A MANUAL OF TEST VALIDATION

Leo S. Goldstein, Ph.D.
Supervisor, Test Validation
Division of Examinations
State of New Jersey
Department of Civil Service

November, 1975

This document was produced under contract #75 NJ04C, funded under the Intergovernmental Personnel Act. Thanks are extended to the funding agency, to Mrs. Barbara Companick who typed the copy, and to my colleagues in the Test Validation Unit who conducted the studies referred to in the Manual.

L.S.G.

Table of Contents

Manual of Test Validation

1. Introduction

A. Need. The advent of the EEOC Guidelines and resultant litigation has created great concern within state selection systems which are merit-based. The concern, of course, is that selection instruments and methodologies be in compliance with the Guidelines. This eagerness to comply with the Guidelines has resulted in a sharp increase in "validation studies" across the nation. Most, if not all of these studies, are content validity studies. Some lesser proportion, are concurrent validations. Few, if any, are in the realm of predictive validity.

The number of merit-based systems venturing into the content validation stream is certain to increase over the next three to five years. Various cooperative efforts will be made to keep redundancy to a minimum but, even so, similar (if not identical) titles and jobs will be validated in multiple jurisdictions. This manual has been developed as a means of "standardizing" the studies and their methods. The information contained in this manual is intended to guide those personnel technicians, who are without previous experience in this area, through the entire process from job analysis to data analysis and final documentation.

B. Approach. If this manual is successful in accomplishing its goal, that success will be due chiefly to the experience gained from three validation studies (see Bibliography) conducted, under grant (No. 75NJ04C) from the Intergovernmental Personnel Program, by technicians in the Test Validation Unit, New Jersey Department of Civil Service. In delineating each step in the validation process, the manual will refer to the studies as exemplars of what is being described. Hopefully, there is sufficient diversity among/between the studies to present a fairly broad range of instances.

C. The Studies referred to deal with these job titles:

1. Custodial-Maintenance Series (C)

2. Firefighter (F)

3. Finance Series (A)

and will be designated as C, F or A in the remainder of this report. Two of the studies represent families of titles. Building Maintenance Worker, Custodial Worker, Park Caretaker, Park Maintenance Worker, and Groundskeeper are subsumed under C. Account Clerk, Audit Account Clerk, Clerk Bookkeeper, and Statistical Clerk comprise the A series.

The following sections delineate the procedures to be followed for each step in a validation study.

II. Job Analysis. Since major concern is on the examination, a thorough job analysis of the title(s) involved is of prime importance. The objective of the job analysis is to identify the tasks and related knowledge, skills, and abilities (KSA) required to perform acceptably in the job title. The selection instrument then is based on the findings of the job analysis, i.e., it is structured around the most important elements of the title.

A. Preliminary Phase. Some preparatory library research is necessary before the job analysis can be accomplished. A minimum of two weeks, perhaps as many as four weeks, should be reserved for "getting acquainted" with the job title(s) being studied. The information collected during this time will assist the technician in preparing the forms which will be used in the pilot and field study phases.

Some excellent sources of information are:

1. Previous studies conducted by other agencies and organizations.
2. The Dictionary of Occupational Titles (U.S. Gov't. Printing Office).
3. Job specifications prepared by the agency's Division of Classifications or those available from other state's agencies.

B. Pilot Study. It is generally good practice for the technician to conduct a preliminary pilot study and "iron out the bumps" before launching a full field investigation.

The earlier research should have identified the location(s) where incumbents in the job title(s) are employed. For example, the geographic distribution and complement size of all paid firefighter jurisdictions in the state had been recorded previously by a technician in the Division of Examinations. Therefore, this information was available for use in Study F. However, similar information was not available for the researchers of

Studies A and C. Mail surveys were conducted in both of these studies to identify the locations and numbers of incumbents in the pertinent titles (Appendix).

Some sampling scheme which takes into consideration factors such as geographical location, agency size, worker's characteristics such as sex, ethnicity, and/or experience should be used to select the group of employees which will be studied.

The preliminary phase library research will also have provided some basic information about the job which the technician can use in constructing forms for collecting data. Data collection may be accomplished through observation, interview, questionnaire, or diary. Each of these methods has advantages and disadvantages which must be weighed before deciding which to use.

1. Observation necessitates the greatest expenditure of time. It demands of the observer constant vigilance over an extended time period as well as an effective system of recording observations. Depending on circumstances, observations can be recorded using shorthand or speedwriting, tape or motion picture. If more than one observer is to be used, a training session will be required to increase inter-observer reliability. Whichever recording method is used, it will require transcription and reduction of the observation material into job task statements.

2. Interviewing, while less time consuming than observation, is still usually a one-on-one method; although interviewing small groups of 3 - 5 persons may be practicable. The technician should construct an interview schedule based on the preliminary phase research. With this as a guide, the technician can probe for the information necessary to construct the job task statements. The schedule also is an element of standardization, if more than one technician is acting as interviewer.

The pilot study for C involved both observation and interview methods. The technician discovered very early that the incumbents were not too adept at describing their daily routine; they lacked sufficient verbal ability. Therefore, a sample of supervisors, who were more verbal, was selected and interviewed.

3. Questionnaires simplify the collection of large amounts of data. However, they must be constructed carefully so that the respondents' replies are readily comprehensible, unambiguous, and complete. The ease of adjusting to sudden changes, which is a virtue of the interview and observation techniques, is not a characteristic of this method. Again, the preliminary phase information, or information gained through interviews or observations, is used as the framework for the questionnaire. If questionnaires are to be used to collect information from a large number of respondents, the technician should plan in advance for the data analysis phase. The questionnaire should be designed so that responses can be easily transcribed into a data processing medium.

4. Diaries or logs may be kept by incumbents in which they record the tasks they perform during each day. The period for which such records are kept must be determined carefully. If the nature of the tasks does not vary, a short time segment,e.g., several days or a week may be adequate. However, if tasks vary over time, a longer recording period or a sequence of shorter intervals may be required. Usually, some interviewing will need to be done, to clarify terms used by the incumbents in their recordings. The technician then consolidates the diary material into task statements.

Analysis of the pilot study data will identify the tasks and corresponding KSAs which define the job title(s). The next requirement is to determine the importance (relative) of each task and the frequency with which it is performed during some standard time segment, e.g., a week or a month. This

information is derived from a larger sample of incumbents than that utilized in the pilot study.

C. Field Study data is usually collected with a mailed questionnaire, because the desired large sample size usually makes other methods of data collection impractical.

1. Instrument Development. The questionnaire will, of course, use whatever information and experience has been garnered from the pilot study. It should be designed to elicit responses which are pertinent, complete, and unambiguous.

In Study C, the pilot study revealed that most job incumbents were not very verbal and, in addition, their average reading ability was quite low. Most could not adequately deal with a questionnaire specially constructed at a low reading difficulty level. Therefore, it was decided to use their supervisors, in most cases the foreman who was in charge of the work team, as the information source. The questionnaire was developed around the pilot study interviews. It was administered, using an interview modality, to small groups (three to five) of supervisors. Each supervisor was required to assign importance ratings from 0 to 100 to each of ten traits which had been identified during the earlier pilot program. The distribution of ratings across the traits was to be such that the sum of ratings equalled 100. (Appendix).

A questionnaire was developed for Study F, based on the pilot information which identified the job tasks and KSAs. The respondent firefighters and their supervisors were requested to assign importance ratings to the KSAs. Again, there was a restriction that the sum of ratings across KSAs should equal 100. (Appendix).

The pilot interviews with incumbents in Study A were analyzed and twelve traits common to the job titles were identified. These were put into a questionnaire format and distributed to the field sample for time and importance ratings. (Appendix).

2. Sample Selection. The sample should be representative of the larger population to which the study's findings are to be applied. Therefore, the population must be adequately described, prior to sampling. Basic information concerning incumbents must be available, e.g., number of workers in each title, location (geographical), agency size, state or local jurisdiction, ratio of males to females, minority representation. These factors, and others which might influence test performance or criterion assessment, should be controlled through proper sampling procedures.

Relevant information about incumbents' location and agency size was available for Study F; selection of a representative sample was a relatively simple chore for the pilot phase. A questionnaire developed from the pilot data analysis was distributed to all firefighter agencies in the State, for the field investigation.

Similar information was not available for Studies A and C. Survey questionnaires had to be developed and distributed state-wide in order to determine the locations of incumbents in the relevant titles, agency size, and other pertinent data. (Appendix). Representative samples were selected after the basic information had been obtained.

3. Implementation. Defining the population and selecting a representative sample are the steps preparatory to the distribution of the data collection instrument. Contact must be made with the appropriate official at each agency who can assist in the distribution of forms and the collection of completed forms. The person to contact might be the personnel director,

the chief of a fire department or police department, or a division director. If possible, telephone contact should be attempted first, since it is usually easier to describe verbally what the study is attempting to do (improve the selection of qualified candidates in the specific title) and how the agency can assist. Once agreement to cooperate with the study has been received, arrangements should be made for the distribution and collection of the data instruments.

In Study F, a cover letter addressed to "Dear Chief" accompanied each packet of questionnaires. Completed questionnaires were to be returned by a specified date in an enclosed, stamped and addressed envelope. The dimensions of Study C permitted the administration of the job analysis questionnaire in a series of small-group interviews conducted by the technician responsible for the study.

Small-group interviews were also conducted in Study A. Agency personnel directors were contacted by phone. They scheduled and assembled incumbents at the required levels for each title. A standard form listing the salient tasks which had been identified through the pilot analysis was used to assess importance and frequency of performance of each task.

4. Data analysis. No matter which mode of collection has been used, the resultant data must be analyzed to identify those job tasks which constitute the important facets of each job title. Analysis can range from simple counting to more complex procedures such as analysis of variance.

For Study F, the average percent importance rating was computed for the KSAs within each defined task. In Study C, mean importance ratings were computed for each of the ten job traits within each of the five titles.

A one-way analysis of variance (ANOVA) was then employed to test for significant differences among the title means for each trait. Study A incorporated the computation of mean importance ratings and mean time (frequency) ratings by title and level. Two-way ANOVAs were then computed in order to identify those tasks for which no significant differences in importance or time occurred between titles or between levels. These tasks would then furnish the core around which the examination would be constructed.

III. Test Development

A. Categorizing Job Tasks. The job analysis will have identified a wide variety of tasks associated with performance in the job title. From among these we must select those which are most important and which can be assessed with an available examination mode, e.g., written, oral, or performance. A task's importance may be evaluated on at least two levels: 1) criticalness, i.e., what amount of harm could be done if the task were not performed or were not properly performed? and 2) frequency, i.e., how often is the task performed during a stated period such as a day, a week, a month?

Some tasks are not readily amenable to assessment. For example, in Study A, the tasks "seeing that others are performing their jobs properly" and "seeing that the assigned work of others is completed on time" don't lend themselves to evaluation by some standard mode. Those tasks whose performance can be assessed and which are important in terms of criticalness and frequency should form the basis for the examination.

B. Item Selection. Identification of the important tasks is followed by a delineation of the KSAs needed to perform the task. In Study C, for example, the most important trait or ability required was that of being able to follow oral directions and demonstration. This ability underlies

the performance of a great variety of tasks which fall within the five job titles studied. Therefore, the examination was built specifically to test this one ability in a wide variety of situations.

In Study F, the major tasks could be categorized into seven areas. Those deemed most important were related to six KSAs: familiarity with general construction principles, with mechanical principles, with general mathematics, ability to read, understand, and apply technical material, ability to transcribe observations into written form, and ability to anticipate critical situations and to react accordingly. These were the focus of the firefighter examination. In Study A, the KSAs which were central to the examination were the abilities to do checking, posting, and mathematical operations.

For Studies A and C, the written examination developed used four-choice multiple choice items. However, Study C pilot interviews had elicited the information that the incumbents generally had low reading level skills. Therefore, the examination was constructed using an oral (taped) presentation synchronized with slides projected on a screen. Respondents were to mark their answers on a special answer sheet.

The items of an examination, whatever mode of administration is used, must be designed so that the behavior which is to be evaluated is elicited from the respondents. Writing good items is a talent inherent in some, acquired by others. Many professional papers and books have information on what constitutes a "good" item. Workshops on item writing and test construction are available through universities and professional organizations. Some sources are referenced at the end of this manual.

If the project director has neither time nor talent for item writing, he may be able to engage the services of a consultant to relieve him of this

chore. However, the onus of selecting items which relate to the KSAs is his whether the items are written "in-house" or obtained through some other means.

IV. Test Validation. The third and final phase is concerned with validation of the examination which has been developed from the job analysis findings. Validation implies relating the examination to some standard. One standard is, of course, the group of KSAs identified through the job analysis as being important in the performance of the job tasks. Validation, then, would be determined by comparing the content of the examination with the description of the KSAs in order to gauge whether the items will elicit the appropriate behaviors. Content validation is a rational, logical, non-statistical assessment of an examination's validity.

An empirical, statistical measure of validity can be obtained by comparing performance on the examination with a standard, such as performance on the job. Correlating examination scores with ratings of performance results in a concurrent validity coefficient, if scores and ratings are derived (within a short time span) from job incumbents and their supervisors. If the examination is given to an unselected group of applicants, and job performance ratings are made subsequently (perhaps six months to two years later) on those who passed the examination and were appointed to the job, a predictive validity coefficient can be obtained by correlating these scores and ratings.

A. Criterion Measures. Validity is concerned with the relationship of an examination to a standard, the criterion. Selection or development of an appropriate criterion measure is of paramount importance. It is not unusual for researchers to develop a predictor instrument which is highly reliable but has a low validity coefficient because of some deficiency with

the criterion. In some job titles, e.g., police officers, supervisors or peers may be reluctant to give middling or low ratings of job performance to people with whom they've worked for some time. The constriction in variation in the criterion measure then results in a low correlation with the predictor.

For some occupations, ability to perform the job can be measured directly in terms of number of units produced, repaired, or sold or something similar. For a large number of job titles, however, such measures of "production" are not available or appropriate. For these, performance can be most readily assessed by asking persons familiar with the worker on-the-job to make a judgment about the individual's ability. In order to have a more reliable estimate independent judgments made by two co-workers (peers or supervisors) should be obtained. This is the procedure which was used in Studies A, C, and F.

In Study A, each incumbent was rated on a number of job tasks by two supervisors: one who was quite familiar with his/her work; the other, somewhat more remote. Each supervisor made independent ratings before joining with the other to make a series of combined ratings. Study C used four raters for each incumbent; two near and two remote. Study F also used four raters; two peers and two superior officers.

In each study the raters were asked to make a judgment about the incumbent's performance. Rating forms were developed to assist in this procedure. Study C used two criterion forms. One presented ten trait situations which might occur on the job; the rater was to select, for each situation, one of five suggested behaviors which the incumbent would exhibit. The second form asked a direct question about each of the ten traits; the rater was to indicate which of five sealed responses was appropriate. Each rater was briefed beforehand on the proper use of the scales. (Appendix).

The criterion forms for Study F instructed the raters on how to assess each incumbent's ability or familiarity with specific job features. Ratings of "inadequate", "below average", "average", "above average", and "superior" were split in two thereby permitting the rater to use a scale of 1-10 in making his judgment. These forms were self-administered, as were most of those in Study C. (Appendix).

The researchers of Study A administered each rating scale to each rater after first determining which task(s) the incumbent performed as part of his/her job. For each such task quantitative and qualitative ratings of accuracy were obtained. The quantitative scale was based on the percent of the incumbent's work which was judged to be error free ranging from 100% to "less than 75%." The qualitative scale used five adjectives ranging from "exceptional" to "weak" to assess the accuracy of the incumbent's performance of each task. In addition, a qualitative measure of the incumbent's overall performance and a quantitative assessment of the "average worker's" performance were obtained.

B. Concurrent Validation. When the examination and the criterion measures have been developed, data collection for the concurrent validation phase can begin. First, a representative sample of incumbents must be selected. All the rules and caveats which applied to the selection of a job analysis sample hold for the validation sample. To the greatest extent possible, sources of bias should be controlled by using appropriate sampling techniques. Often, however, the researcher comes face-to-face with harsh reality and his/her beautiful sampling design is upset and adjustments have to be made. Illness, vacations, holidays, layoffs, non-cooperation are some of the causes of adjusting plans.

In Study C, a proportional sampling scheme had been worked out; the authorities were cooperative in releasing workers' time for taking the examination. Unfortunately, in one locality, a jurisdictional dispute arose between two rival unions fighting for the right to represent some of the incumbents who were to be tested. One of these unions was influential in persuading the incumbents not to participate in the research. Therefore, a sizable portion of the sample which was to represent Black incumbents could not be obtained. No adjustment for this loss could be made.

The examination probably will be administered to the selected incumbents in a number of different locations and/or on different dates. The researcher has to guard against the examination's security being breached. This can be done by permitting only a small number of qualified examiners to administer the examination and by taking the usual precautions such as counting the number of test booklets distributed and returned.

V. Data Preparation. The data collected will consist of examination scores (subtests and total), criterion measures (units "produced" or ratings), and background information such as, incumbent's sex, age, ethnicity, job title and level, education, experience, and any other pertinent information which might be helpful in explaining the results of the study.

If the size of the test sample is large (30 or more), the researcher must consider the value of using computer systems for data analysis. This determination should be made before the data are collected. The forms used for data collection can (and should) be designed to minimize the steps needed to prepare the data for delivery into the computer system. It will be helpful for the researcher to consult with the data processing person early in the study so that the data forms can be properly designed.

Another important element to consider, at about the same time, is the nature of the data analysis that will be performed, i.e., the statistics and statistical tests that will be required in order to interpret the data.

VI. Data Analysis. In a concurrent validation study the key statistic is, of course, the correlation coefficient for test score with a job performance rating. However, a much more intensive analysis of the data is required for the final report.

Data analysis in the validation phase may be categorized into three sections: 1) Examination predictor variable(s), 2) Criterion instrument(s), and 3) Validation statistics. The basic statistics which should be computed for the examination scores include the mean, standard deviation, median, and a measure of reliability such as a split-half correlation (odd-even) and its Spearman-Brown correction. In addition, item statistics, particularly the difficulty level (proportion correct) and a measure of internal consistency such as the point biserial correlation, should be computed. This information will give some indication of whether the examination is "easy" or "hard", how consistently it measures the attributes it was designed to evalute, and how individual items relate to either subtest or total test performance.

The statistics of interest in regard to the criterion instrument(s) are those concerned with intra and inter rater reliability. These are measures of the consistency with which the raters use the rating scales. The former is the within rater consistency; the latter is the between rater consistency. These measures are important chiefly because inconsistency (error) in the use of the criterion will be reflected in the validity coefficient's reduction in size.

In addition to the total test validity, the researcher might be interested in the validity of the subtests, i.e., the correlation of subtest score with the job performance measure. The role of subtests in prediction of a criterion is also examined with multiple correlation and multiple regression techniques. These indicate which combination of subtests yields the highest correlation with the criterion and the proportional weight that each subtest contributes to the prediction.

If the data warrant examination for specific effects of background variables or for testing hypotheses, the research will use statistical procedures such as: t tests, for testing differences between means; analysis of variance, to assess the effect of specific attributes on test performance; factor analysis, to help "explain" the structure of an examination from the item intercorrelations.

VII. Final Report. All the effort that has gone into the study will have been for naught,if some documentation, a record of what was done and what was found, is not produced. Although it is the final step in the entire procedure, much thought should be given to the final report, and appropriate preparation made, as the study develops. The need for maintaining a history of the study from start to finish cannot be overstressed. A diary would be very appropriate; and a diarist. Copies of all forms which are developed, memoranda of meetings, interview schedules, in short, everything relevant to the study should be filed against the day when the final report is to be written. The use of cryptic notes and symbols should be avoided. If something isn't clearly expressed in January it will be completely muddled six months hence.

The final report, as all good documentation, has order to its structure. This structure can be determined early in the study; once the research design has been developed. The project director or research

director should write down what the procedures will be so that all involved in the study will be aware of the direction in which the study is headed.

At the outset, the final report should establish the need for the study. What conditions have given rise to it? What are the goals? How are they to be achieved?

In validation studies, the procedure is basically that which has been described in this manual. These will be the major headings of the report: preliminary research for the job analysis, pilot study, field study; test development, test validation, results, discussion and recommendations.

The report writer should sacrifice all for clarity; too often the results of good research are buried in a morass of verbiage. Tell the reader what was done, why it was done, how it was done, and what were the results. Use tables where they will help clarify the text; put less important materials in the appendix.

The final report is the chief avenue of contact between the study and the outside world. Care should be taken in its construction so that the reader will be able to reap full benefit from the researcher's efforts.

This manual has attempted to picture, for the unexperienced researcher, the processes which comprise a validation study. Many topics, of necessity, have been treated somewhat cursorily. However, the broad strokes have been applied to the canvas. It is left to the researcher to apply the finer touches. To assist those who will be embarking on validation studies of their own, the following references will be found of great value.

VIII. Suggested Readings

A. Job Analysis.

Job Analysis: developing and documenting data. United States Civil Service Commission. Bureau of Intergovernmental Personnel Programs. Washington, D.C. BIPP 152-35 December, 1973.

B. Item Writing.

Wesman, A.G. Writing the Test Item in Educational Measurement (second edition), Thorndike, R.L. editor, American Council on Education, Washington, D.C. 81-129.

C. Reliability, Validity, General Statistical Concepts.

Thorndike, R.L. and Hagen, Elizabeth. Measurement and Evaluation in Psychology and Education, John Wiley and Sons, Inc. New York.

BIBLIOGRAPHY

Gerczak, T. A Concurrent Validation Study of Custodial-Maintenance Titles. New Jersey Department of Civil Service, Division of Examinations. July, 1975.

Skinas, N. and Goldstein, L.S. Construction and Validation of an Entry-level Firefighter Examination. New Jersey Department of Civil Service, Division of Examinations. September, 1975.

Goldstein, L.S. Wisniewski, Leslie H., and Faley, R.H. Concurrent Validation of a Prototype Examination for Selected Clerical Titles. New Jersey Department of Civil Service, Division of Examinations. November, 1975.

APPENDIX

JOB ANALYSIS QUESTIONNAIRES

a) Study C

b) Study F

c) Study A

Job Title _____

Name _____

Agency _____

Date _____

The following factors have been found to be important in the performance of maintenance jobs. If you were given the job of hiring a maintenance worker how much consideration would you give to each of these factors when making your decision? Please weigh all the factors below by giving each a number value from 0-100, giving a higher value to those factors you feel are the most important. Tht total of all the numbers should equal exactly 100. Look at all the factors before beginning.

1. How well the worker would be able to follow oral directions and demonstrations. _____

2. How well the worker would be able to work together with a group so that he would get along with them by being friendly, cooperative, and doing his share of the work. _____

3. How quickly the worker would be able to learn and remember where supplies are found and where work areas are located so that it would not slow down his on-the-job training. _____

4. How well the worker would be able to adjust to situations which he does not like, such as being told to do something that he does not really want to do. _____

5. How well the worker would be able to perform tasks that are routine and repetitive, like sweeping, mopping, window washing, etc. _____

6. How well the worker would be able to perform tasks requiring either strength or physical fitness, like loading heavy equipment onto trucks, shoveling snow, mowing grass, etc. _____

7. How well the worker would be able to perform tasks which require him to take on responsibility, like taking good care of his equipment, serving as a watchman or doing extra work that was needed without having to be told to. _____

8. How well the worker would be able to relate to the public or make a good impression by being courteous and having a neat appearance. _____

9. How good the worker's attendance would be. _____

10. How punctual the worker would be. _____

_____

Job Analysis Rating Form

_____ FORM _____

NAME _____

TITLE _____

LENGTH OF TIME IN TITLE _____

AGENCY _____

DATE _____

1_____ Writing or typing information on forms (either selected items or the entire form)

2_____ Reviewing information (words, numbers, codes, calculations, etc.) on written forms or calculation sheets, and making changes where necessary

3_____ Transferring words and/or numerical information from one form to another or to a master form

4_____ Performing addition, subtraction, multiplication, division, and other basic mathematical operations such as interest rates or percentages either with a calculator or by hand

5_____ Composing original letters and/or memos of a business nature

6_____ Making and/or receiving calls to or from other offices, businesses, or the public

7_____ Directly teaching another employee the duties, responsibilities, knowledges and skills required in the performance of the job

8_____ Classifying or categorizing verbal or numerical information

9_____ Answering job-related questions from co-workers at one's level or at a lesser level

10_____ Seeing that others are performing their jobs properly

11_____ Seeing that the assigned work of others is completed on time

12_____ Dealing with the public on a face-to-face basis

Instructions for Filling Out Job Analysis Rating Forms

(Time Spent Rated First)

Introduction

On the basis of interviews with Audit-Account Clerks, Clerk Bookkeepers, Statistical Clerks, Payroll Clerks, and Account Clerks, we have come up with a list of twelve tasks, all of which or some of which, are performed by employees in these titles. We would like to rate these tasks, in terms of time spent and relative importance, so that we will know what types of items to emphasize on the test.

We prefer that you do these ratings individually as we are interested in your personal opinion - not in a group consensus. We want you to rate the tasks as they apply to your particular level, not to the overall titles.

On the forms that will be handed out are listed descriptions of tasks that may or may not be performed by someone in your title.

Here is the first form. Please write a number "one" in the blank before the word "form" and the word "time" in the blank after the word "form". This tells that you are first rating these tasks according to the amount of time spent on them.

Now fill in your name, exact title including level, length of time in that specific title (including any time that you were provisional or temporary), agency, and today's date.

On this form we would like you to fill in the percentage of time, 0% to 100%, spent in performing each task. In assigning these percentages of time, please view your job over a one year period. The reason for this is that some aspects of your job may be infrequently done or seasonal in nature. If a task does not apply - if you do not perform it - fill in 0%. The total for all tasks must equal exactly 100%.

Let's read through all the tasks together before we begin filling in any percentages.

(Read one task at a time and ask for questions after each one.)

Remember, we want you to assign a percentage rating to each task based on the amount of time spent on that task. Each blank must have a rating, even if it is zero; and the total must be 100%. A useful approach is to first assign zeroes to those tasks that you do not perform. Then divide the 100% among the remaining tasks.

If there are no further questions, you may begin.

(Wait until form is completed.)

Does everyone have a percentage rating for each task, including zeroes for those that do not apply? Does the sum of all ratings equal 100%?

Now, let's read through the tasks again making sure you understand each one and are satisfied with the rating given. Make any changes you deem necessary.

(Collect first form; pass out second form.)

Here is the second form. Please fill in a number "two" in the blank before the word "form" and place the word "importance" in the blank after the word "form". Now fill in your name and other information.

On this form we would like you to rate the relative importance of each task using a scale of 0-100 points. A higher rating indicates a more important task. If a task does not apply, it should be rated zero. If you spend no time on a task, it cannot be important to your job; if you spend time on a task, it must have some importance to your job. Therefore, if you assign a zero to a task on one form, there must be a corresponding zero for that task on the other form. As before, the total for all tasks must equal 100.

There may be a question as to what we mean by importance. We are interested in how essential each task is to the overall performance of your job. Importance in this instance is not related to time spent on the task. You can spend a great deal of time on a task that has little importance or little time on a task that is very important. As an example, a dental assistant may tidy up the waiting room five times a day but develop x-rays only once a day. Developing x-rays may take less time than attending to the waiting room, but it is more important. Failure to perform an important task may result in a job improperly done or a deadline not met.

If there are not questions, please assign a rating to each task.

(Wait until form is completed.)

Does everyone have a rating for each task, including zeroes for those that do not apply? Do the zeroes on one form correspond with the zeroes on the other form? Does the sum of all ratings equal 100 points?

Now let's read through the tasks again to make sure you are satisfied with the ratings you have given. Make any changes you deem necessary.

(Collect the forms. Check that each form totals 100, all blanks are filled, and zeroes correspond.)

Instructions for Filling Out Job Analysis Rating Forms

(Importance Rated First)

Introduction

(Same as for the first set of instructions.)

On the forms that will be handed out are listed descriptions of tasks that may or may not be performed by someone in your title.

Here is the first form. Please write a number "one" in the blank before the word "form" and the word "importance" in the blank after the word "form". This tells us that you are first rating these tasks according to importance.

Now fill in your name, exact title including level, length of time in that specific title (including any time that you were provisional or temporary), agency, and today's date.

On this form, we would like you to rate the relative importance of each task using a scale of 0-100 points. A higher rating indicates a more important task. If a task does not apply - if you do not perform it - fill in zero. The total for all tasks must equal exactly 100 points.

There may be a question as to what we mean by importance. We are interested in how essential each task is to the overall performance of your job. Importance in this instance is not related to time spent on the task. You can spend a great deal of time on a task that has little importance or little time on a task that is very important. As an example, a dental assistant may tidy up the waiting room five times a day but develop x-rays only once a day. Developing x-rays may take less time than attending to the waiting room, but it is more important. Failure to perform an important task may result in a job improperly done or a deadline not met.

Let's read through all the tasks together before we begin filling in any points.

(Read one task at a time and ask for questions after each one.)

Remember we want you to assign a rating of points to each task based on its relative importance. Each blank must have a rating, even if it is zero; and the total must be 100 points. A useful approach is to first assign zeroes to those tasks you do not perform. Then divide the 100 points among the remaining tasks.

If there are no further questions, you may begin.

(Wait until form is completed.)

Does everyone have a rating for each task, including zeroes for those that do not apply? Does the sum of all rating equal 100 points?

Now, let's read through the tasks again making sure you understand each one and are satisfied with the rating given. Make any changes you deem necessary.

(Collect first form; pass out second form.)

Here is the second form. Please fill in a number "two" in the blank before the word "form" and place the word "time" in the blank after the word "form". Now fill in your name and other information.

On this form we would like you to fill in the percentage of time, 0% to 100%, spent in performing each task. In assigning these percentages of time, please view your job over a one year period. The reason for this is that some aspects of your job may be infrequently done or seasonal in nature. If a task does not apply, it should be rated zero. If you spend time on a task, it must have some importance to your job. Therefore, if you assign a zero to a task on one form, there must be a corresponding zero for that task on the other form. As before, the total for all tasks must equal 100.

If there are no questions, please assign a percentage rating to each task.

(Wait until form is completed.)

Does everyone have a rating for each task, including zeroes for those that do not apply? Do the zeroes on one form correspond with the zeroes on the other form? Does the sum of all rating equal 100%?

Now let's read through the tasks again to make sure that you are satisfied with the rating you have given. Make any changes you deem necessary.

(Collect the forms. Check that each form totals 100, all blanks are filled, and zeroes correspond.)

MAIL SURVEY FORMS

a) Study A

b) Study C

State of New Jersey

DEPARTMENT OF CIVIL SERVICE

WILLIAM DRUZ
CHIEF EXAMINER & SECRETARY

EAST STATE & MONTGOMERY STREETS
TRENTON, N. J. 08625

February 14, 1975

Dear Personnel Director:

The Test Validation Unit, Division of Examinations, is presently involved in developing a new test for all levels (entrance, senior, principal, and head or supervising) of the following State and local titles: Audit Account Clerk, Account Clerk, Clerk Bookkeeper, Statistical Clerk and Payroll Clerk. In order to develop a valid test, we need to determine how many permanent full time employees are presently in these job titles.

For future planning, we would also like an estimate of your anticipated needs in these titles for the fiscal year 1975-76. This information will be useful to the Division of Examinations in determining future testing needs.

The survey form and a stamped self-addressed envelope are enclosed. We would appreciate receiving your completed form by March 7, 1975. If more information is needed, contact Ms. Leslie Wisniewski at 609-292-6940.

Thank you for your cooperation.

Sincerely,

Leo S. Goldstein

Leo S. Goldstein
Supervisor
Test Validation and Staff Development

LSG:ac

Enclosure

(Survey Letter)

Name of person completing this form _____

Name of jurisdiction _____

Name of agency _____ Phone No. _____

INSTRUCTIONS *

In the table below under the columns headed "Present", indicate the number of permanent full time employees there are at each title and level. If at present there are no employees at a specific title and level, write an "X" in that space. Under the columns headed "Estimate", indicate the number of new employees you estimate will be needed at each title and level for the fiscal year 1975-76.

	Entrance		Senior		Principal		Head or Supervising	
	Present	Estimate	Present	Estimate	Present	Estimate	Present	Estimate
udit Account Clerk								
ccount Clerk								
lerk Bookkeeper								
tatistical Clerk								
ayroll Clerk								

If any of these titles are non-competitive (candidates are certified without a test) please note by placing a check mark next to that title.

After completing the survey form, please return it in the enclosed stamped, addressed envelope by March 7, 1975.

STATE OF NEW JERSEY
TEST VALIDATION AND STAFF DEVELOPMENT
ARNOLD CONSTABLE BUILDING
FRONT AND MONTGOMERY STREETS
TRENTON, NEW JERSEY 08625

(Survey Form)

State of New Jersey

DEPARTMENT OF CIVIL SERVICE
ARNOLD CONSTABLE BUILDING
FRONT AND MONTGOMERY STREETS
TRENTON, N. J. 08625

January 2, 1975

WILLIAM DRUZ
CHIEF EXAMINER & SECRETARY

Dear Personnel Director:

The Test Validation Unit, Division of Examinations, is presently involved in developing a new test for several maintenance titles. In order to develop a valid test, we need an estimate of how many workers are presently in these job titles. We also need to know if these job titles are competitive (tested for) or non-competitive (certified without test) in your jurisdiction. (Check the appropriate box for each title.)

The survey form and a self-addressed envelop are enclosed. We would appreciate receiving your completed form by January 17, 1975. If more information is needed, contact Mr. Tony Gerczak at 609-292-6940.

Thank you for your cooperation.

Sincerely,

Leo S. Goldstein

Leo S. Goldstein, Ph.D.
Supervisor
Test Validation and Staff Development

LSG:ac

Enclosure

	Competitive	Non-competitive	Total number of employees
Building Maintenance Worker	☐	☐	_____
Custodial Worker	☐	☐	_____
Groundskeeper	☐	☐	_____
Park Caretaker	☐	☐	_____
Park Maintenance Worker	☐	☐	_____

Name of your jurisdiction _____

Name of person completing this form _____ Phone No. _____

CRITERION INSTRUMENTS

a) Study C (Appendix L)

b) Study C (Appendix M)

c) Study F (D.)

d) Study A

A. You have just received a new riding mower to replace the hand mower you previously used. You must explain to the workers how to run and use this mower. How well would the worker be able to use this mower after hearing your detailed instructions?

He would:

/___/ be able to use it properly with no further explanations.

/___/ be able to use it properly with some further explanations.

/___/ have to hear some parts of the instructions again and again before he could use it properly.

/___/ have to be shown and told repeatedly.

/___/ be unable to use the mower properly even with repeated instructions.

B. You have assigned this worker and five others to pick up papers in a large open area. How well would the worker perform this task?

He would:

/___/ pick up more paper than anyone else in the group.

/___/ pick up as much paper as the next worker.

/___/ work only if others work.

/___/ allow other workers to do most of the work.

/___/ not work if put in a group.

C. You are changing the locations of all your supplies. Previously, all your supplies were located in one place.Now they are to be found in five different locations. How quickly would the worker learn and remember where to find the supplies?

He would:

/___/ learn the locations after being shown only once.

/___/ learn the locations after being shown a number of times.

/___/ would need occasional help to remember.

/___/ have great difficulty in remembering.

/___/ not be able to remember.

D. This worker has always disliked taking out the trash, but the worker who usually is assigned this duty is out ill. You have asked the worker to take out the trash.

He would:

/___/ willingly do it as best he could.

/___/ do it.

/___/ do it but show unhappiness.

/___/ argue bitterly and try to get out of doing it.

/___/ not do it unless he was forced to.

E. You have assigned this worker a repetitive task such as washing windows on a day to day basis. How well would he perform this duty?

He would:

/___/ perform this duty well all the time.

/___/ do it but not as well as he would perform other types of duties.

/___/ slack off as work became boring.

/___/ perform this duty poorly from the start.

/___/ try to get out of doing it.

F. A large shipment of supplies has arrived and must be unloaded. There is no other help around, and most of these supplies are rather heavy (75 pounds). You have assigned this worker to do this duty. How well would he perform this duty?

He would:

/___/ be able to unload the shipment in the least amount of time.

/___/ complete the job in a reasonable time.

/___/ complete the job but work slower than usual.

/___/ become too tired to finish the job.

/___/ be unable to do heavy work.

G. You have assigned this worker a new lawn mower. You told him that it must be oiled everyday before mowing so that it would be kept in good working condition. How frequently would he oil this mower?

He would:

/___/ never forget to oil the mower everyday as directed.

/___/ rarely forgot to oil the mower.

/___/ occasionally forget to oil the mower.

/___/ need continual reminding to oil the mower.

/___/ not oil the mower at all.

H. While mowing the lawn, a private citizen comes to the worker and complains that the mower is making too much noise. The citizen then asks the worker to mow the lawn at a later time. How would the worker react to this citizen?

He would:

/___/ go out of his way to be polite.

/___/ listen to the citizen.

/___/ be indifferent to the citizen's complaint.

/___/ avoid the public at all costs.

/___/ be very rude.

I. You have a job coming up the next day that you know this worker does not like to do. Would this worker come to work anyway?

/___/ yes

/___/ depends how much he dislikes the work.

/___/ cannot say

/___/ only if he can get out of doing the work.

/___/ no

J. You have certain equipment that must be carried to another area. To get it there on time, the workers must begin transporting it immediately after lunch. If one worker is late from lunch, it will delay everyone.

This worker would be:

/___/ always on time.

/___/ a little late but for good reason.

/___/ a couple of minutes late.

/___/ very late.

/___/ very late all the time.

1. How well does the worker perform tasks once you have shown or demonstrated the procedure to him?

He:

/___/ learns all demonstrations very quickly.

/___/ needs more demonstrations only for more difficult tasks.

/___/ does them correctly if they are not too difficult.

/___/ needs additional demonstrations even on easier tasks.

/___/ has great difficulty in learning from demonstrations.

2. How well does the worker cooperate with others when working in a group?

He:

/___/ cooperates with all members of the group.

/___/ cooperates with most members of the group.

/___/ cooperates with some members only.

/___/ cooperates if he feels like it.

/___/ doesn't cooperate with any group members.

3. How well does the worker remember where supplies and work areas are found?

He:

/___/ remembers where everything can be found.

/___/ remembers where most things are found.

/___/ remembers where some things are found.

/___/ needs help in finding most things.

/___/ cannot remember where things are found.

4. How well does the worker perform duties that he doesn't want to do?

He:

/___/ performs them to the best of his ability.

/___/ performs them but slacks off.

/___/ perform them but shows his unhappiness.

/___/ tries to get out of doing them.

/___/ has to be forced into doing them.

5. How well does the worker perform routine everyday tasks such as dusting, picking up papers, or sweeping?

He:

/___/ does an outstanding job in all types of tasks.

/___/ does a consistently good job.

/___/ does an acceptable job.

/___/ does a "half-hearted" job.

/___/ tries to avoid these repetitive tasks.

6. How well does the worker perform duties which require strength and endurance?

He:

/___/ has no problems even with the heaviest work.

/___/ has some difficulty with the heaviest work.

/___/ can do this work if he has help.

/___/ can do this work but for a very short time.

/___/ has great difficulty with all work of this type.

7. How well does the worker perform duties which require him to take on responsibility such as taking good care of his equipment?

He:

/___/ can be relied upon at all times.

/___/ keeps his equipment in good condition most of the time.

/___/ is lax at times.

/___/ usually takes improper care of his equipment.

/___/ cannot be given responsibility to take care of equipment.

8. Is the worker polite in his contact with the public?

He:

/___/ goes out of his way to be polite to the public.

/___/ is generally polite.

/___/ is indifferent.

/___/ tries to avoid the public.

/___/ is rude.

9. How good is the worker's attendance?

He:

/___/ is never out except for illness or other good reason.

/___/ is rarely out.

/___/ has occasional unexplained absences (about once every two months).

/___/ has frequent unexplained absences (more than once a month).

/___/ is out very often.

10. How often is the worker more than five minutes late?

He is:

/___/ always on time for work.

/___/ usually on time.

/___/ late for work about once every two weeks.

/___/ late for work about half the time.

/___/ always late for work.

Date _____ Location _____

Name of Worker _____ Social Security No. _____

Name of Supervisor _____ Title _____

Length of Time You Have Supervised This Employee: /___/ years /___/ months

/___/ Male /___/ Female /___/ Black /___/ White /___/ Hispanic

SPEED - Q		ACCURACY - Q		SPEED - A		ACCURACY - A	
Single	Combined	Single	Combined	Single	Combined	Single	Combined
AV							
AV							
AV							
AV							
AV							

Single	Combined

Q = Quantitative
A = Adjective

Original Rating Scales

Interviewer _____

Date _____ Location _____

Name of Worker _____ Social Security No. _____

Name of Supervisor _____ Title _____

Length of Time You Have Supervised This Employee: /___/ years /___/ months

/___/ Male /___/ Female /___/ Black /___/ White /___/ Hispanic

	ACCURACY - Q		ACCURACY - A	
	Single	Combined	Single	Combined
AV				
AV				
AV				
AV				
AV				

Single	Combined

Revised (Final) Rating Scale

Interviewer _____