



094208

# REPORT TO THE CONGRESS

## Improvements Needed In Examining And Selecting Applicants For Federal Employment B-179810

Civil Service Commission

*BY THE COMPTROLLER GENERAL  
OF THE UNITED STATES*

~~710702~~

7094008



COMPTROLLER GENERAL OF THE UNITED STATES

WASHINGTON, D.C. 20548

B-179810

To the President of the Senate and the  
Speaker of the House of Representatives

This is our report on improvements needed in examining  
and selecting applicants for Federal employment.

We made our review pursuant to the Budget and Account-  
ing Act, 1921 (31 U.S.C. 53), and the Accounting and Audit-  
ing Act of 1950 (31 U.S.C. 67).

We are sending copies of this report to the Director,  
Office of Management and Budget, and to the Chairman, Civil  
Service Commission.

A handwritten signature in cursive script that reads "James B. Stacks".

Comptroller General  
of the United States

## C o n t e n t s

	<u>Page</u>
DIGEST	i
CHAPTER	
1 INTRODUCTION	1
Examining	2
Selection	3
2 RELIABILITY OF UNASSEMBLED EXAMINATIONS NEEDS IMPROVEMENT	5
Descriptions of unassembled examinations	5
Testing the unassembled examination rating process	6
Results of test	6
Effect of differences in ratings	7
Causes of unreliability	8
Conclusions and agency actions	8
3 MODIFICATIONS NEEDED IN SELECTION SYSTEM	10
Precision of testing	10
Validity	11
Reliability	12
The rule of three	13
Alternative selection systems	17
Category-ranking system	17
Rule of ranks	17
Rule of reliability	18
Other systems	18
Agency comments	18
Conclusions	19
Matters for consideration by the Congress	20
4 SCOPE OF REVIEW	21

APPENDIX

I	Letter dated March 26, 1974, from the Executive Director of the United States Civil Service Com- mission	23
II	Principal CSC officials responsible for administering activities discussed in this report	27

ABBREVIATIONS

GAO	General Accounting Office
CSC	Civil Service Commission

## GLOSSARY

Certify	The process by which a CSC area office prepares a list, or certificate, of the names of eligible applicants from a register and sends it to an agency for appointment consideration.
Competitive examining program	A program which measures the relative qualifications of applicants in competition for given positions.
Eligible	Any applicant who meets the minimum entrance requirements and is rated "eligible" for Federal employment or receives a numerical rating of 70 points or more in a CSC examination.
Examination	A means of measuring, in a practical and suitable manner, the qualifications of applicants for employment in specific positions in the Federal service.
Federal merit system	A complete system of personnel selection and management based on an integrated set of personnel policies, procedures, and practices designed to (1) recruit a competent work force, (2) insure a stable work force, and (3) provide equal opportunity for employment.
Psychometrician	One who measures the duration, force, interrelations, or other aspects of mental processes, as by psychological tests.
Register	A list of qualified applicants compiled in order of relative standing for certification. Same as "eligible" list.

Reliability

The degree of consistency or repeatability in measuring the qualifications of an applicant in an examination.

Selective certification

Certifying only the names of eligibles who have special qualifications required to fill particular positions.

Spoils system

A personnel system characterized by the political appointment and removal of employees without regard to merit.

D I G E S T

WHY THE REVIEW WAS MADE

Each year over two million persons apply for Federal employment in the competitive civil service. In fiscal year 1973 nearly 180,000 were hired.

The Civil Service Commission (CSC) administers the examination and selection system that screens these applicants. Its objectives are to evaluate and rank applicants on merit and fitness and provide Federal agencies with the names of well qualified applicants. (See p. 1.)

In fiscal year 1973 approximately 55 percent of the applicants took an "assembled examination," whereby candidates assemble at an appointed time and place to take a written examination or a performance test.

The remaining 45 percent were evaluated under an "unassembled examination," whereby CSC assesses the candidate's education, training, and experience from his or her formal application.

Those passing the examinations are designated as "eligibles" and are placed on a civil service employment register according to their scores.

According to law a Federal agency must select new employees from the first three names listed on the register. This process is commonly referred to as the "rule of three." (See pp. 2 and 3.)

Both the first and the second Hoover Commissions, in 1949 and 1955, respectively, recommended replacing the rule of three for many jobs in the Federal Government. CSC agreed for certain high-level positions and for midlevel positions, and in the early 1950s recommended that legislation be enacted along the lines proposed by the first Hoover Commission.

According to CSC the recommendations were not adopted because necessary amendments to the statutory requirement for the rule of three were not enacted. (See p. 4.)

GAO reviewed CSC's unassembled examination procedures to determine whether applicants are being reliably evaluated, scored, and ranked on employment registers. GAO also examined the fairness of the rule of three. (See p. 4.)

FINDINGS AND CONCLUSIONS

GAO found that the reliability of CSC's unassembled examinations needed improvement so that applicants' scores can be more consistently determined. At GAO's request, CSC rerated several hundred applications selected at random from CSC files.

Rerated scores varied an average of about 3.5 to 5.5 points from the originally assigned score. Thus, rerated scores, if used, would have made a difference for many of the

sample group in their relative standing on the employment register and in their chances of being considered for employment under the rule of three. (See pp. 5 to 7.)

CSC has initiated corrective actions in the unassembled examining process to increase examination scores' reliability. These include:

- Improving the clarity and usefulness of CSC rating guidelines.
- Developing a training program for CSC rating examiners to increase their proficiency.
- Developing a quality assurance program to provide continuing and statistically valid review of the quality and consistency of ratings. (See pp. 8 and 9.)

Practical limitations in personnel testing and measurement prevent both assembled and unassembled examinations from being accurate--that is perfectly reliable (consistent in measurement) and perfectly valid (job related).

Improved evaluation procedures can reduce some of this imprecision, but a large part is irreducible because subjective judgments and other uncontrollable elements will always be present.

Because of its limitations, the examination process is not precise enough to judge the potential job success of persons with identical or nearly the same scores.

Examining techniques, however, are accurate enough to insure that a person with a high score will generally be more competent than one with a much lower score. (See pp. 10 to 14.)

CSC registers often include many applicants with the same or nearly the same scores. In case of identical scores, names are usually placed on the register in alphabetical order or other means which do not relate to the applicant's job qualifications. (See p. 14.)

The effect of these factors is that, under the rule of three, many applicants who may have qualifications equal to others could be denied the opportunity to be considered for employment.

Officials of several major agencies said the rule of three has hindered them in hiring applicants they considered the most qualified available. (See p. 16.)

The CSC examining process can, if improved, reliably screen applicants into quality levels or score groups in which applicants would be considered comparably qualified.

Federal agencies should, within practical limits, be permitted to consider and fill vacancies from among all applicants in the best qualified level or group.

These changes would reduce the degree of arbitrariness currently in the rule of three selection process.

Available data indicated that among all States less than one-third are using a rule of three or a procedure more restrictive. (See p. 18.)

Several alternate methods of accomplishing these changes, such as category ranking, the rule of reliability, and the rule of ranks, would also be fairer to the applicants, provide agencies greater opportunity to obtain well qualified applicants, and still provide



necessary merit system safeguards.  
(See pp. 17 to 20.)

RECOMMENDATIONS

None.

AGENCY ACTIONS AND UNRESOLVED ISSUES

In July 1973 CSC furnished GAO a plan for improving unassembled examination procedures. Since then CSC has taken and planned other actions designed to improve the rating guidelines, training of rating examiners, and quality assurance program. If adequately implemented, these actions will increase the reliability of its unassembled examination rating process. In its continuing reviews of Federal personnel matters, GAO plans to evaluate the effectiveness of CSC's actions. (See pp. 8 and 9.)

CSC believes the rule of three has served the country well but that this issue is an important policy question that deserves to be aired and considered.

CSC said it has not advocated changes to the rule of three since the Hoover Commission reports were issued because the examining system has become flexible through other changes. (See p. 19.)

MATTERS FOR CONSIDERATION  
BY THE CONGRESS

In GAO's opinion, the rule of three selection requirement is unrealistically rigid. GAO recommends that the Congress amend the requirement and allow CSC to prescribe alternative selection procedures, similar to those discussed in this report. CSC should design the procedures in such a manner that they will provide

- applicants with a reasonable opportunity to be considered for employment,
- Federal agency management greater opportunities to identify and select well qualified applicants,
- safeguards needed to insure compliance with merit principles.

Each agency should then be permitted to choose the selection procedures best suited to its particular needs-- just as the Tennessee Valley Authority and the Atomic Energy Commission have done. The individual agency's application of the procedures should be subject to CSC's continuing review and evaluation to assure compliance with CSC guidelines and policies.

Any change in the rule of three selection requirement need not affect the preferential rights of veterans. (See p. 20.)

## CHAPTER 1

### INTRODUCTION

Under Presidential direction the United States Civil Service Commission (CSC) provides the overall leadership for personnel management in the Federal Government. The Civil Service Act of 1883 (Ch. 27, 22 Stat. 403) established CSC to administer a merit staffing system, including planning and carrying out a competitive examining program, to meet the personnel needs of the Federal competitive service.

The major objectives of the merit staffing system are to (1) provide the public with a systematic means to compete for available Federal jobs in the competitive civil service, (2) evaluate and rank applicants solely on the basis of merit and fitness, and (3) refer the best qualified applicants to Federal agencies for consideration in filling job vacancies.

At the time of its creation, CSC's principal function was to establish and maintain a merit system and to protect it against advocates of the spoils system. The Civil Service Act, and later the Veterans' Preference Act of 1944, provided procedures under which applicants were appointed to Federal jobs on the basis of their qualifications and fitness, rather than on the basis of personal preference or political considerations.

CSC conducts its operations in headquarters offices in Washington, D.C., and in 10 regional offices and 65 area offices. During fiscal year 1973 it spent about \$27 million, or 41 percent of its total budget, on examination and selection. It processed 1.7 million applications and sent 1.3 million names to Federal agencies, from which 179,000 selections were made.

The examination and selection system must be designed to insure that applicants are fairly examined and evaluated and that their opportunity for employment depends on their qualifications. It must also insure Federal agencies that the candidates they are considering hiring are well qualified.

## EXAMINING

CSC gives competitive examinations as part of the process of filling the employment needs of the Government. Section 2 of the Civil Service Act of 1883 (5 U.S.C. 3304) states the objective of examinations:

"The President may prescribe rules which shall provide, as nearly as condition of good administration warrant, for--

"(1) open, competitive examinations for testing applicants for appointment in the competitive service which are practical in character and as far as possible relate to matters that fairly test the relative capacity and fitness of the applicants for the appointment sought."

Competitive examinations are the means by which equality of opportunity is united with efficiency of performance.

CSC uses two different evaluation procedures. The most familiar procedure is the written test, which is particularly well adapted to the testing of aptitudes or specific knowledge. CSC refers to written tests and some performance tests, as "assembled examinations" because candidates are required to assemble at an appointed time and place to take the test. These tests are used for some clerical and lower level jobs and for the Federal Service Entrance Examination, the recent college graduate's first means of entry into most occupations in the Federal service.

The second procedure used is CSC's assessment of the candidate's education, training, and experience as shown on his or her formal application. This is called an "unassembled examination" since individuals are not required to assemble for a written examination. Unassembled examinations are used to evaluate the qualifications of most applicants for trade and labor jobs and mid- and upper-level management jobs.

During fiscal year 1973, approximately 55 percent of all applicants took a written examination or performance test; the remainder were evaluated by unassembled examination procedures. Those applicants who receive a score of at least 70 points (on a scale of 100) are considered to have passed the examination and are designated as "eligibles." They are then

placed on job registers in order of their score. If appropriate, 5 or 10 points are added to veterans' scores. As required by law, preference eligibles, such as veterans, are listed ahead of other eligibles receiving the same score. The higher an applicant's score, the better qualified he or she is deemed to be. The higher an applicant is listed on the register, the greater the probability that he or she will be considered for Federal employment.

### SELECTION

When an agency does not fill a vacancy through promotion or reassignment from within, it requests CSC to provide the names of those eligible for appointment. Generally, there is a separate register for each job category and grade level. Each register is maintained by one or more CSC offices serving Federal agencies in their geographical areas.

According to law (5 U.S.C. 3318), the agency must select from the three eligibles with the highest scores. This procedure is commonly referred to as the "rule of three." The agency may pass over the three highest eligibles only for legitimate, documented reasons satisfactory to CSC.

Both the first and the second Hoover Commissions on Organization of the Executive Branch recommended adopting a category ranking system like that of the Atomic Energy Commission and Tennessee Valley Authority to replace the rule of three for many jobs in the Federal Government. In February 1949 the first Hoover Commission recommended that officials in the Federal Government be given more leeway than the rule of three permits in selecting personnel and proposed that applicants be grouped into categories, such as "outstanding," "well qualified," "qualified," and "unqualified."

CSC agreed with the proposal for category ratings in certain high-level positions, use of a "rule of five" for midlevel positions, and retention of a "rule of three" for lower level positions. During the early 1950s CSC recommended that legislation be enacted along the lines proposed by the first Hoover Commission.

In February 1955 the second Hoover Commission recommended that CSC (1) be authorized by law to allow appointing officers greater leeway in selecting personnel by permitting

them to make their selection from among five certified eligibles rather than three, except for appointments to lower grade positions and the postal field service, and (2) make full use of its authority to substitute category rating for numerical rating of applicants' qualifications for certain higher scientific, technical, and administrative positions.

According to CSC, it did not adopt either of the first or the second Hoover Commissions' recommendations because necessary amendments to the statutory requirement for the rule of three were not enacted.

Since the CSC examination and selection system assume a large part of the responsibility for the quality of the men and women who fulfill the demanding tasks of civilian government, we examined CSC's practices in evaluating, scoring, and ranking applicants who do not take a written test--applicants whose qualifications are determined by an unassembled examination. (See chapter 2.) We also examined the fairness of the rule of three which applies to candidates evaluated by either assembled or unassembled examinations. (See chapter 3.) We did not, however, evaluate Federal agencies' recruiting efforts or determine whether the Federal Government is as effective as it can be in attracting the most qualified candidates available.

## CHAPTER 2

### RELIABILITY OF UNASSEMBLED EXAMINATIONS

#### NEEDS IMPROVEMENT

The shortcomings in the reliability of CSC's unassembled examining procedures markedly affect applicants' standings on employment registers. Under the rule of three, applicants' ratings determine whether or not they are considered for employment. If these ratings are unreliable, they will minimize the opportunity for Federal agencies to identify and consider many well qualified candidates.

#### DESCRIPTIONS OF UNASSEMBLED EXAMINATIONS

A CSC examiner evaluates an unassembled examination on the basis of an applicant's education, training, and experience as disclosed on his or her application for Federal employment. The examiner compares this information against written standards for certain types of jobs and rates the applicants eligible or ineligible.

For most jobs, a rating of eligible is accompanied by a numerical score ranging from 70 to 100 points. An additional 5 or 10 veterans' preference points are added, if appropriate. The score determines the applicant's position on an employment register--the higher the score, the higher the position--and thus directly affects the candidate's chances of being considered for a Federal job, since CSC refers names to Federal agencies from the top of the register.

The qualifications standards and rating guidelines CSC rating examiners use are to provide a uniform and objective means of rating applicants. These guidelines include detailed qualification requirements for each job position, and a scale of weights and points for varying kinds and amounts of education, training, and experience.

CSC believes that rating scores assigned in unassembled examinations should be highly reliable. It recognizes the variety of education, training, and experience among applicants and the rating examiners' necessity to use some subjective judgment in assigning scores. Nevertheless, CSC believes that using the same standards for all applicants should insure reasonably consistent ratings.

TESTING THE UNASSEMBLED  
EXAMINATION RATING PROCESS

To test the reliability of CSC's unassembled examination rating process, we selected 351 employment applications from 30 CSC area offices around the country. The applications were chosen from five nationwide examinations: Accountant and Auditor, Biological and Agricultural Sciences, Computer Specialists, Engineer and Scientists, and Nurse. At the time of our test in February 1973, these registers contained the names of about 58,000 eligible applicants in 39 CSC area offices.

Our samples included 205 applications rated eligible at 70 to 110 points, 97 additional rated eligible at 90 points or above, and 49 rated ineligible below 70 points (no score assigned). We randomly selected the 302 eligible applications, and the results are projectible to the five examinations at a 95-percent level of confidence. We could not randomly select the ineligible applications because CSC does not normally retain them. Accordingly, while the test of ineligible applications may indicate a problem, the results are not projectible to all ineligible applications.

We deleted the original rating scores and other identifying information on the applications and sent them for independent rerating to the office from which they were originally chosen and to one other randomly selected office which maintained the same register.

Results of test

The eligible applications which were independently rerated by the originating city deviated an average of about 3.5 points from their original scores. The same applications, when rerated by a second city, varied an average of about 5.5 points from original scores. Sixty-two of the 302 applications originally rated eligible were rerated with scores that deviated 10 or more points from original scores, or were rerated ineligible.

Since applications rated ineligible do not have scores, we could not compare original ratings and reratings. However, 12 of the 49 ineligible applications were rerated eligible.

After our test we gave CSC officials, at their request, several eligible applications from our random samples for their use at a workshop on examining held in June 1973. At the workshop, panels of four to five supervisory staffing specialists from each region and CSC's national office independently rerated applications. The results of the reratings varied from panel to panel and deviated from original scores by an average of 5.9 points.

Effect of differences in ratings

Since the rating score assigned on an unassembled examination generally determines the applicant's standing on a job register, a different score could put the applicant in a different position and affect other applicants' chances for employment.

To determine the extent to which a small difference in point score could affect an applicant's position on a register, we counted the number of eligible applicants on selected registers in February and March 1973. The following schedule shows the results of those counts at 5 CSC offices for applicants with scores of 91 to 100 points.

<u>CSC area office</u>	<u>Grade</u>	<u>Job</u>	<u>Applicants on register</u>	<u>Applicants with scores of</u>				
				<u>100-99</u>	<u>98-97</u>	<u>96-95</u>	<u>94-93</u>	<u>92-91</u>
Boston, Mass.	GS-5	Civil engineer	62	3	8	8	12	22
Salt Lake City, Utah	GS-7	Fisheries biologist	150	30	34	28	8	14
Portland, Ore.	GS-5	Accountant-auditor	247	22	27	34	77	55
San Diego, Calif.	GS-7	Nurse	328	15	4	19	11	23
Mobile, Ala.	GS-7	Computer operator	101	15	8	2	5	1

As can be seen from the table, a difference of even a point or two in a rating could have made a difference in an applicant's standing.

As another test, we calculated the original and the new register position for selected applicants included in our samples. This test showed that applicants' positions could be affected--in some cases by 50 places or more--by a minor variation in score.



## CAUSES OF UNRELIABILITY

To provide CSC with an opportunity to evaluate the results of our test and to determine the causes for the deviations in ratings, we furnished the CSC Seattle region with statistical data on the test results and with copies of several original and rerated applications.

Regional officials analyzed the results and said that incorrect interpretations by raters, inattention to detail, insufficient supervisory review, and other problems, as well as rater judgment, were probable causes for the discrepancies. The region's analysis also disclosed the following needed improvements in rating guidelines:

- Increased explanations or examples of qualifying experience and education.
- Better explanation for awarding supplemental points.
- Better definitions of college majors.
- Clearer descriptions of quality and grade levels.
- More discussion on judging relatedness of experience or education to qualification standards.

## CONCLUSION AND AGENCY ACTIONS

We recognize that unassembled examinations have a certain amount of impreciseness which cannot be eliminated but believe that CSC can improve procedures so that examination scores more consistently reflect the probable qualifications of applicants.

CSC officials agreed on the need for improving unassembled examination procedures. In July 1973 CSC gave us a plan designed to improve both the short- and long-range effectiveness of the unassembled examination system. In his letter to us of March 26, 1974 (app. I), the CSC Executive Director outlined the following actions taken or being taken under the plan:

- Each CSC region reviewed ratings under the five examinations GAO audited, to identify specific causes of inconsistent ratings.

- CSC is paying more attention to reviewing and revising rating schedules. New rating guides for the 5 examinations have been tested in all 10 CSC regions and CSC headquarters. Test results indicate that the new guides will improve the uniformity of rating practices and thereby increase their reliability. Final revisions based on the test results are now being prepared and will be issued shortly.
- CSC officials conducted workshops in January and February 1974 in which examiners from all 65 area offices applied the new guides to actual cases. The workshops also provided training on job analysis and on identifying appropriate qualifications and ranking factors.
- From the review of ratings, CSC regions identified CSC examiners' specific training needs and took immediate steps to improve the general quality and scope of the training given to their staffing specialists.
- Regions have reevaluated and, where necessary, revised their quality assurance plans so that the responsible area office and the staffing division can review ratings regularly; errors are corrected as discovered and individual staffing specialists are counseled on mistakes in rating and misinterpretations of rating guides or qualification standards.

In addition to these measures, CSC is taking further action at its headquarters level to develop an integrated training program for personnel staffing specialists, covering all areas of examining and selection, to (1) improve consistency of approach and help maintain a skilled work force, (2) prepare a standardized sampling procedure to insure statistical adequacy of quality assurance efforts in all regions and (3) emphasize refresher training and continued professional development for journeyman and senior examiners.

In our opinion, the above CSC actions, if adequately implemented and continued, will increase the reliability of its unassembled examination rating process. In our continuing reviews of Federal personnel matters, we plan to evaluate the effectiveness of CSC's actions.

## CHAPTER 3

### MODIFICATIONS NEEDED IN SELECTION SYSTEM

Practical limitations in the art of personnel testing and measurement restrict the degree of accuracy attainable and prevent either the assembled or unassembled examination from being perfectly reliable or valid. Improved evaluation procedures can reduce some of this imprecision, but a large part is irreducible. As a result, the examining process can not accurately rate and rank comparably qualified applicants in exact order of competence.

CSC registers often include many applicants with the same or nearly the same scores. In case of identical scores, names are usually placed on the register, including the top three positions, in alphabetical order or by other means which do not relate to the applicant's job qualifications.

The effect of these and other factors is that, under the rule of three, many applicants who have qualifications equal to others could be denied the opportunity to be considered for employment. In our opinion, this situation could be substantially overcome if the rule of three were replaced with a selection procedure that screened applicants into quality levels or score groups in which all persons are considered comparably qualified. In this respect, present-day examining techniques are accurate enough to insure that a person with a high examination score will generally be more competent than one with a much lower score.

### PRECISION OF TESTING

Examinations are to determine, as accurately as possible, the degree to which an applicant's attributes are related to successful job performance. To do this, the traditional examining system evaluates applicants and gives them an exact rating on a scale of 100 (plus 5 or 10 veteran's preference points, if appropriate); the three highest scoring applicants are then referred to agencies for employment consideration.

Underlying the rule of three is the assumption that examinations are precise, that they can reliably measure the variations or degree of applicants' attributes necessary for job success, and that the three highest candidates are the

three best qualified. However, as discussed below, examinations are not that precise. The degree of precision in today's examinations is such that large differences in scores usually predict real differences in job performance but differences of a point or two are often only chance differences.

Two factors determine the precision of tests--validity and reliability.

### Validity

A valid examination is one which accurately measures whatever it is intended to measure. Personnel selection testing measures potential for successful job performance. A perfectly valid employment test would rank prospective employees in exactly the same order as they would rank after working on the job. However, perfect validity, in practice, is impossible to attain.

Because of the necessity of testing hundreds or thousands of applicants, some merit system examinations rank candidates for broad classifications of jobs, only considering factors common to all positions. They do not necessarily take into account the different factors affecting success in specific positions. Thus they cannot be as precise as a test designed for a particular position.

To reduce the impact of this problem, CSC permits Federal agencies, under certain conditions, to specify additional characteristics affecting job success. CSC then may selectively certify applicants with these characteristics.

Another factor affecting validity is that CSC examinations typically test only for job knowledge, or what a person can do in a given set of circumstances. Job success, however, may depend more on initiative, motivation, or on factors other than capability. In some jobs, such factors as promptness or regular attendance are the most important criteria. In others, success or failure may depend on ability to work under extreme pressure, to adapt quickly to changing conditions, or to work harmoniously with others.

Also, CSC examinations generally are confined to information provided in writing by the applicant (written tests or employment application forms). They do not include (1) interviews with the applicant, (2) contacts with former employers, teachers, or others having knowledge of the applicant's personal characteristics and abilities, or (3) other in-depth evaluations.

Because of these factors, the ranking of applicants on a CSC register may not accurately reflect their potential as successful employees.

### Reliability

If a test is highly reliable, a person taking it at two different times would make the same score.

Perfect reliability is also, in practice, unattainable. Factors which affect reliability include the length, difficulty, construction, and understandability of the test. Other factors which limit examination reliability include the mental and physical condition of the candidate at the time of testing, the physical conditions of the test site, the homogeneity of the sample, the "test-wiseness" of the candidate, and even the element of luck in guessing.

We contacted a number of personnel experts and psychometricians in consulting firms, State and local governments, and private industry to determine if the precision of testing is limited. Although there was divided opinion as to whether validity or reliability was the biggest problem in predicting job success, they agreed that there is a practical limit to precision. Many felt that a 5-percent error factor was the minimum that could be expected for reliability, and some felt that even 10 percent or higher was the practical minimum. They also indicated that written tests are generally more precise than unassembled examinations because of the greater subjectivity of unassembled examinations.

Our review of a wide range of literature on personnel testing and measurement confirmed these judgments and disclosed that most tests, including CSC's, are not reliable enough to distinguish between examinees with reasonably similar scores.

It is therefore generally recognized that, because of these limitations on validity and reliability, the examination process is not precise enough to judge the potential job success of persons with reasonably similar scores. Personnel selection techniques are accurate enough to insure that a person with a high examination score is generally more competent than one with a much lower score. Small differences in scores, however, often do not predict any real differences in competence.

### THE RULE OF THREE

The origin of the rule of three is rooted in regulations which predate the Civil Service Act of 1883 but which were not incorporated into that act. Although traditionally accepted as an important element of the merit system, the rule of three was not given statutory authority until it was incorporated into the Veteran's Preference Act of 1944 (ch. 287, 58 Stat. 389).

The rule of three was aimed at eliminating political patronage, favoritism, and other abusive elements of the spoils systems to make selection a matter of merit and competency only. It was based on the assumption that civil service tests could make sufficiently accurate distinctions among individuals to reflect different levels of competence.

The assumption is statistically unsound because of the practical limitations on attaining precision in testing. At best, according to employment authorities, the rule of three provides a fast means of distinguishing among candidates and, if applied without abuses, can reduce outside influence. However, it can also prevent persons well qualified for specific jobs from being considered for employment.

As mentioned earlier, selective certification is one procedure which can increase the probability that well qualified applicants will be among the top three names. CSC permits this procedure when the selecting agency can show that requirements not included in the applicable qualifications standards are essential to satisfactorily perform in a particular position. Under selective certification, the three persons certified do not necessarily stand at the top of their register, but are instead the three highest scoring applicants with the special or selective requirements of the

position as specified by the hiring agency. Although this procedure reduces the number of examination categories needed and permits specific job categories to be recognized and certified, it still considers only three persons from a list of eligibles.

Another problem caused by the rule of three is the manner in which tie scores are handled. Since most CSC registers have hundreds or thousands of names which fall within the relatively narrow range of 70 to 110 points (a maximum of 10 veterans' preference points), tie scores occur frequently. Because only three names can be certified for a vacancy, tie scores must be arbitrarily resolved.

The law (5 U.S.C. 3313) requires that ties first be decided in favor of preference eligibles. CSC instructions (FPM Supplement 332-71, May 1970) state that, if ties persist, they should be decided in favor of:

- The person with the highest score on the most heavily weighted part of the examination, if any.
- Federal employees over non-Federal employees.
- Employees of the agency to which the certificate of eligibility is issued over other Federal employees.
- Ex-employees of the agency over other applicants.

As a tie-breaking device of last resort, names are listed alphabetically, in reversed order from Z to A, or randomly. Since it is not uncommon for a high proportion of scores on registers to be tied under the rule of three, ranking of eligibles is often determined by procedures unrelated to competence for the job.

Examples of the number of tie scores which can occur in CSC registers are shown on page 7 and in the following listings from the Seattle area office.

Number of Eligibles on Selected Registers  
April 24, 1974

<u>Rating score</u>	Card Punch Operator <u>GS-2</u>	Technical Assistant <u>GS-4</u>	Air Traffic Controller <u>GS-7</u>
105	-	-	2
104	1	-	6
103	-	1	6
102	-	9	7
101	-	47	13
100	8	9	24
99	2	8	22
98	3	26	26
97	5	16	31
96	4	102	21
95	8	63	22
94	4	23	25
	<u>35</u>	<u>354</u>	<u>205</u>
Subtotal			
	<u>154</u>	<u>589</u>	<u>587</u>
Total on register			

For each of the above registers, a list containing the top three names from which an agency could choose to fill a vacancy would exclude at least five persons who received the same score on the examination as one or more of those certified. In fact, the agency with the vacancy would have no opportunity to examine the applications of, or interview, any of those excluded individuals to see whether they had special characteristics or qualifications which might be useful on the job. In view of the degree of imprecision in ranking candidates, it seems unreasonable to consider the three top scoring candidates as the best qualified candidates in all cases. This problem is compounded in the case of tie scores.

Federal agency managers are responsible for conducting operations and achieving the missions and goals prescribed for their agencies in the most efficient, effective, and economical manner. To do this, they are authorized to formulate personnel and financial budgets which provide necessary resources. Within the area of personnel administration,



Federal executives are responsible for determining the size and mix of personnel needs, the duties which personnel will perform, and the means of filling positions. Except for the limited discretion permitted by the rule of three, agency managers do not have the authority or flexibility to fill vacancies by hiring from CSC registers those persons whom they find are the most qualified available for a particular job. Further, many agency managers feel that many jobs, even though in the same general category as those of other agencies, such as a clerk or accountant, have such different requirements that only the agency can select specific people who meet the detailed needs.

Although several agency officials felt that the rule of three was appropriate and said they were satisfied with the applicants from the top of the registers, personnel management and operating officials of 17 of 21 agencies expressed some reservations about it. Many felt that the rule of three was too restrictive and has hindered them in hiring applicants they considered to be the highest quality available. Others referred to the time-consuming "techniques" and "games" they sometimes had to use to obtain candidates (other than the top three) whom they felt were highly qualified. Additional comments were directed at the methods by which tie scores are broken and the imprecision of the examination process.

Several agency officials feel that the examination system does not take into account some aspects of human behavior which may be important to the job, nor does it always distinguish between jobs in a job category. Eight agencies advocated changes to give agencies more choice in the final selection of applicants than provided by the rule of three.

## ALTERNATIVE SELECTION SYSTEMS

### Category-ranking system

For many years other Federal agencies which are exempt from the Civil Service Act have used a selection procedure that does not involve the rule of three. Two of these agencies are the Atomic Energy Commission and the Tennessee Valley Authority. Under their employment systems, applicants are rated and grouped into broad categories, such as "outstanding," "well qualified," "qualified," and "unqualified." Selections are made from the names in the highest category until the names in that group have been exhausted. Veterans are considered ahead of nonveterans within each category. This system generally distinguishes between persons' capabilities; it groups comparably qualified candidates together but leaves the final ranking to the manager whose job vacancy is being filled. It does not assume a degree of precision in testing and ranking which is impossible to attain.

Although the category-ranking system differs in concept and practice from the rule of three, CSC has found it to comply fully with merit principles. As a result CSC and these agencies have signed interchange agreements which permit employees to change from one system to another without losing status.

Two more recent selection systems that have been based on the need for increased flexibility and for more objective, statistically sound procedures are the "rule of ranks" and the "rule of reliability."

### Rule of ranks

A rank is defined as a specific rating score, such as 100, 99, 98, etc. Some States and local governments have selection systems whereby a set number of ranks, rather than individuals, are certified to an employing agency. For example, California has switched from a rule of three to a rule of three ranks for many of its job categories. Under the rule of three ranks, all persons scoring in the top three scores are certified. As a result, the average number of applicants from which an agency can select is increased and the need for arbitrarily resolving tie scores is eliminated.

## Rule of reliability

Under this rule, the number of applicants to be certified for a given job is based on the examination's statistical reliability. On May 1, 1973, after a lengthy study and improvement of the reliability of its examinations, the Michigan Department of Civil Service adopted a rule of reliability for its certification system. Using statistical formulas to calculate reliability coefficients, Michigan determined that the reliability of its examinations ranged from .96 to .75. (1.00 is 100 percent reliability.) For each examination a bandwidth, or number of ranks, is calculated; an agency requesting a list of eligibles to fill a vacancy is given the names within the computed bandwidth. For example, for the position of storekeeper the test reliability was determined to be .86 and the bandwidth was calculated as four ranks (i.e., 100-97). Thus, all applicants scoring from 100 to 97 were considered statistically comparable and all were equally eligible for a job opening. For the position of computer programmer, the test reliability was determined to be .78 and the bandwidth was calculated as seven ranks (i.e., 100-94).

These two examples show that the more reliable the examination, the narrower the bandwidth and the fewer the number of certified applicants from that register for a given opening.

## Other systems

Several State agencies are using other systems more liberal than the rule of three. In fact, available data indicated that among all States, less than one-third are using a rule of three or a procedure more restrictive.

## AGENCY COMMENTS

In informally discussing this matter with CSC officials, they stated that the objective of identifying the best qualified applicants is tempered in the Federal competitive service by a number of public-interest considerations contained in law, such as veterans' preference, apportionment of positions in the Washington, D.C., area, members-of-family restrictions, and employment-of-relatives (nepotism) prohibitions--all of which may at times prevent appointment of the best qualified applicant. No less important, according to CSC, is the public-interest requirement for minimizing the opportunities for

personal and political favoritism and other arbitrary selection factors that are antimerit. To fulfill this requirement CSC believes that it is vital to keep the selection base as small as possible and consistent with the needs of the selecting agency for reasonable flexibility, even if a larger base might at times permit the agency to select a better qualified eligible. CSC agrees that the agency must have a choice, but it believes that the more candidates the agency can choose from, the greater the chances are that non-merit factors will be applied, the less the agency is protected from outside pressures to appoint favored candidates, and the greater are the strains on public confidence in the system.

Therefore, we believe CSC needs a selection system which (1) more appropriately considers the accuracy of employment examinations, (2) allows the selecting agency reasonable flexibility in identifying and hiring the best available applicants, and (3) meets the need for merit principle safeguards. Such a system need not affect the preferential rights of veterans.

The Executive Director, CSC (app. I), stated that one of the reasons CSC has not actively continued to advocate changes in the rule of three since the Hoover Commission reports (see page 3), is because the examining system has become flexible through changes, such as broad-band examining, open continuous announcements, improved qualification requirements and single-agency standards, deferred ratings, and selective certification. The Executive Director stated, however:

"While we believe the rule of three has served the country well, we recognize that other jurisdictions have different provisions for referral of names of eligibles and that this issue is an important policy question that deserves to be aired and considered."

## CONCLUSIONS

Even with improvements in the reliability of CSC's un-assembled examinations, practical limitations in employment testing and measurement will continue to prevent examinations from being perfect predictors of job performance. Our review

showed that, even under the best of conditions, the degree of precision attainable in testing is not nearly as high as the degree of test precision required by the rule of three. As a result the rule of three unfairly denies to many applicants who have equal qualifications the opportunity to be considered for Federal employment.

CSC's examining process can, if improved, reliably screen applicants into quality levels or score groups in which all persons are considered comparably qualified. Federal agencies should, within practical limits, be permitted to consider and fill vacancies from among all applicants in the best qualified level or group available.

#### Matters for consideration by the Congress

In our opinion, the rule of three selection requirement is unrealistically rigid. We recommend that the Congress amend the requirement and allow CSC to prescribe alternative selection procedures, similar to those discussed in this report. CSC should design the procedures in such a manner that they will provide

- applicants with a reasonable opportunity to be considered for employment,
- Federal agency management greater opportunities to identify and select well qualified applicants,
- safeguards needed to insure compliance with merit principles.

Each agency should then be permitted to choose the selection procedures best suited to its particular needs--just as the Tennessee Valley Authority and the Atomic Energy Commission have done. The individual agency's application of the procedures should be subject to CSC's continuing review and evaluation to assure compliance with CSC guidelines and policies.

Any change in the rule of three selection requirement need not affect the preferential rights of veterans.

## CHAPTER 1

### SCOPE OF REVIEW

Our review of CSC's policies and procedures for examination and selection was conducted at CSC headquarters in Washington, D.C., and at the regional and area offices in Seattle, Washington. We conducted parts of our review at CSC's Portland, San Francisco, San Diego, Salt Lake City, Mobile, Boston, and Washington, D.C., area offices. At these locations we reviewed Federal employment applications, employment registers, certification forms, and other records pertaining to examination and selection and interviewed operating and management personnel.

We also interviewed and obtained data from:

- Personnel management and operating officials of 20 other major departments and agencies in the competitive service, at both Washington, D.C., and field locations.
- Personnel management officials of the Atomic Energy Commission and the Tennessee Valley Authority, two agencies not in the competitive service.
- Officials of the National Civil Service League, the Veterans of Foreign Wars, the American Legion, the National Federation of Federal Employees, and Ralph Nader's Public Interest Research Group.

We reviewed literature on test development, test validation, and test measurement and interviewed a number of nationally recognized authorities in the field of personnel evaluation and testing. In addition, we sent a questionnaire on the precision of testing and the adequacy of various selection systems, including the rule of three, to 40 State and local government personnel directors, 24 personnel representatives of private industry organizations, and 12 personnel and management consulting firms. Thirty-six replies were received. We also conducted a statistically valid random sample of CSC unassembled examination ratings to evaluate their reliability.



## UNITED STATES CIVIL SERVICE COMMISSION

IN REPLY PLEASE REFER TO

WASHINGTON, D.C. 20415

MAR 26 1974

YOUR REFERENCE

Mr. Forrest R. Browne  
 Director, Federal Personnel and  
 Compensation Division  
 U.S. General Accounting Office  
 Washington, D.C. 20548

Dear Mr. Browne:

Thank you for the opportunity to comment on the General Accounting Office's draft report, "Improvements Needed in Examining and Selecting Applicants for Federal Employment." We do want to respond to certain of your findings and especially to report on steps taken or planned to improve the consistency of ratings in examinations in which experience and education alone are the basis of the rating.

When first informed of your findings, as your report accurately reflects, we identified the need for certain measures to improve the consistency of ratings under the five examinations covered in your review. We also reevaluated the types and levels of training and quality assurance efforts at various organizational levels (area office, region, central office). Based on this review we have, to date, taken the following actions:

1. An extended review of ratings under the five examinations audited by GAO was made by each Commission region to identify specific rating problems and causes of inconsistent ratings. Regional plans were then developed to correct the identified causes of rating variances through consultation with appropriate staffing specialists and supervisors. (This was a special effort in addition to those rating reviews which were a normal aspect of periodic evaluations.)
2. We are giving increased attention to review and revision of rating schedules. Tentative changes were developed in the rating guides for the five examinations in response to the particular problems identified in the regional reviews. We tested the new guides in all 10 CSC regions and Washington, D.C. Test results indicate that the new guides will significantly improve uniformity in rating practices and thereby increase the reliability of ratings. Final revisions based on the test results are now being prepared and will be issued shortly.

THE MERIT SYSTEM—A GOOD INVESTMENT IN GOOD GOVERNMENT

## APPENDIX I

3. The Bureau of Recruiting and Examining (BRE) conducted workshops in January and February 1974 on the five revised rating guides in which examiners from all 65 area offices applied the new guides to actual cases. The workshops also provided training, through the case-study approach, on job analysis and the identification of appropriate qualifications and ranking factors processes which are essential in referral of the best qualified candidates under examinations in which deferred rating techniques are used or selective certification is appropriate. This training supplemented ongoing, and special, training efforts at the regional and local levels.
4. From the review of ratings, CSC regions identified specific training needs of CSC examiners and took immediate steps to improve the general quality and scope of the training given to their staffing specialists. These actions included ensuring that reference materials are current and complete, reviewing ratings through systematic sampling to identify training needs on a continuing basis, developing needed training materials, and ensuring that new examiners receive thorough training in the application of rating guides.
5. Regions have reevaluated, and where necessary, revised their quality assurance plans to provide regular review of ratings by both the responsible area office and the staffing division; errors are corrected as discovered and individual staffing specialists are counseled on mistakes in rating and misinterpretations of rating guides or qualification standards. Suggestions or problems that cannot be handled at the regional level are referred to BRE.

In addition to these measures, we are taking further action at the bureau level:

1. We have had a variety of training programs which covered specific aspects or types of examining. We are now developing an integrated program for personnel staffing specialists that covers all areas of examining and selection. The program will consist of practice units developed from real life situations involving rating guides, name cases, certification, job-relatedness determinations and documentation, special placement efforts, staffing assistance, etc. The training should improve our consistency of approach and help maintain a skilled work force of staffing specialists.
2. The extent of review of ratings has varied among examining offices depending to a large extent on the experience level of the examiners. We are preparing a standardized sampling procedure to assure statistical adequacy of quality assurance efforts in all regions. We will issue guidelines covering goals, minimum standards of adequacy, sampling and review techniques, and the responsibilities of the bureau, regions, and area offices.



3. Much of our past training efforts have been devoted to bringing trainees to the journeyman level. We plan to place equal emphasis on "refresher" training and continued professional development of journeymen and senior examiners.

Your report recognizes our concern with your recommendation that Congress replace the rule of three. We feel the most productive way to improve the reliability and validity of the examining program is to work on the rating process itself. While we believe the rule of three has served the country well, we recognize that other jurisdictions have different provisions for referral of names of eligibles and that this issue is an important policy question that deserves to be aired and considered. Of course, any significant modification of the rule of three inevitably involves a consideration of the nature and extent of veterans preference, which has been longstanding public policy and about which the Congress has expressed few, if any, doubts.

One of the reasons the Civil Service Commission has not actively continued to advocate changes in the rule of three since the Hoover Commission reports is that we have achieved a great deal of flexibility through changes in the examining system itself with such improvements as broad-band examining, open continuous announcements, improved qualification requirements and single-agency standards, deferred rating, and selective certification--all of which have resulted in a much better match of person to job.

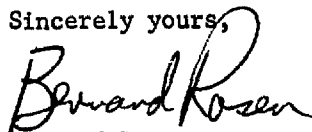
[See GAO note.]

APPENDIX I

[See GAO note.]

We would be happy to discuss further any of these comments with you.

Sincerely yours,

A handwritten signature in cursive script that reads "Bernard Rosen". The signature is written in dark ink and is positioned above the printed name.

Bernard Rosen  
Executive Director

GAO note: Deleted material suggested minor changes to the report. We have considered these changes in this final report.

Copies of this report are available at a cost of \$1 from the U.S. General Accounting Office, Room 4522, 441 G Street, N.W., Washington, D.C. 20548. Orders should be accompanied by a check or money order. Please do not send cash.

When ordering a GAO report please use the B-Number, Date and Title, if available, to expedite filling your order.

Copies of GAO reports are provided without charge to Members of Congress, congressional committee staff members, Government officials, news media, college libraries, faculty members and students.

AN EQUAL OPPORTUNITY EMPLOYER

UNITED STATES  
GENERAL ACCOUNTING OFFICE  
WASHINGTON, D.C. 20548

OFFICIAL BUSINESS  
PENALTY FOR PRIVATE USE, \$300

POSTAGE AND FEES PAID  
U. S. GENERAL ACCOUNTING OFFICE



THIRD CLASS

Mr. T. E. Sullivan  
Transportation  
Room 5033