluating the performance of aircrew members in electronic flight simulators.

The obvious relevance of the behaviors involved in critical incidents and the specific details included make such incidents an ideal basis for developing training programs and training materials.

A recent study by Collins (5) uses critical incidents as a basis for evaluating the effectiveness of a training program. The types of incidents reported by mothers after a two-week training course were significantly different from those reported at the beginning of the program in a number of aspects relevant to the objectives of the program. The critical incidents appeared to provide a much more sensitive basis for revealing changes than other procedures used.

Selection and classification. Until recently, the customary approach of the research psychologist to the development of tests for selection and classification purposes has been as follows: A very brief period was given to study of the job. Following this, a wide variety of selection procedures was administered to a group of applicants or employees, and follow-up data were gathered. Since the research psychologist had little confidence in the accuracy of his analysis of the psychological elements required by the job, there was a tendency to try everything that was available and seemed even remotely related to the tasks involved. This has been called the "shotgun approach." It was hoped that with a wide scatter at least a few of the tests would pay off. The critical incident technique has lent substantial support to the more thorough study of the job prior to initiating testing procedures. There is increasing feeling at the present time that a much larger percentage of the investigator's time should be spent on determining the critical requirements of the job, so that the psychologist will have sufficient confidence in his tentative conclusions as

to the nature of the important selection procedures to permit their use on a tentative basis prior to the collection of empirical follow-up data. This is especially important in those situations where the follow-up requires a very long period of time or where the number of cases that can be followed up is so small that definitive findings cannot be anticipated.

One of the most important requirements for developing a system of job analysis that will facilitate a relatively accurate identification of the important job elements for a specific task is to establish a clear and specific set of definitions for these job elements in behavioral terms. The American Institute for Research has carried out a series of projects on this problem. The first of these was a study undertaken by Wagner under the sponsorship of the United States Air Force School of Aviation Medicine to define the requirements of aircrew jobs in terms of specific job elements (67, 68). Several thousand critical incidents were gathered from aircrew members, and these were classified into 24 job elements. These job elements were inductively formulated from the critical incidents and were grouped under the four area headings: (a) learning and thinking; (b) observation and visualization; sensory-motor coordination; and motives, temperament, and leadership.

The development of more than 100 proficiency tests to measure each of the various critical behaviors included in the 24 tentatively proposed job elements was reported by Hahn (30) for the School of Aviation Medicine. These tests were administered to a group of approximately 500 high school senior boys, and the intercorrelations were used to reformulate the tentative job elements. In a study just completed by Taylor (62) for the Human Resources Research Center, the results of applying an analytical procedure developed by Horst to study the interrelationships involved are reported. This analysis led to the formulation of a new set of 20 job elements for each of which a selection test has been developed. These tests have been administered to several hundred aviation cadets and follow-up data on their success in aircrew training should be available soon.

A similar project based on critical incidents collected from various civilian jobs has been reported by the present author (2, 18, 19). The Flanagan Aptitude Classification Test Series, published in 1953, provides aptitude measures for 14 critical job elements. The Applicant Inventory, also published in 1953, measures attitudes predictive of job adjustment for hourly wage employees.

An effort to adapt the critical incident technique to the problem of developing civil service examinations is reported by Wager and Sharon (64). In an exploratory study they collected about 100 incidents regarding on-the-job behaviors of maintenance technicians. These incidents were used as a basis for determining job requirements in terms of behavior, and test items were developed for use in selecting applicants who could be expected to meet these requirements.

Another study that used critical incidents as a basis for developing tests to predict performance was carried out by O'Donnell (51). His test, designed to predict success in dentistry, was based on critical incidents collected by Wagner. The test includes items designed to predict, in part, the following three general areas: (a) demonstrating technical proficiency; (b) handling patient relationships; and (c) accepting professional responsibility. A follow-up study indicated moderate validity for these materials.

One of the few studies known to the author in which the critical incident technique was used in a project carried on outside the United States is Emons' doctor's dissertation (8). This study, carried out at the University of Liege, investigated the aptitudes of

effective sales personnel in a large department store. A group of 40 supervisors provided 228 critical incidents. Nine categories were formulated from this group of incidents and recommendations made for an aptitude test to improve current selection procedures.

5. Job design and purification. Inadequate attention has been given to the scientific design of jobs to promote over-all efficiency. Where a team has several different types of tasks to perform, it is frequently possible to design each of the team member's jobs so that only a few of the several tasks are involved. If the jobs have been studied by use of the critical incident technique, it may be possible to select and train each team member for only two or three of the critical job elements. This tends to maximize the effectiveness of performance with respect to each of the various types of tasks. Although such procedures have nearly always been informally used in planning the work of teams, the critical incident technique facilitates the collection of the data essential to this type of job purification.

Some preliminary work on this problem has been carried out at the American Institute for Research. Recommendations resulting from these studies for reducing the number of job elements required in certain common maintenance jobs are expected to lead to a saving of millions of dollars in training costs as well as to improving the effectiveness of job performance.

Operating procedures. Another application of critical incidents which has not been adequately exploited is the study of operating procedures. Detailed factual data on successes and failures that can be systematically analyzed are of great importance in improving the effectiveness and efficiency of operations. Such information can be efficiently collected by means of the critical incident technique.

Examples of such studies are provided by a series of three projects carried out by the American Institute for Research under the sponsorship of the United States Air Force School of Aviation Medicine. The first of these involves the collection of critical incidents relating to near accidents in flying reported by Vasilas, Fitzpatrick, DuBois, and Youtz (63). More than 1,700 critical incidents were collected from pilots and other aircrew members by procedures developed for this study. These incidents pointed to possible improvements in training job design and equipment design as well as in operating procedures.

The second of these studies was specifically concerned with the effect of the age of pilots and other crew members on aircrew operations. This study was reported by Shriver (56), and included tentative suggestions regarding various modifications in operating procedures.

The third study in this series, reported by Goodenough and Suttell (26), involved the collection of critical incidents regarding the impairment of human efficiency in emergency operations. These incidents provide a detailed statement of both the types of stresses that impair performance and the types of performance that are impaired under specific conditions. More than 2,000 critical incidents were collected in which impairment in performance on operational assignments was observed. These incidents were collected in Alaska and the Far East as well as in operational commands in the United States. This report contains suggestions for improving operations in emergency situations.

Equipment design. An application closely related to that just discussed involves the collection of critical incidents to improve the design of equipment. Reports of specific in-

cidents from the field have always been a basis for equipment modifications. The critical incident technique facilitates the collection and processing of this type of information. Too often in the past action was taken on the basis of informal reports from operating personnel. The collection of large numbers of critical incidents representative of operating experience provides a sound basis for modifying existing equipment and designing new models.

In the study by Fitts and Jones (12), mentioned above, which was carried out at the Aero-Medical Laboratory, 270 critical incidents relating to errors in reading and interpreting aircraft instruments were collected and analyzed. These led to a number of specific suggestions regarding modifications in instrument displays.

Other recent studies conducted at the American Institute for Research have used data from the critical incident technique along with other sources to develop procedures for designing jobs. The reports on these projects are classified for military security reasons.

Other projects at the American Institute for Research have used the critical incident technique as a supplemental procedure for task analysis of equipment in the design stage of development (9, 10, 34, 35, 39). These procedures have been found very effective when used by psychologists working closely with engineers on the preparation of design specifications for new equipment.

Motivation and leadership. The study of attitudes has been somewhat limited and difficult to interpret because of the almost exclusive reliance on verbal statements of opinions and preferences. The critical incident technique has been applied in a few instances to gather factual data regarding specific actions involving decisions and choices. These studies suggest that critical incidents of this type may be a very valuable supplementary tool for the study of attitudes.

A recent study carried out by Preston of the American Institute for Research for the Air Force's Human Resources Research Center (53) used critical incidents as a basis for studying decisions of airmen to re-enlist in the Air Force. It is believed that these specific incidents provide valuable information not contained in studies utilizing only data on opinions.

A series of reports by Ruch (55) contains critical incidents on combat leadership collected from senior officers in the Far East Air Forces. These incidents provide a factual basis for the study of motivation and leadership of Air Force personnel engaged in combat operations.

Counseling and psychotherapy. Another field in which current techniques emphasize over-all impressions, opinions, and reports of single cases is counseling and psychotherapy. There appears to be a trend, however, in this field toward emphasizing the collection of factual incidents. This suggests that the critical incident technique may be useful in this area also.

Exploratory work has recently been done at the University of Pittsburgh with the critical incident technique to establish areas of change accompanying psychotherapy. A series of three master's theses were carried out by Speth, Goldfarb, and Mellett (25, 38, 60). They collected 243 critical incidents from 11 psychotherapists. These incidents were collected about patients who had shown improvement and were replies to the question, "What did the patient do that was indicative of improvement?" Although these studies were primarily exploratory in nature, the tentative finding that different therapists stress different criteria of improvement and nonimprovement suggests that the critical incident

approach may be of use not only in developing objective measures of improvement but also in experimental studies of the types of improvement resulting from the therapists' use of specific procedures.

A somewhat related type of study initiated by Diederich and reported by Allen (1) describes the use of the technique to obtain critical incidents from students reporting things that caused them to like a fellow high school student either more or less than before. This study is being continued to provide the basis for tests of specific value areas. An incidental finding of the study was that when an example of the kind of incident desired was shown on the form, 53 per cent of the positive and 23 per cent of the negative behaviors reported were in the same category as the example given.

SUMMARY AND CONCLUSIONS

This review has described the development of a method of studying activity requirements called the critical incident technique. The technique grew out of studies carried out in the Aviation Psychology Program of the Army Air Forces in World War II. The success of the method in analyzing such activities as combat leadership and disorientation in pilots resulted in its extension and further development after the war. This developmental work has been carried out primarily at the American Institute for Research and the University of Pittsburgh. The reports of this work are reviewed briefly.

The five steps included in the critical incident procedure as most commonly used at the present time are discussed. These are as follows: (a) Determination of the general aim of the activity. This general aim should be a brief statement obtained from the authorities in the field which expresses in simple terms those objectives to which most people would agree. (b) Development of plans and specifications for collecting factual incidents regarding the activity. The instructions to the persons who are to report their observations need to be as specific as possible with respect to the standards to be used in evaluating and classifying the behavior observed, (c) Collection of the data. The incident may be reported in an interview or written up by the observer himself. In either case it is essential that the reporting be objective and include all relevant details. (d) Analysis of the data. The purpose of this analysis is to summarize and describe the data in an efficient manner so that it can be effectively used for various practical purposes. It is not usually possible to obtain as much objectivity in this step as in the preceding one. (e) Interpretation and reporting of the statement of the requirements of the activity. The possible biases and implications of decisions and procedures made in each of the four previous steps should be clearly reported. The research worker is responsible for pointing out not only the limitations but also the degree of credibility and the value of the final results obtained. It should be noted that the critical incident technique is very flexible and the principles underlying it have many types of applications. Its two basic principles may be summarized as follows: (a) reporting of facts regarding behavior is preferable to the collection of interpretations, ratings, and opinions based on general impressions; (b) reporting should be limited to those behaviors which, according to competent observers, make a significant contribution to the activity.

It should be emphasized that critical incidents represent only raw data and do not automatically provide solutions to problems. However, a procedure which assists in collecting representative samples of data that are directly relevant to important problems

such as establishing standards, determining requirements, or evaluating results should have wide applicability.

The applications of the critical incident technique which have been made to date are discussed under the following nine headings: (a) Measures of typical performance (criteria); (b) measures of proficiency (standard samples); (c) training; (d) selection and classification; (e) job design and purification; (f) operating procedures ; (g) equipment design; (h) motivation and leadership (attitudes); (i) counseling and psychotherapy.

In summary, the critical incident technique, rather than collecting opinions, hunches, and estimates, obtains a record of specific behaviors from those in the best position to make the necessary observations and evaluations. The collection and tabulation of these observations make it possible to formulate the critical requirements of an activity. A list of critical behaviors provides a sound basis for making inferences as to requirements in terms of aptitudes, training, and other characteristics. It is believed that progress has been made in the development of procedures for determining activity requirements with objectivity and precision in terms of well-defined and general psychological categories. Much remains to be done. It is hoped that the critical incident technique and related developments will provide a stable foundation for procedures in many areas of psychology.

REFERENCES

1. ALLEN, CD. Critical requirements in interpersonal behavior. Unpublished senior thesis, Princeton Univer., 1950.
2. AMERICAN INSTITUTE FOR RESEARCH. *The applicant inventory*. Pittsburgh: American Institute for Research, 1954.
3. AMERICAN PSYCHOLOGICAL ASSOCIATION. COMMITTEE ON ETHICAL STANDARDS FOR PSYCHOLOGY. *Ethical standards of psychologists*. Washington, D. C: Amer. Psychol. Ass., 1953.
4. BARNHART, R. E. The critical requirements for school board membership based upon an analysis of critical incidents. Unpublished doctor's dissertation, Indiana Univer., 1952.
5. COLLINS, MARJORIE G. Selected methods applied to the evaluation of a parent training course in child management. Unpublished doctor's dissertation, Univer. of Pittsburgh, 1954.
6. DOMAS, S. J. *Report of an exploratory study of teacher competence*. Cambridge: New England School Development Council, 1950.
7. EILBERT, L. R. A study of emotional immaturity utilizing the critical incident technique. *Univer. Pittsburgh Bull.*, 1953, 49, 199-204.
8. EMONS, V. L'analyse de la fonction de vendeuse de grand magasin par la methode des exigences critiques. Unpublished doctor's dissertation, Univer. of Liege, Belgium, 1952.
9. ERICKSEN, S. C. Development of an objective proficiency check for private pilot certification. Washington: Civil Aeronautics Administration, 1951. (*Div. of Res., Rep. No. 95.*)
10. ERICKSEN, S. C. Development of a light plane proficiency check to predict military flying success. *USAF, Hum. Re-sour. Res. Cent. Tech. Rep.*, 1952, No. 52-6.
11. FINKLE, R. B. A study of the critical requirements of foremanship. *Univer. Pittsburgh Bull.*, 1950,46,291-297. (Abstract)
12. FITTS, P. M., & JONES, R. E. Psychological aspects of instrument display. 1. Analysis of 270 "pilot error" experiences in reading and interpreting aircraft instruments. Dayton, O.: U. S. Air Force, Air Materiel Command, Wright-Patterson Air Force Base, 1947. (*Mem. Rep. TSEAA-694-12A.*)
13. FLANAGAN, J. C. The aviation psychology program in the Army Air Forces. Washington: U. S. Government Printing Office, 1947. (*AAF Aviat. Psychol. Program Res. Rep. No. 1.*)

14. FLANAGAN, J. C. A new approach to evaluating personnel. *Personnel*, 1949, 26, 35-42.
15. FLANAGAN, J. C. Critical requirements: a new approach to employee evaluation. *Personnel Psychol.*, 1949, 2, 419-425.
16. FLANAGAN, J. C. Use of comprehensive rationales. *Educ. psychol. Measmt*, 1951, 11, 151-155.
17. FLANAGAN, J. C. Principles and procedures in evaluating performance. *Personnel*, 1952, 28, 373-386.
18. FLANAGAN, J. C. *Flanagan aptitude classification tests*. Chicago: Science Research Associates, 1953.
19. FLANAGAN, J. C. Improving personnel selection. *Publ. Personnel Rev.*, 1953, 14, 107-112.
20. FLANAGAN, J. C, et al. *Critical requirements for research personnel*. Pittsburgh : American Institute for Research, 1949.
21. FLANAGAN, J. C, MILLER, R. B., BURNS, R. K., HENDRIX, A. A., STEWART, B., PRESTON, H. O., & WEST, E. D. *The performance record for hourly employes*. Chicago: Science Research Associates, 1953.
22. FLANAGAN, J. C, MILLER, R. B., BURNS, R. K., HENDRIX, A. A., STEWART, B., PRESTON, H. O., & WEST, E. D. *The performance record for non-supervisory salaried employes*. Chicago: Science Research Associates, 1953.
23. FLANAGAN, J. C, MILLER, R. B., BURNS, R. K., HENDRIX, A. A., STEWART, B., PRESTON, H. O., & WEST, E. D. *The performance record for foremen and supervisors*. Chicago: Science Research Associates, 1953.
24. FOLLEY, J. D., JR. Development of a list of critical requirements for retail sales personnel from the standpoint of customer satisfaction. Unpublished master's thesis, Univer. of Pittsburgh, 1953.
25. GOLDFARB, A. Use of the critical incident technique to establish areas of change accompanying psychotherapy: II. Relationship to diagnostic group. Unpublished master's thesis, Univer. of Pittsburgh, 1952.
26. GOODENOUGH, D. R., & SUTTELL, BARBARA J. *The nature and extent of impairment of human efficiency in emergency operations: the field study*. Pittsburgh: American Institute for Research, 1952.
27. GORDON, T. The airline pilot: a survey of the critical requirements of his job and of pilot evaluation and selection procedures. Washington: Civil Aeronautics Administration, 1947. (*Div. of Res., Rep. No. 73.*)
28. GORDON, T. The development of a standard flight-check for the airline transport rating based on the critical requirements of the airline pilot's job. Washington: Civil Aeronautics Administration, 1949. (*Div. of Res., Rep. No. 85.*)
29. GORDON, T. The development of a method of evaluating flying skill. *Personnel PsychoL*, 1950, 3, 71-84.
30. HAHN, C. P. Measurement of individual differences with respect to critical job requirements. *USAF Sell. Aviat. Med.*, 1954, Proj. No. 21-29-014, No. 2.
31. HOBBS, N. The development of a code of ethical standards for psychology. *Amer. Psychologist*, 1948, 3, 80-84.
32. JENSEN, A. C. Determining critical requirements for teachers. *. exp. Educ*, 1951, 20, 79-86.
33. KONIGSBURG, D. The development and preliminary evaluation of an instructor check list based on the critical incident technique. Unpublished doctor's dissertation, Univer. of Pittsburgh, 1954.
34. KRUMM, R. L. Critical requirements of pilot instructors. *USAF, Hum. Resour. Res. Cent., Tech. Rep.*, 1952, No. 52-1.
35. KRUMM, R. L. *Development of a measure of pilot instructor proficiency based on the critical requirements of the instructor's job*. Pittsburgh: American Institute for Research, 1953.
36. MARLEY, F. W. Individual differences in critical aircrew elements. I. The determination of critical proficiency requirements for B-29 combat crews. *USAF Sch. Aviat. Med.*, 1952, Proj. No. 21-29-014, No. 1.

37. MARLEY, F. W. The development of performance flight checks for B-29 combat crews. *USA F, Hum. Resour. Res. Lab.*, 1952, Rep. No. 19.
38. MELLETT, T. P. Use of the critical incident technique to establish areas of change accompanying psychotherapy: III. Differences among therapists. Unpublished master's thesis, Univer. of Pittsburgh, 1952.
39. MILLER, G. S. *Development of B-36 bombardment objective flight checks and proficiency measures for use with reconnaissance crews.* Pittsburgh: American Institute for Research, 1953.
40. MILLER, N. E. Psychological research on pilot training. Washington: U. S. Government Printing Office, 1947. (*AAF Aviat. Psychol. Program. Res. Rep.* No. 8.)
41. MILLER, R. B. Anticipating tomorrow's maintenance job. *USA F, Hum. Resour. Res. Cent., Res. Rev.*, 1953, No. 53-1.
42. MILLER, R. B. A method for determining human engineering design requirements for training equipment. Dayton, O.: U. S. Air Force, Air Development Center, Wright-Patterson Air Force Base, 1953. (*Tech. Rep.* 53-135.)
43. MILLER, R. B. Handbook on training and training equipment design. Dayton, O.: U. S. Air Force, Air Development Center, Wright-Patterson Air Force Base, 1953. (*Tech. Rep.* 53-136.)
44. MILLER, R. B. A method for man-machine task analysis. Dayton, O.: U. S. Air Force, Air Development Center, Wright-Patterson Air Force Base, 1953. (*Tech. Rep.* 53-137.)
45. MILLER, R. B. Human engineering design schedule for training equipment. Dayton, O.: U. S. Air Force, Air Development Center, Wright-Patterson Air Force Base, 1953. (*Tech. Rep.* 53-138.)
46. MILLER, R. B., & FLANAGAN, J. C. The performance record: an objective merit-rating procedure for industry. *Amer. Psychologist*, 1950, 5, 331-332. (Abstract)
47. MILLER, R. B., & FOLLEY, J. D., JR. *The validity of maintenance job analysis from the prototype of an electronic equipment.* Pittsburgh: American Institute for Research, 1952.
48. NAGAY, J. A. The development of a procedure for evaluating the proficiency of air route traffic controllers. Washington: Civil Aeronautics Administration, 1949. (*Div. of Res., Rep.* No. 83.)
49. NAGAY, J. A. The airline tryout of the standard flight-check for the airline transport rating. Washington: Civil Aeronautics Administration, 1949. (*Div. of Res., Rep.* No. 88.)
50. NEVINS, CHARLOTTE I. An analysis of reasons for the success or failure of bookkeepers in sales companies. Unpublished master's thesis, Univer. of Pittsburgh, 1949.
51. O'DONNELL, R. J. The development and evaluation of a test for predicting dental student performance. *Univer. Pittsburgh Bull.*, 1953, 49, 240-245. (Abstract)
52. PRESTON, H. O. *The development of a procedure for evaluating officers in the United States Air Force.* Pittsburgh: American Institute for Research, 1948.
53. PRESTON, H. O. *Events affecting reenlistment decisions.* Pittsburgh: American Institute for Research, 1953.
54. RONAN, W. W. *Training for emergency procedures in multi-engine aircraft.* Pittsburgh: American Institute for Research, 1953.
55. RUCH, F. L. Incidents of leadership in combat. *USAF, Hum. Resour. Res. Inst., Res. Memo.*, 1953, No. 3.
56. SHRIVER, BEATRICE M. Age and behavior: a study of the effects of aging on aircrew performance. *USAF Sch. Aviat. Med.*, 1953, Proj. No. 21-0202-0005, No. 3.
57. SIVY, J., & LANGE, C. J. Development of an objective form of the Leaders Reaction Test. *AGO, Personnel Res. Branch, PRS Rep.* 1952, No. 930.
58. SMIT, JO ANNE. A study of the critical re-requirements for instructors of general psychology courses. *Univer. Pittsburgh Bull.*, 1952, 48, 279-284. (Abstract)
59. SMITH, R. G., JR., & STAUDOCHAR, F. T. Technical requirements of basic training tactical instructors. *USAF, Hum. Resour. Res. Cent., Tech. Rep.*, 1954, No. 54-7.

60. SPETH, E. W. The use of the critical incident technique to establish areas of change accompanying psychotherapy: I. Function of age and education. Unpublished master's thesis, Univer. of Pittsburgh, 1952.
61. SUTTELL, BARBARA J. *Development of a situation test to measure non-intellectual officer qualities*. Pittsburgh: American Institute for Research, 1953.
62. TAYLOR, M. V., JR. *The development of aircrew job element aptitude tests*. Pittsburgh: American Institute for Research, 1953.
63. VASILAS, J. N., FITZPATRICK, R., DUBOIS, P. H., & YOUTZ, R. P. Human factors in near accidents. *USAF Sch. Aviat. Med.*, 1953, Proj. No. 21-1207-0001, No. 1.
64. WAGER, C. E., & SHARON, M. I. Defining job requirements in terms of behavior. *Personnel Admin.*, 1951, 14, 18-25.
65. WAGNER, R. F. A group situation compared with individual interviews for securing personnel information. *Personnel Psychol.*, 1948, 1, 93-107.
66. WAGNER, R. F. A study of the critical re-requirements for dentists. *Univer. Pittsburgh Bull.*, 1950, 46, 331-339. (Abstract)
67. WAGNER, R. F. Development of standardized procedures for defining the requirements of aircrew jobs in terms of testable traits. *USAF Sch. Aviat. Med.*, 1951, Proj. No. 21-29-010, No. 1.
68. WAGNER, R. F. Using critical incidents to determine selection test weights. *Personnel Psychol.*, 1951, 4, 373-381.
69. WEISLOGEL, MARY H. *Development of a test for selecting research personnel*. Pittsburgh: American Institute for Research, 1950.
70. WEISLOGEL, MARY H. *Procedures for evaluating research personnel with a performance record of critical incidents*. Pittsburgh: American Institute for Research, 1950.
71. WEISLOGEL, MARY H. *The development of tests for evaluating research proficiency in physics and chemistry*. Pittsburgh: American Institute for Research, 1951.
72. WEISLOGEL, R. L. Critical requirements for life insurance agency heads. *Univer. Pittsburgh Bull*, 1952, 48, 300-305. (Abstract)
73. WEISLOGEL, R. L., & SCHWARZ, P. A. Some practical and theoretical problems in situational testing. Unpublished manuscript, Pittsburgh, American Institute for Research, 1953.
74. WICKERT, F. Psychological research on problems of redistribution. Washington: U. S. Government Printing Office, 1947. (*AAF Aviat. Psychol. Program Res. Rep. No. 14.*)

Received November 10, 1953.