

11632 110946

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



110546

FOR RELEASE ON DELIVERY
EXPECTED AT 9:30 a.m. EDT
Wednesday, October 10, 1979

STATEMENT OF

CLIFFORD I. GOULD, DEPUTY DIRECTOR
FEDERAL PERSONNEL AND COMPENSATION DIVISION

BEFORE THE

SUBCOMMITTEE ON ELEMENTARY, SECONDARY, AND
VOCATIONAL EDUCATION
HOUSE COMMITTEE ON EDUCATION AND LABOR

ON

HSE01102

[EXTENDING H.R. 3564 and H.R. 4949
TO FEDERAL EXAMINING]

We appreciate the opportunity to appear before you today to discuss our recent report on Federal employment examining and to comment on the potential impact of extending concepts contained in H.R. 3564, the Truth in Testing Act of 1979, and H.R. 4949, the Educational Testing Act of 1979, to tests administered by the Federal government. You asked that we focus on our report, "Federal Employment Examinations: Do They Achieve Equal Opportunity and Merit Principle Goals."

Our report, released on May 15th of this year, highlighted problems confronting those responsible for ensuring that the Federal work force is competent, productive, and

007211

to the extent possible, reflective of all segments of our Nation's society. The goal of Federal employment programs is to hire qualified individuals based on merit principles giving everyone an equal opportunity to compete for a job. To achieve this goal the Civil Service Commission, the predecessor agency to the Office of Personnel Management, designed various examining procedures to help assure that the most competent and productive people are employed.

We found that black job applicants were screened out of the competition at a much higher rate than white applicants on two major written tests used by OPM to develop hiring registers. The implications of this are presented in detail in our report where we recommended that OPM further investigate the use of valid alternatives to written tests.

The written test for the Professional and Administrative Career Examination (PACE) is the most thoroughly researched test in the history of Federal civil service examining. According to OPM the test has been fully validated and its research shows a clear relationship between performance on the test and subsequent performance on the job. OPM says that not only does the test result in hiring individuals who perform better on the job, but its use enhances overall Government productivity.

The PACE is particularly important because it is the entry route into Federal service for more college graduates

than any other single method and because the type of jobs covered by the test often leads to a higher level career position. For that reason, it is important that the Subcommittee explore the potential impact of the bills it is considering on this examination.

I would like, first, to put PACE in perspective in terms of its place in obtaining a Federal job. PACE is used to hire new employees for entry level jobs in 118 different occupations. According to OPM, the most common feature of PACE occupations is their information burden--that is, all of the occupations require individuals who can read, understand and interpret a large body of written laws and regulations or a large body of knowledge related to a specific profession or specialty area. OPM determined that five abilities were important to successful performance in these occupations and could be measured with a written test. These were: inductive and deductive reasoning, verbal comprehension, judgment, and quantitative reasoning.

Individuals who pass the test and receive a rating of 70 or more are listed on a PACE job register according to their final rating--test score plus points for Veterans' preference or for being an outstanding scholar, when applicable. When entry level jobs become available in one of the 118 occupations, agencies request OPM to send them a list of names from the PACE register. In response, OPM

sends the agency the three names with the highest ratings. Agencies generally must select one of the three individuals on the list, and if one of the individuals on the list is an eligible Veteran, the Veteran must be selected. The process is somewhat more complex than what I have just described and substantial changes are contemplated as a part of the authority delegations under the Civil Service Reform Act. But this should help explain what has been required of an individual seeking an entry level PACE position and who has not had a Government job.

During fiscal year 1978, OPM processed 1.6 million job applications for all types of competitive civil service jobs, and about 152,000 people were hired. These numbers include all competitive jobs, both wage grade and general schedule. During that same year 135,000 individuals took the PACE, 77,000 passed with a score of 70 or above, and 7,600 were hired from PACE registers.

COMMENTS ON EXTENDING CONCEPTS
IN H.R. 3564 AND H.R. 4949 TO
FEDERAL EXAMINING

I would like to begin my comments on H.R. 3564 and H.R. 4949 with a caveat. In the time available to prepare this testimony, we were unable to perform a detailed legal analysis of these two bills. We were also unable to audit cost and workload estimates provided to us by OPM. We can provide additional information for the record at a later

date or the Subcommittee may want to obtain that information directly.

While we support the intent of both H.R. 3564 and H.R. 4949 with respect to improving the quality and availability of information about tests, we do have reservations about the potential cost and feasibility of extending certain provisions to Federal examining.

The Federal government may well be at the forefront in providing information and preparatory assistance to those taking standardized written tests. For instance, in 1971 it published a 60-page booklet designed to provide teachers, students, and other interested persons with samples of the types of items used in government tests for filling vacancies in a wide variety of office occupations. This booklet has undergone five revisions over the years and each time it has been revised OPM produced 10,000 copies for distribution to those who might be interested. In addition, it has been available through the Superintendent of Documents.

Again, in 1974 the government published a 12-page booklet containing samples of the types of items used in the PACE written test. This booklet was attached to the basic PACE qualifications announcement in 1975 and has been distributed to every applicant since then.

The most important reservations we have about these bills relate to Section 5 of H.R. 4949, requiring the public disclosure

of all test items, and to that part of Section 6 in H.R. 3564 which prohibits the assignment of scores to knowledge and achievement tests on the basis of rank order.

We agree with the various arguments which have been made to the effect that the Federal government would experience a substantial economic burden if it were required to provide copies of test questions and correct answers to examinees. Such a procedure would preclude reuse of any items. So new ones would have to be prepared each time a test is given. This has enormous cost and technical implications.

The cost implications are illustrated in an example prepared by OPM. The government's clerical exam is administered nationwide every working day. If test items were disclosed, 250 separate and unique exam booklets would be needed each year. It takes about one staff year, at an average salary of \$20,000, to prepare a clerical exam booklet using existing questions. OPM estimated that at least three staff years would be required to prepare a booklet using totally new questions. This does not include additional administrative costs required to insure that the same test is given nationwide on each day so that applicants in one area do not benefit from disclosure in another area.

The administrative problems, budget, and staff requirements are multiplied many times when more complicated,

higher level exams are used. To develop the current pool of items for the PACE, for instance, required a total of 52 staff years of effort over a 5-year period.

It should be noted that the cost and workload estimates I have just provided are for test development only. They do not cover the very time-consuming and costly process of validating the test once it is developed.

The technical implications are no less severe. In the first place, there is a finite limit to the number of new questions that can be developed in any subject matter area without quality deteriorating to a level that the exam would no longer be an effective means of evaluating people for employment. In the second place, a group of common questions is needed in each version of a test so that adjustments can be made for differences among the versions in average difficulty. Only in this way can a score have the same meaning on one version as it has on another. Disclosure would make it impossible to have a common group of questions. Therefore different versions of the same test could not be equated and changes over time in test performance--such as the decline in SAT scores--could not be identified.

The concern we have with H.R. 3564 is with the distinction made in Section 6(c) between aptitude test and achievement test and the prohibition against grading the

latter on the basis of the relative distribution of scores of other test subjects.

The distinction between aptitude and achievement tests is a point of great controversy even among experts in the field of testing and assessment. Sometimes the distinction is made on the basis of the way the test scores are used: if they are used to assess someone's current level of performance, the test might be considered an achievement test; if the scores are used to predict how well someone might perform in the future, the test might be called an aptitude test. By this distinction, the clerical exam I mentioned earlier would seem to be an achievement test, while the PACE would seem to be an aptitude test.

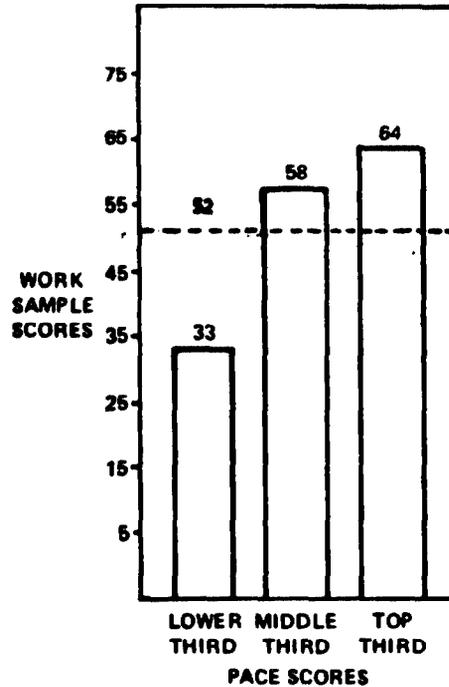
Whether or not this type of distinction is made, we would be concerned about the extension of a prohibition such as described in Section 6(c) to Federal examining. We would be particularly concerned about extending the prohibition to aptitude tests like PACE; but we would also be concerned about extending it to achievement tests like the clerical exam where rank order on the exam may be related to relative performance on the job. The reason for our concern can be seen in data on the value of PACE reported in our study.

As soon as the PACE written test was developed, OPM started research designed to determine its value in terms

of the extent to which scores on it were related to job-performance in a number of occupations. The first study was conducted on social security claims examiners. The written portion of PACE was given to 250 examiners and measures of their work effectiveness were collected. These measures were specially designed for the study. They included a work sample consisting of a standardized claim to be adjudicated by the examiner and scored for correctness.

OPM's analysis of the relationships between test scores and the various job performance measures showed that those who scored higher on the test tended to be the ones who also scored higher on the measures of job performance. The graph on the following page gives some indication of these relationships. It shows that those who scored in the top third on PACE did almost twice as well on the work sample (average score of 64) as those who scored in the bottom third on PACE (average score of 33).

**Average Work Sample Scores at
Each Range of PACE Scores**



What the chart shows is that the average Work Sample Score for all three ranges of PACE scores combined is 52. That is the score you would expect a large group of people to obtain on average if they were selected on some basis unrelated to their relative performance on PACE. However, there is a relationship between performance on PACE and performance on the Work Sample. So if you want to maximize productivity you would select only those who score in the highest third on PACE. Their Work Sample scores average

out to 64--an improvement of 12 points over selection among a group which has not been rank-ordered.

Another way to look at this is in terms of how much productivity is increased by selecting on the basis of relative rank order as opposed to selecting on some other basis. OPM researchers have constructed economic models which show that for PACE the increase in productivity is over twice as high for selection done in a strict rank-order based on test scores as it is for selection done on a random-order basis from among those who pass the test.

In addition to these productivity considerations, it should be noted that Veteran's Preference legislation does require a rank ordering of applicants.

Mr. Chairman and Members of the Subcommittee, I would like to conclude by repeating that we support the intent of these bills to improve and make public the processes and information so important to enjoying an equal chance for achievement in our society. The hearings of this subcommittee and the legislation it is considering have added materially to public scrutiny and accountability in this area. However, with respect to those provisions about which we have some concern, we believe that their enactment could bring about the end of testing in the Federal government.

This completes my prepared statement. We will be happy to answer any questions you may have.