

1974 WL 222
United States District Court, N.D. Ohio, Eastern
Division.

Shield Club et al., Plaintiffs
v.
City of Cleveland et al., Defendants.

Civil Action No. C 72-1088 | July 6, 1974

THOMAS, J.

Memorandum and Order

*1 Plaintiff The Shield Club, an organization composed principally of black Cleveland police officers, acting on behalf of all of its members who are past or present officers of the Cleveland Police Department (CPD), individual plaintiffs who have taken promotional examinations for positions above the rank of patrolman, individual plaintiffs who took the 1972 entrance examination for patrolman, and other persons who claim an interest, bring this action pursuant to 42 U.S.C. §§ 1981 and 1983. Alleging violation and deprivation of their constitutional rights to equal protection of the laws, the complaint challenges “a broad range of practices used by the officials in the recruitment, testing, screening and hiring of new patrolmen, and in the assignment, treatment and promotion of current police officers.” The alleged jurisdiction exists in this court under 28 U.S.C. § 1343(3) and (4).

By its memorandum and order of December 21, 1972 [5 EPD P 8406] and by supplemental order [5 EPD P 8527], this court directed that 18% of the 194 new patrolmen to be hired by the City defendants under a federal Impact Cities grant be blacks or Hispanics (hereafter minorities) who had passed the 1972 Civil Service entrance examination for patrolmen. This order followed a trial at which the single issue of the validity of the 1972 entrance examination was decided adversely to the City defendants. The prima facie invalidity of the examination inferred from the racially disproportionate impact of the examination on blacks and Hispanics who took the exam was not overcome by any showing by the City defendants that the examination was job related.

At a hearing that began April 16, 1974, the remaining issues of the case were tried. Following oral argument on June 24, 1974, this court issued an oral provisional order that directs the Cleveland Civil Service Commission (CSC) and the CPD to proceed to devise new screening procedures to be used hereafter in the appointment of new patrolmen. That provisional order has since been amplified and is filed as a separate written order of this court. This memorandum and order relates to a different issue—alleged discrimination in promotions to positions above the rank of patrolman.

[Allegations of Complaint]

I. As generally stated in the complaint plaintiffs allege:

“The defendants and their agents and employees have discriminated and are continuing to discriminate on the basis of race, color, and national origin against black and Hispanic policemen within the Police Department. At the present time there are only eight (8) black officers above the rank of patrolman (6 men and 2 women). These are all sergeants with the exception of one black male lieutenant. There are no blacks above the rank of lieutenant in the Cleveland Police Department. There are no Hispanics above the rank of patrolman.”

The City defendants and Richard J. Faragher, intervenor and new party defendant (appearing as president and representative of the members of the Fraternal Order of Police), do not contest the assertion of the plaintiffs that the only black and Hispanic police officers above the rank of patrolman are “six men and two women.” Each of the defendants, however, denies in their respective answer all of plaintiffs’ allegations of discrimination in the complaint.

*2 Title VII of the Civil Rights Act of 1964, was extended to public employers, 42 U.S.C. §§ 2000e-2000e-17, effective March 24, 1972. Plaintiffs do not base their case on Title VII, relying instead on sections 1981 and 1983, as seen. However, decisions rendered in Title VII cases, such as *Griggs v. Duke Power Co.*, [3 EPD P 8137] 401 U.S. 424 (1970) apply equally to actions brought under the century old Civil Rights Acts, 42 U.S.C. §§ 1981 and 1983. In its opinion of December 21, 1972, this court noted that “courts have been quick to apply *Griggs* to public employment cases involving policemen, firemen, and teachers.” See *Castro v. Beecher*, [4 EPD P 7783] 459 F.2d 725, 732 (1 Cir. 1972), and the recent case of *Vulcan Society of New York, City Fire Department, Inc. v. Civil*

Service Comm'n, [6 EPD P 8974] 490 F.2d 387 (2 Cir. 1973).

Particularizing their claim of discrimination in promotions, plaintiffs identify a series of practices of the City defendants, including but not limited to

“(a) the use of a written test, as a prerequisite for promotion, which excludes a disproportionately high number of black applicants for promotion, relative to white applicants. Only eight blacks hold a position above the rank of patrolman, the beginning rank in the police department. The tests have not been professionally developed or shown by validation studies or otherwise to predict with reasonable accuracy job performance or job relevance.

“(b) the awarding of extra points based upon seniority, said practice having the effect of reinforcing the impact of past racially discriminatory hiring practices.

“(c) the discriminatory assignment of black and Hispanic officers to positions within the Department which afford little opportunity to study for and gain experience relevant to the promotional examination.”

[Test Performances]

A. Plaintiffs prepared and submitted exhibits, the accuracy of which is unchallenged, that reflect the racial composition of (1) persons taking the 1972 promotional examination for sergeant and (2) sergeant appointments from February 14, 1973 to January, 1974.¹

*3 On October 6, 1972, a promotional examination for the position of sergeant was given. Of the 695 persons taking the examination, 70 (10.1%) were minorities (used herein interchangeably with blacks and Hispanics) and 625 (89.9%) were white. Twenty-nine (41.4%) of the minorities and 332 (53.1%) of the whites passed the examination.² White patrolmen had an average score of 70.07 on the examination and minority patrolmen averaged 67.04.

Computations indicate that during the two-year life³ of the promotional list prepared on the basis of the examination score supplemented by seniority points, a minimum raw score of 79 would be required for a person actually to be eligible for promotion. Three of 70 (4.3%) minorities and 115 of 625 (18.7%) whites achieved a raw score of 79 or higher. As of January, 1974 20 of 625 (3.2%) of the whites taking the examination, or 6.0% of the 325 whites

passing the examination, were promoted. One black, which is 1.4% of the 29 minorities passing the examination, was promoted.⁴

When asked about the statistics concerning the performance of whites and minorities on the 1972 Sergeant's Promotional Examination, Dr. Norman Henderson, Professor and Chairman of the Department of Psychology,⁵ testified initially on the difference between the mean examination scores of minorities (67.04) and the mean examination scores of whites (70.07):

*4 “Yes. There is a mean difference there of—a mean difference of slightly less than 3 points [actually 3.03 points] and it is statistically significant. In other words, that has a reliability result, and is not due to chance fluctuation. . . . It is a T value of 2.55 which is significant at the 1% level, which means that the odds that that would have occurred by chance is less than one in 100. I'll point out immediately that that mean difference relative to the variation within groups is not particularly large. There is a Z score difference of .33, and one of the reasons we obtain a significant difference with only a 3 point spread is that the samples are large and the (statistical) methods are powerful and can detect small absolute differences.”

Plaintiffs' counsel then asked Dr. Henderson for calculations concerning the statistical difference in the performance between minorities and whites in that instance.

“A. What I did was to analyze a simple four-fold [2 x 2] table on the number of individuals above or below the 79 cutting point as indicated for blacks and whites. One would predict, for example, that among minority applicants, if there were no differences, one would expect 12 to have scored 79 and above, rounding off to a total number; whereas only 3 did, and differences of course in the rest of the table as well; and with the expectation of 12 scoring over 79, you would expect 58 below; whereas the split was 3 and 67, and this discrepancy is significant. Using the Chi-square test, the likelihood of it occurring by chance is quite low. . . .

* * *

“ . . . It is clearly better than 50 to 1. . . .

* * *

“ . . . In other words, it is highly unlikely that if the tests were producing similar scores for blacks and whites, that you would get only three blacks scoring above 79. One would expect 12.”

Thus, Dr. Henderson testified that statistically significant differences are shown. With reference to the test results of the 1972 Sergeant's Promotional Examination, he said “one would expect 12 [minorities] to have scored 79 and above [the cutting point below which raw score no promotion is likely to be made] . . . whereas only 3 did;” and the mean differences in scores were 3.03 points more for whites than for minorities. Another marked difference is that 29 of 70 minorities taking the test (41.4%) passed the examination, while 332 of 625 whites who took the exam (53.1%) passed. To date, of the 26 patrolmen promoted to sergeant from the 1972 eligibility list (based on the 1972 examination) only one is black. It is concluded and determined that these statistically significant differences between the passing and promotion figures of minorities and whites create a prima facie case that the 1972 Sergeant's Promotional Examination has, in the words of Judge Coffin, “a racially disproportionate impact.” When this is proved,

“The public employer must, we think, in order to justify the use of a means of selection shown to have a racially disproportionate impact, demonstrate that the means is in fact substantially related to job performance. *Castro*, *supra* at 732.”

*5 After approving this language, Judge Friendly elaborates in *Vulcan*, *supra* at 393:

“ . . . A showing of a racially disproportionate impact puts on the municipal or state defendants not simply a burden of going forward but a burden of persuasion. . . . But if the public employer succeeds in convincing the court that the examination was “substantially related to job performance,” an injunction should not issue simply because he has not proved this to the hilt.”

[Job Relatedness of Tests]

Before reviewing the evidence it is essential to consider the means by which the public employer (the City defendants) may sustain its burden of persuasion that the 1972 Sergeant's Promotional Examination was “substantially related to job performance.” Under recognized and accepted standards of educational and psychological testing, prevailing at least as early as 1972, employment examinations, whether for entrance or

promotion, may be shown to be job related by the procedures of criteria validation (predictive or concurrent), content validation, or construct validation.⁶

No construct or criteria validation was performed before McCann Associates prepared and the CSC gave the 1972 Sergeant's Promotional Examination. Furthermore, there is no time now nor is it practical to undertake construct or criteria validation of the 1972 Sergeant's Promotional Examination. Through its chairman, Vincent Francioli, the CSC has informed the court of the Commission's intention to conduct a new promotional examination with a target date for giving the test and certifying the eligibility list set for November 30, 1974. Even more important, Dr. Henderson explains why a content validation of this promotional exam is feasible, but criteria validation is not:

“I think content validation is certainly more feasible and perhaps a more reasonable approach, again assuming you have an honest job evaluation and the items being tested are fair samples of the kinds of knowledge and skills that the job requires, so I would say yes, content validation procedures would be, I would suspect, more common procedure for promotional exams; *and the practical reason is, of course, you usually don't have a large number of samples of existing sergeants, particularly of different races and sections to do a criteria related study. Consequently, we almost always have to fall back on a content validation procedure.* (Emphasis added.)”

[Content Validity]

Dr. Richard S. Barrett,⁷ called by the plaintiffs to testify at the hearing on the 1972 entrance examination, agreed the “test has content validity if it is in fact a sample of the skills or knowledges that are actually required on the job.” The terms “content validity” and “job relatedness” must be differentiated, as Dr. Henderson makes clear. Under questioning by plaintiffs' counsel, he gave “comparative definitions” of these terms.

*6 “Content validity refers to the validity of an item or series of items with respect to actual duties on the job. They are static, they involve the list of duties that must be performed on the job.

“Job relatedness could be achieved by content validity of all items or most items, but it can be achieved by other methods as well. The statistical relationship between a battery of tests and a performance on the job, even if the specific items will have no one-to-one correspondence

with the job can be a job related test. It is another form of validation.”

Dr. Henderson continued:

“And much of the confusion is coming in, in crossing over two lines. Think of content valid items as items that have specific duties. There is another whole realm of things dealing with general attitudes or motivation or what have you. These do not show up as items on a list in a job description.

“The degree of motivation of an individual, for example, should be reflected in that individual knowing the specific aspects of the job well enough to do well on an examination. One doesn’t ask a question about motivation.”

As seen from his earlier testimony Dr. Henderson says that content validation assumes “an honest job evaluation and the items (exam questions) being tested are fair samples of the kinds of knowledge and skills that the job would be.” With Dr. Henderson’s assumption in mind, it is essential to consider the circumstances surrounding the development of the Sergeant’s Promotional Examination.

[Job Analysis]

B. McCann Associates is a public personnel management consultant firm headquartered in Hunting Valley, Pennsylvania, a suburb of Philadelphia. From its inception in 1960, its major interest has been in providing entrance and promotional examinations for “a wide range of municipal jobs, specifically in the public safety field, of police and firemen.”

William Howeth⁸ of that organization received a telephone call from Jay Loeb, then personnel administrator of the CSC.⁹ As a result of that call Mr. Howeth met with Mr. Loeb in Cleveland on June 15, 1972. Though McCann Associates was officially hired later by written agreement with the CSC, Howeth undertook at that first meeting an exchange of information concerning the preparation of the promotional examinations that were later given. Because of breeches of security in the administration of promotional examinations in prior years the CSC delegated authority to its chairman to hire a firm to prepare the tests.

^{*7} At the meeting of June 15, Mr. Howeth left a “subject matter check list for planning a deluxe POLICE PROMOTION Examination.” Mr. Howeth states that he

“discussed the job of police sergeant with Mr. Loeb, and discussed the advisability of getting written job descriptions.” He received the job descriptions that are in evidence, either at that meeting or later by mail.

Mr. Howeth recalls that he suggested to Mr. Loeb “a possible examination plan” and that this be reviewed by the Police Department “to get their concurrence that it was an appropriate subject matter.” Despite this suggestion, CSC did not submit the proposed examination plans to the CPD. The reason given by Loeb for not informing the CPD was that security problems had existed in previous examinations.

A reading reference list, approved by the CSC, based on suggestions made by Howeth was listed in the CSC’s announcement of the promotional examination. When Howeth received this announcement together with the approved subject matter check list he discovered that under the heading of “Legal Knowledges” the completed check list included five items each for the following subject matter areas: “237–Laws regarding arrest;” “238–Laws regarding search and seizure;” and “239–Legal definitions of crimes, (the elements necessary to charge various degrees of crimes).” Howeth noted that the approved reading references did not include any reading reference that related to these three subject matter areas. After a telephone conference with Mr. Loeb the 15 items devoted to “Legal Knowledges” were stricken and Howeth substituted, with Loeb’s consent, other items. Howeth explained:

“It became evident that there were no references sources on the announcement which in our judgment would allow for appropriate testing in the areas of laws of arrest and laws regarding search and seizure.

“In addition, I probably would have recommended against “legal definitions of crimes,” had there been a reference on the list [because they are rote memory type questions] . . . I discussed this by telephone, and a decision was made at that time not to change the announcement, because I felt that was not a serious problem, that we could still create a job related test for the rank of sergeant without testing legal knowledges; so we made the change.”

For the stricken “Legal Knowledges” Howeth substituted five items in subject matter area “243–Ability to read and comprehend written instructions, laws, ordinances, etc.,” five items in subject matter area “244–Ability to understand and evaluate tables, charts, and graphs,” and five items in subject matter area “245–Problems relating to minority groups.” Howeth explained his reasons for adding these subject matter areas:

“Based upon my understanding of the job of sergeant,¹⁰ and based upon reading the job descriptions and the supervisory and training responsibilities, I felt that this area should be—this behavior area should be tested for—at the rank of sergeant as well as at the other ranks, and therefore reassigned 10% that had been assigned to “legal knowledges,” to subjects 243 and 244, which explains the notation on the exhibit.”

*8 Howeth further testified:

“I noticed that problems relating to minority groups had not been assigned for sampling at the sergeant’s level, and made the decision that this would be a good additional subject to include in the test.”

With the approval of Mr. Loeb, Mr. Howeth made other changes on the check list submitted by the CSC. For example, five items in subject matter area “256—Case preparation and court room procedures” were substituted for subject matter area “207—First aid and rescue work.” The check list as submitted by Loeb under the heading “Supervision,” assigned 20 items to “226—Police supervisory methods—general,” and five items to “228—Supervising traffic, accident functions.” Presumably, after discussion with Loeb, Howeth struck out subject matter area 228 and substituted five items in subject matter area “227—Supervising crime investigation.”

In his testimony of July 2, 1974, Dr. Henderson was asked whether the job descriptions furnished McCann by the CSC were sufficient for him to determine the content validity of the Sergeant’s Promotional Examination. He answered that the materials furnished (job descriptions and the subject matter check list) were topic headings beneath which many individual items could be filled in. Consistent with Dr. Henderson’s observation, the description of the duties of sergeant of police generally state[s] that he is “to supervise the activities of a platoon, detail, or squad of men engaged in general and special police duty,” etc. The “typical tasks” which a sergeant is likely to perform may be illustrated by several excerpts:

“To supervise the work of subordinates engaged in patrol or field duty or on general or special detail or in an assigned district . . . to perform general police duty in connection with the protection of life and property, the prevention and detection of crime, the preservation of law and order, etc. . . . to receive, check and make reports; to apprehend offenders; to assist detectives in investigation; to prepare evidence for prosecutions; to attend court. . . .”

Dr Henderson conceded that with such general topics it would be difficult for a person with his limited knowledge of police work to develop a promotional examination for the position of sergeant. However, as recognized by Dr. Henderson McCann Associates had knowledge of sergeants’ duties and experience in preparing sergeants’ promotional examinations. Howeth’s testimony bears out Dr. Henderson. Prior to preparing the Cleveland promotional exams, Mr. Howeth “conducted numerous studies of the rank of police sergeant.” Indeed, his testimony on the degree of specificity of the job descriptions furnished him by Loeb is illumination. Alluding to the subject matter areas that Howeth substituted for “Legal Knowledges,” plaintiffs’ counsel asked Howeth on cross examination:

“Would you glance over that [job descriptions] and tell me where it is in there that it shows that it is necessary for a sergeant to read and comprehend written instructions?—or to understand and evaluate tables, charts, and graphs?”

*9 Mr. Howeth answered:

“The job description does not set forth the knowledge, skills and abilities required for the rank of sergeant, nor of any of the other ranks. It is part of normal specification that is missing; however, a review of the tasks would indicate a number of them that would lead one to perform those tasks would be to obtain information through the normal manner, which would be analyzing technical material and data presented in other manners, which would be graphs, charts, and tables.”

It is concluded and determined that the topical headings of the job descriptions and the subject matter check list furnished by CSC to McCann Associates, where supplemented by Howeth’s prior knowledge of the tasks and duties of sergeant and McCann Associates’ general knowledge in the same area, provided the firm with a sufficient job analysis from which to prepare a promotional examination for the position of sergeant.

To support its defense that the promotional test is job related, the Fraternal Order of Police, Cleveland Lodge 8 (FOP) engaged George W. O’Connor, presently Commissioner of Public Safety for the City of Troy, New York,¹¹ to perform “a job analysis of the positions of police sergeant and police lieutenant; and to give his “opinion as to whether [he] felt the [1972] examination was related in terms of job knowledge.” On April 12 and 13, 1974, O’Connor interviewed 22 Cleveland police sergeants and 11 Cleveland police lieutenants (a 10% sample of the rosters of both sergeants and lieutenants

with names drawn at random from the respective units of the department).¹²

[Duties of Sergeants]

*10 Mr. O'Connor had never before prepared a job analysis; and it is evident that his job analysis of the 11 Cleveland police lieutenants and 22 police sergeants was conducted on the eve of trial. Dr. Barrett, plaintiffs' rebuttal witness, stated that O'Connor's lack of experience in doing a job analysis and evaluating the validity of an examination raised problems. He testified:

"Well, the point is, that it is not enough to know what a job entailed in order to determine whether the test for selection of people to that job is adequate. It is necessary to understand testing to understand what is meant by some of the technical but also fundamental concepts in the field, such as validity and reliability of tests;"

Reserving for the moment O'Connor's qualifications to pass on the content validity and job relatedness of the 1972 promotional examinations, it is concluded and determined that his 22 years of experience in the area of police administration and direct awareness of the duties of sergeants of police in Cleveland, Chicago, Oakland, Denver, Puerto Rico, Skokie, Fridley, Minnesota, and Troy, New York—more than make up for his previous lack of experience in conducting a job analysis. Asked whether the duties of sergeant in the City of Cleveland "varied from the duties as observed in the cities of Oakland and Chicago," he answered,

"I think there are always some things unique to a particular department, but I think that the great bulk of the duties of sergeant here [Cleveland] aren't at all dissimilar from those performed in most other police departments."

In his written job analysis O'Connor "felt that an indicator of the initial nature of the sergeants' job would be the first assignment." He ascertained that 12 were assigned to basic patrol—district stations—field supervisors. Five were assigned to desk duty, one each in accident investigation, community relations, records, communications, and recruitment. The report goes on to note that "12 of the 17 have always served in basic patrol, two have remained in administration units, and one each in Records, Communications, and Traffic." O'Connor observed that the

"[T]able of organization indicates that more than half of the sergeants are assigned to basic patrol. An additional

30% are working in line units such as detective and traffic."

In tabular form O'Connor summarized the "duties/responsibilities identified by 22 sergeants." In four columns he recorded the "assignment of sergeants" as to the four main categories of field, desk, detective, and traffic and administration. The sergeants in each of the four assigned areas identified their duties and responsibilities. Totalling these figures for the most frequently appearing "type of duty/responsibility," 19 sergeants indicated "direct supervision of personnel," 14 sergeants indicated "assigned personnel on daily basis," 9 sergeants "maintain time/attendance records," 19 sergeants "reviewed reports for completeness, accuracy, etc.," 13 sergeants "conduct roll calls," 3 sergeants "investigate license applicants," 4 sergeants indicated "book and handle prisoners," 8 sergeants "investigate complaints about men and investigate accidents involving departmental vehicles," 3 sergeants "receive calls for service," 2 sergeants indicated "handle public relations," and 3 sergeants "handle calls for service in field."

*11 The text that follows his table indicates that "direct supervision of personnel" includes several of the following recited duties and responsibilities. The text in part states:

"In each group of interviewees, the point was made clearly that sergeants were first of all police officers and secondly, supervisory personnel. Thus, this basic assumption eliminated the need for the interviewees to deal with fundamental police tasks such as the preservation of lives, protection of property, rendering of medical services, detection, and apprehension, prevention and control of crime and delinquency, etc."

The report continues:

"To further substantiate this point, some groups were asked to assess the extent to which their operational time was divided between supervisory duties and those which every policeman was required to perform. Of the seven individuals asked to make this judgment, they divided the time roughly into a ⁶⁰/₄₀ split with supervisory duties occupying the larger portion. In three cases, administrative duties were said to occupy their time also. In two of these latter cases, supervisory duties (those tasks considered to be appropriate for sergeants) took at least 60% of their time and attention."

In his "conclusions regarding the duties of sergeants," Mr. O'Connor states, among other things,

"The normal or average work of the sergeant is a mixture

of engaging in operational tasks and in providing direct line supervision to operating personnel of lower status. The men interviewed agree almost without exception that their duties did not require them to engage in routine development of reports other than those required in response to a specific incident or complaint. In other words, sergeants are essentially reviewers of the reports of others. Seldom do they generate documents peculiar to their position. For example, except in the case of probationary recruits, there are no regular performance appraisals of employees required of sergeants. They need not prepare regular reports of conditions within their unit or area of assignment. They must originate reports of incidents which they handle just as any police officer might.”

When asked to state his conclusion concerning O’Connor’s job analysis, Dr. Barrett testified:

“The report really gets to the job analysis part on page 7, where there is a table headed, “^{Duties/Responsibilities} Identified by 22 Sergeants,” and it is broken down by their assignment in the field, desk, traffic, administration, and so forth, and then he lists the type of duty or responsibility.

“The first one is “Direct supervision of personnel.” Now, I don’t find any more about the direct supervision of personnel than that in this job description.”

Later in his answer, Dr. Barrett continued:

“Now, I am sure O’Connor knows a lot more about this, but it is not available for me to examine. In fact, it doesn’t tell me anything I didn’t know before, and it says, ‘He assigns personnel on a daily basis.’”

“Well, I have no idea from this, nor does anyone else reading it, what that means. Is it a totally routine operation in which he has so many slots to fill, and he puts the people on; or is it a more complicated type of job, in which he has to match the quality and qualifications on an individual with the work to be done, and he has to make sure nobody gets unpleasant if he sees his name, and that person has a tough assignment, and he thinks he should be given an easier one. There is no way of knowing whether it is an important activity, trivial, or whether it is easy.”

*12 Dr. Henderson found Mr. O’Connor’s topic headings even broader than those in the job descriptions furnished McCann Associates by the CSC. He made it clear that with his own background he would not be able to devise a test given the generality of the job descriptions and O’Connor’s job analysis. However, he added “we have to

remember George O’Connor could flesh out the descriptions.” At another point, Dr. Henderson stated, “I would have to agree [O’Connor] could fill in short hand that would fill out the general duties [of police sergeant].”

It is concluded and determined that the job analysis prepared by George W. O’Connor is not in sufficient detail to permit its employment by an educational psychologist to ascertain the content validity or job relatedness of the 1972 Sergeant’s Promotional Examination. However, it is concluded and determined that, assuming O’Connor otherwise has the qualifications to render an opinion as to the content validity or job relatedness of the 1972 Sergeant’s Promotional Examination, his job analysis is sufficient, coupled with his wide experience in police administration and police work, to form a basis for any opinion which he may be qualified to render.

[Job Relatedness of Questions]

C. The court must now consider whether upon all the relevant evidence the City defendants have sustained their burden of proving the job relatedness of the 1972 Sergeant’s Promotional Examination. In undertaking this task the court acknowledges that it has not been able to avoid “the burden of judicial examination-reading and the risk that a court will fall into error in umpiring a battle of experts who speak a language it does not fully understand,” *Vulcan, supra* at 396.

Howeth stated that it was his opinion that the several 1972 promotional examinations, including the sergeant’s, “did test the knowledges that were directly related and required for successful performance of the ranks.” In weighing Howeth’s conclusory and unelaborated opinion it must be realized that he is being asked to self-evaluate a test whose items were pulled together in sets of five under his direction from McCann’s bank of previously prepared items. Howeth testified later in the trial that “if an item was used in a previous exam it would be subjected to item analysis.” Nevertheless, the Cleveland item analysis¹³ discloses that 40 of the test’s 100 items were determined by McCann to be nondiscriminating.¹⁴ For McCann to have used 40 nondiscriminating items, which McCann must have previously found to be nondiscriminating or which 40 items were used in the Cleveland examinations without previous testing experience, casts doubt on the necessary predicate of Howeth’s opinion that the examination could test “for successful performance of the ranks.”¹⁵

*13 Turning to the testimony of George W. O'Connor concerning the job relatedness of the sergeant's promotional examination, O'Connor stated on direct examination:

"I believe I indicated that I do feel that the examination is related to the job of that of sergeant in the Cleveland Police Department."

Counsel for the City defendants further asked O'Connor:

"Do you have an opinion as to whether in fact that examination does measure the knowledge, the skills, and knowledge required to perform the job of police sergeant?"

He answered:

"I believe it does measure some of the areas of knowledge, not all of them—some skills, but not all of them; in order to perform adequately as a police sergeant."

The probative weight of O'Connor's opinion of job relatedness, not particularly strong to begin with, is further weakened by his self-evaluation of his opinions in the testing area. As to items 16 through 20 of section 1, designed to test "reading comprehension," he gave certain "impressions" which he acknowledged were nothing more than "a layman's opinion on that particular point."

He was further asked:

"Now, do you know what the term "validity" means as it is used in test evaluation?"

He answered:

"I think, in a general way, but I am not sure that I know at what particular point a question, for example, might be considered valid, or a whole test might be considered valid."

When asked if he held himself out as "an expert in the area of evaluating the validity of examinations," he answered "no."

No attempt will be made to pass judgment on the content validity or job relatedness of the sergeant's promotional examination on an item-by-item basis. Nevertheless, it is worthwhile to evaluate the examination through a comparison of the views of Dr. Barrett and Dr. Henderson.

The examination consists of 100 multiple choice items. Dr. Barrett was critical of multiple choice tests. He said, "Life doesn't come with a series of five alternatives, one to be chosen." He continued:

"Multiple choices are a test of recognition. You recognize the right answer when it is presented with three or four incorrect answers; but in a situation where one is facing someone being arrested or something like that, there is no one there reading off a list of five items, one of which is correct.

"He has to act and do something. This is a very different thing psychologically. Acting is very different from recognizing the right answer on the test."

However, he later agreed that "there are applications where a written multiple choice examination is appropriate as a selection device;" and that he does not have "any blanket objections to such tests overall."

Although not asked specifically about the use of multiple choice questions, Dr. Henderson voiced no objection to their use. Upon the present record, it is concluded and determined that there is no basis to find that an examination consisting of multiple choice items is inherently culturally biased or not job related.

*14 Dr. Barrett was also critical of what he termed "a very strange response format for some questions." There are 22 such items in the examination whose format may be described as double-tiered. These items comprise in the upper tier a list of several numbered pieces of information upon which a question is premised. In the lower tier of the item this question is followed by five alternative responses which require the examinee to determine what combination of numbered pieces of information best answer the question. Dr. Barnett commented on these questions as follows:

"What this does is introduce into the testtaking procedure a great deal of emphasis on the skill at an unraveling kind of response format. A person has to bear in mind which of the items, one, two, three or four, he agrees with, and which ones he disagrees with; and then after that he has to go through these five possible alternatives and pick out the one that fits his prior conclusion."

He stated that this format in his experience is "not commonly used." He further stated that this type question would have an impact

"on the individual who did not have what we call in our trade the "test wiseness;" that is, he did not know the art

of taking the test as separate from the art of knowing what is in the test. He would have to know how to juggle this sort of abstract and rather complex way of answering questions, which is a separate thing from knowing the right answer to the question.”

In contrast, when Dr. Henderson was asked to give his opinion as to the effect of these double-tiered items on the job relatedness of the examination, he answered, “there is nothing inherently wrong in using double-tiered questions.” He said that they are difficult to construct and that any ambiguity could be confusing among people who are test anxious. However, he found only one or two questions of this type with which he was unhappy, and pointed out that McCann did a good job considering the complexity of this type of question.

Dr. Barrett was asked whether there were any questions on the test which could be regarded as testing for reading comprehension. He answered:

“Well, there are some that are deliberately designed to be reading comprehension items. There is one passage that covers almost all of a legal size piece of paper. Section 1, questions 16 through 20, are based on that passage, and they take another whole page, and it is intended as a reading comprehension test; and now, of course, the rest of the test being so long and so verbal, is in itself a reading comprehension test.”

With reference to these same items, Dr. Henderson was asked to give his opinion as to whether the knowledges or abilities which these items seek to test are job related. He answered, “Again my opinion is intuitively, these are job related.” He added that the ability to read certain materials and make interpretations is related to the job even though there is not much in the job descriptions that would suggest that. He also concluded that these questions tested reading skill; and, that they overlap the other questions of the exam only to a small degree.

***15** On the same subject plaintiffs’ counsel asked him:

“Is not this particular examination that you have before you heavily loaded with questions that require skills of reading comprehension simply to wade through the very questions themselves? Let me add, given the fact this is a time test.”

Dr. Henderson answered,

“Yes. But again as I testified earlier on the most complex questions, the two-tiered questions, the scores were reasonably good. And the second issue I think we have to

remember is that McCann Associates received the check sheet which included an item that said reading instructions, etc. And one might regard all items as reflecting that to some degree.

“I think it would be fair to say that considering the subject matter involved in this test, the difficulty level or the attempt to mislead through complicated writing is far less than one often finds in such examples. So I believe on this point, while it is—it does require wading through a lot of written material, it is a reasonably clear test given the subject matter. . . .”

Thus, differing from the opinions of Dr. Barrett, Dr. Henderson found that the reading comprehension items (Section 1, 16-20) overlapped the other questions only to a small degree and that although the test required “wading through a lot of written material, it is a reasonably clear test given the subject matter.”

The plaintiffs have questioned the job relatedness of the examination because it omitted questions relating to Ohio laws and City ordinances. As seen, although the CSC originally requested questions on “Legal Knowledges,” its failure to include text references on this subject in the public announcements of the promotional examinations caused Howeth and Loeb to substitute other items. Although O’Connor recognized in his report that “the work of Sergeants required an in-depth understanding” of laws, ordinances, and regulations, he nevertheless sought to justify the omission of items in this area by the assumption that it would seem reasonable that “candidates for Sergeant possess a sound working knowledge of the basic laws and codes.” Dr. Henderson generally testified that an examination can be job related and not touch all bases. Thus, standing along the omission of the “Legal Knowledges” subject matter area is not deemed sufficient to invalidate the examination. However, the day-to-day importance of this subject warrants the CSC to consider the inclusion of these matters in the preparation of new promotional examinations.

Mr. O’Connor had reported that seven of the 22 interviewed sergeants roughly distributed their duties and tasks as 60% supervisory and 40% duties which every policeman is required to perform. Of the 100 items in the examination 25 were designated as falling within the supervisory subject matter area, and 13 of these 25 items were identified on McCann’s item analysis as nondiscriminating.

Asked the effect on job relatedness of giving a test with 25% supervisory items to candidates for a position requiring 60% supervisory duties, Dr. Henderson gave

these opinions. He said that if the questions were drawn from independent aspects of supervision, covering a variety of duties, he would be far more comfortable than if the items were tapping a narrow range of duties. Additionally, he said that while 25% supervision items was “rather modest,” he would be far less concerned about these supervision items if the applicants for promotion would be expected to go back to the police academy to learn the supervisory skills. The record shows that the CPD has not provided such post-examination in-service training. He did later say that his concerns would be “ameliorated” if he were to assume that the substantial portion of the examination dealt with the practices, procedures, and subject matters as to which sergeants supervise patrolmen.

***16** In considering the job relatedness of the examination it is apparent that both Dr. Barrett and Dr. Henderson were concerned that McCann’s item analysis disclosed that 40 out of the 100 items are non-discriminating. Dr. Barrett noted this when he testified:

“40% of the items were indicated as being worthy of a designation of “N” which meant that they were inadequate in discriminating between the high and low test takers, so those 40 items, really, are not contributing much to the examination and should not have been used in the first place.”

When asked, “In your opinion and experience, is there any validation for having as many as 40 out of 100 questions on an examination not particularly discriminatory?” Dr. Barrett answered in the negative.

With reference to the 40 items that are nondiscriminating Dr. Henderson indicated that as the number of nondiscriminating items increases the reliability of the test begins to drop off. It loses its predictive power, although it still may be job related. He indicated that further studies would have to be made to determine the effect of the items rated nondiscriminating by the item analysis of McCann. He volunteered that he believed that at least 16 other items were weak discriminators. He pointed out that it was not unusual that nondiscriminating items show up on a test. However, he would prefer to look more closely at the matter.

The final judgment of Dr. Barrett concerning the validity of the examination was

“that there is no demonstration that the tests are valid, which is distinct from a conclusion that the tests are not valid for different things.”

He was then asked, “And you are not concluding the tests are invalid?” He answered, “That is right.” He was then asked, “You cannot make that determination on the basis of the knowledge that you have?” He answered, “That is right.”

In contrast, Dr. Henderson testified that “intuitively [it] appears to me this is a job related exam –.” On the witness stand for most of July 2, and subjected to repeated questioning first by the court and then by each counsel, Dr. Henderson kept coming back to his “intuitive judgment” thus expressed. Illustrative is this question and answer on cross examination by counsel for the FOP:

“Is it a fair summary of your testimony on this point that despite a number of factors that trouble you, a number of unknowns, the combination of which weaken the strength of your opinion to be not very strong at all, that nevertheless it is your intuitive judgment that the test is probably job related?”

“A. My intuitive judgment.

During the noon intermission in his testimony, Dr. Henderson worked out several equations on the court room blackboard relating to the reliability of the test. Subsequently, counsel for the Shield Club plaintiffs conducted this interrogation of Dr. Henderson:

“Q. I understand that you are defining reliabilities of .1 and .2¹⁶ as essentially findings of trivial reliability?

***17** “A. Well, they are very unreliable, but as I stated earlier, anything better than zero begins to—it’s better than a chance prediction.

“Q. And you would expect that this test reliability probably is in excess of .1 or .2?

“A. Yes.

“Q. But do I understand you to say that you are unable to determine beyond that the nature of the reliability?

“A. That’s right.”

Dr. Henderson was then asked to make comparisons.

“Q. Do you have any reason to suppose the test is reliable within the range of .6 to .8 that you described before?

“A. I’d be surprised if it was that high, because that is so unusual. It takes a number of years of testing and redeveloping of items before you have a solid test that

reaches into the 80's."

Later, counsel for the City defendants asked Dr. Henderson whether there is justification for appointing sergeants on the basis of a list formulated in rank order, taking into account only the raw scores of the applicants and using the appointive power's authority to select one in three. Dr. Henderson answered:

"My answer earlier concerning the job analysis was that it's very hard for me, given these general statements, to argue convincingly that it was directly job related, that I had to lean on my own intuition that the items at least were reasonable and depend somewhat upon the skills of the test makers at that point.

"Leaning on my own intuition, and using this estimate of reliability that exceeds .2, I would have to say that the procedure is rational in that it probably is a little better than drawing lots. Beyond this I would be hard put to make any strong arguments for the test."

In the final questioning of Dr. Henderson by counsel for the plaintiffs the factor of selection ratio was introduced for the first time. Dr. Henderson defined this term as the "proportion of persons taking the test who would be selected." Dr. Henderson stated that there is no relationship between establishing test validity and the selection ratio. But, as he put it, "a crummy test can get by if you only had a 2% selection ratio." He added, "you can take advantage of the numbers game and sacrifice some test validity when the selection ratio is small." Counsel for the plaintiffs finally put this question to Dr. Henderson:

"This gets back to the same question of the . . . slight difference between the utility of this examination and the lottery system, it's the same comparison that you alluded to before?"

Dr. Henderson interjected:

"Now, wait. The validity of this exam may not be high, but the selection ratio is so small that the utility of the exam is probably at least comfortable. In other words, you can afford to make mistakes and you are making mistakes because any individual taking this exam might shift several points in either direction later on, but by selecting so few individuals out of the entire pool you are skimming from the very top. And the likelihood of someone clearly down the list, and probably not qualified for sergeant, will on another taking show up in the high 80's, that is rather unlikely.

***18** "So we have to watch the three terms we are using here, selection ratio, validity, and utility. With the selection ratio as small as this, the utility of this test may not be bad, even if its validity is at best modest."

Dr. Henderson's testimony then ended with this questioning by the court:

"Q. So it would be a matter of finding—of going no higher than 70 [the upper 10% of the 695 test takers], but as you approach 70, the so-called weak predictor has less chance of being a basis for job relatedness, or are you saying that within that 10%—

"A. You increase your chances as you go beyond 10%. As you choose more than 70, each person that you add has a slightly higher probability of being unqualified.

"Q. Now I am talking between the 26 who presumably already have been appointed, or 25, and a top of 70, is there similarly a range where the weak predictor characteristics of this test would have an effect?

"A. Yes. It is a gradient. Starting from the top, each succeeding person has a slightly higher probability of being inappropriate for the job. But the increase is not really serious until you get beyond that selection ratio of 10%. Again we are working with approximations since none of us really know what the validity or reliability of this test is.

"The Court: Very well. . . ."

[Conclusion to Job Relatedness]

D. Upon all the evidence it is concluded and determined that whether the City defendants have carried their burden of proving that the 1972 Sergeant's Promotional Examination is substantially related to job performance depends on the weight to be assigned to the testimony of Dr. Henderson, a thoroughly credible witness. For reasons previously noted the opinions of job relatedness of Mr. Howeth and Mr. O'Connor will be given no controlling weight; and Dr. Barrett's ultimate opinion was that he could not make a determination of validity on the basis of the knowledge he had.

The question this court must now decide is whether it is convinced that the sergeant's examination was "substantially related to job performance," although this need not be proved "to the hilt." See *Vulcan*, *supra* at 393. "Substantially job related," as here used, is understood to

mean “of or pertaining to the substance or main part” of job relatedness.

The examination’s use of multiple choice styled items, the double-tiered format of 22 of the 100 items in the exam, and the reading comprehension items, are concluded and determined to have had no constitutionally impermissible effect on the 1972 Sergeant’s Promotional Examination. But other facts developed during Dr. Henderson’s testimony of July 2, raise serious doubt about the job relatedness of the examination and its reliability in formulating the eligibility list for sergeants’ promotions. Dr. Henderson concluded that 25% supervisory items was “rather modest,” and that he “would be far less concerned” about these supervisory items if sergeants “went back to the police academy” to learn supervisory skills after their appointment. Considering the absence, following appointment, of any in-service training for sergeants in which they may learn supervisory skills and the paucity of evidence that the 12 discriminating supervisory questions provide a sufficient sampling of sergeant’s supervisory duties, it is concluded and determined that the examination’s treatment of supervisory duties seriously impairs the job relatedness of the total examination.

***19** Both Dr. Barrett and Dr. Henderson were alarmed that 40% of the 100 items were found by McCann to be nondiscriminating. Dr. Henderson said that each additional nondiscriminating item causes the reliability of the test “to drop off” and the test “to lose its predictive power.” To these 40 nondiscriminating items Dr. Henderson added “at least 16 other items” that he felt were weak discriminators.

Finally, Dr. Henderson found that the test has a reliability of above .2 although exactly how far above that, without further reliability studies, he could not say. He stated that .2 indicates “that the procedure is rational in that it probably is a little better than drawing lots.” As to its job relatedness, Dr. Henderson leaned on his “own intuition that the items at least were reasonable and depend somewhat upon the skills of the test makers at that point.” Synthesizing these measurements of the reliability and the job relatedness of the test, Dr. Henderson described the test as a “weak predictor,” that is to say, that it is a weak predictor of successful job performance. Because the test is a “weak predictor,” Dr. Henderson concluded that not more than 10% (70) of the 695 test takers should be considered for promotion. The use of a “selection ratio so small” means that “by selecting so few individuals out of the entire pool you are skimming from the very top.”

By a preponderance of all the evidence, it is concluded

and determined that although the examination has some validity, the test is not shown to be “substantially job related.” Therefore, further use of the test and the eligibility list based thereon is enjoined, except as ordered in Part IV of this memorandum and consistent with the determinations reached in Part II on the issue of seniority points.

[Prior Service Credit]

II. Plaintiffs allege that the practice of adding points, based upon seniority, to the raw scores achieved on promotional exams has “the effect of reinforcing the impact of past racially discriminatory hiring practices.” The practice to which plaintiffs refer is specifically described by the CSC as prior service credit and is set out in the CSC rules. In pertinent part the rule states:

“Applicants obtaining passing grades in promotional examinations shall have added to their passing grades credit for seniority. Such credit shall be for all service rendered, pursuant to appointment from an eligible list, in all classifications which are lower in rank and which are considered in the direct line of promotion. The amount of such credit shall be obtained by allowing an amount for each month of an applicant’s prior regular service in accordance with the following schedule: one per cent of the total grade obtainable for each of the first four years of prior service and six-tenths per cent of such total grade for each year of the next ten years of prior service.”

Although this rule has general application to all promotional examinations overseen by the CSC, it is specifically attacked in the instant case as to its racially discriminatory impact on minorities taking the police promotional tests.

***20** Several general observations about prior service credit and its effects upon minorities must first be set out. The aggressive minority recruitment program, the new police entrance examination, and the anticipated development, pursuant to this court’s order, of a more objective screening procedure of patrolmen who have passed the entrance examination, should markedly increase, and properly so, the minority representation in the CPD. In order to be eligible to take the sergeant’s examination it is undisputed, that a patrolman must have a minimum of three years’ service as a patrolman. Once promoted to a higher rank, an officer must have at least one year at that rank before becoming eligible to take the promotional exam for the next higher rank.¹⁷

It is further evident that of the maximum of 10 seniority points that may be awarded, the first four points accrue at the rate of $\frac{1}{12}$ of a point per month, or one point per year for four years. However, after four years the rate becomes $\frac{1}{20}$ of a point per month, or $\frac{6}{10}$ of a point per year for 10 years. Therefore, although the accrual rate is slightly accelerated for the first four years it remains constant at a lesser rate for 10 additional years, requiring a total of 14 years of service in order for an individual to obtain the maximum seniority points. Thus, it is readily apparent, given the increased number of minority patrolmen who have recently joined and will be joining the CPD, that the use of seniority points as they are presently awarded will have a racially disproportionate impact upon promotions within the CPD. Nevertheless, it must also be determined whether seniority points, which obviously award experience, have a sufficient job relatedness that would make their award appropriate to assist in determining the patrolmen most qualified for promotion.

[Job Relatedness of Seniority]

As a prelude to a question regarding the use of seniority points, Mr. O'Connor was asked about the three-year and one-year time-in-grade requirements.¹⁸ Concerning the three-year time-in-grade requirement for a patrolman to become eligible for the sergeant's examination, he stated:

"I think it is appropriate. I think in view of the fact that much of a sergeant's responsibilities relate to supervising and observing and understanding what others are doing, that he himself needs to have an in-depth and broad knowledge of the technical aspects of law enforcement and police work.

*21 "I am not sure that three years is the precise number of years, or whether it is subject to some fluctuation—two years, or three years—

"I think three years is about as much as is necessary."

Similarly as to the one-year time-in-grade requirement to become eligible for each succeeding rank above sergeant, O'Connor stated:

"I believe that it is job-related, and I think the one year on top of three is cumulative. I don't see three years in each case. I think the experience a man gains in the department is cumulative."

O'Connor was then asked his opinion as to the job related justification for awarding up to 10 seniority points for 14

years of service. He stated that "I don't think it has any value in terms of being job related." He then went on to explain his conclusion:

"Well, I just indicated that I believe that experience is an important prerequisite to promotion. Assigning specific points to each additional year of experience I don't think is necessarily any appropriate way to measure the value of that experience. It isn't. Additional experience may or may not add to the core of information the individual brings to a higher rank.

"But providing points for seniority, as it were, and particularly in a highly competitive system as we find in these types of promotional examinations, those seniority points very often make the difference from being promotable really and never having an opportunity."

O'Connor then referred to his April, 1974 interviews.

"In terms of the group that I interviewed, I asked for a profile of their experience within the force, and it is a part of the record which I have presented to you.

"In the average of the 22 men who were sergeants, the average tenure of those men in the rank of patrolmen was 15 years and 1 month before they were promoted.

"In the case of lieutenants their average tenure as patrolmen—I am sorry—I don't show those averages, but I would estimate it is in excess of 10 years, and their average tenure as sergeant prior to promotion to lieutenant, I would estimate that to be somewhere around 7 to 8 years, and I think seniority points are the major factor in that amount of tenure.

"In departments that do not have seniority points it is not unusual to find younger and experienced men achieving promotions more frequently than here. I think the tendency has been to reward age without regard to evaluation of that age."

Thus O'Connor states that three years is an appropriate time in grade requirement to become eligible for the sergeant's examination. But he states that seniority points are not job-related and are not necessarily the "appropriate way to measure the value of . . . experience." Implicit in his conclusion is that experience has a certain value for purposes of promotion and that experience or the knowledge presumably obtained through experience can be tested in other ways.

Dr. Henderson's opinion and approach to the issue of seniority was somewhat different from that of

O'Connor's. Henderson was asked whether the job relatedness of the selection procedures used to determine promotions to the position of sergeant would be enhanced by a system that awarded up to 10 points on the promotional eligibility list for 14 years of service. Dr. Henderson pointed out that the specific number of points awarded for seniority could only be considered in terms of the point spread between the highest and the lowest scores on the 1972 Sergeant's Examination. He stated:

***22** "There was a 48-point spread.¹⁹ . . . [A] total of 10 points is a rather substantial number for those people earning the maximum. I am of the opinion that either you have an examination that is good enough to detect this additional experience by the questions asked and the quantity of the questions asked, or you begin to consider this. I would be far happier to see the number of points go up more rapidly early as any learning curve would rise more rapidly during the first five years and perhaps the maximum be reached in a time shorter than 14 years.

"And I would, secondly, feel that seniority points would be in much better standing if there was more . . . rotation of duties within the force so that additional years of experience results in additional experience in other aspects of the job."

Specifically regarding the present system of awarding seniority points, Dr. Henderson stated "there is some very weak justification for that. It's far from the optimum scheme, but probably some very modest positive relationship."

Based upon the testimony and evidence before the court it is concluded and determined that the seniority points rule is likely to have a racially disproportionate impact upon future promotions and has insufficient job relatedness to support its continued use. *See Allen v. City of Mobile*, [5 EPD P 7958] 466 F.2d 122 (5 Cir. 1972). Promotional testing in the CPD is and will be undergoing substantial change and modification, pursuant to other aspects of this court's order. The CSC is ordered in any future police promotional tests to refrain from the use of seniority points. However, the CSC is authorized and directed to work out and adopt in the CSC rules reasonable jobrelated time-in-grade provisions. Whether the present practice meets this requirement is for the CSC to determine subject to review by this court.

As this court has expressly indicated above, the present manner of awarding seniority points has a racially disproportionate impact and is not substantially job related to justify its continued use. The court nevertheless determines that for the limited number of additional

appointments to be made from the existing eligibility list for the position of sergeant (see Part IV), the court will not reorder the list by using individual raw scores without seniority points. Several reasons compel this conclusion.

The CSC announcements of the 1972 Sergeant's Promotional Examination contain the notice

"SENIORITY: Determined in accordance with the rule of the Civil Service Commission."

Candidates' reliance upon the seniority points rule is a factor which this court should disregard only if it is clear that compliance with the seniority points rule will have a clear and demonstrably racial impact on any appointment from the sergeants' eligibility list, under the limited use that is ordered in Part IV. Indeed, the contrary appears to be true. It is apparent that those minority officers that have an opportunity for promotion would be severely affected should promotions be based upon raw scores alone, since those close to promotion are there because of entitlement to the maximum number of seniority points. Further, as has already been pointed out, although the promotional examination has sufficient reliability and content validity to permit its limited use, the exam does have a variety of weaknesses. Thus, it is not surprising that when Henderson was asked whether a person with a raw score of 85 on the promotional examination is any more qualified to be a sergeant than one with a raw score of 83, he responded:

***23** "I doubt that the test is that precise that one would take seriously a matter of a point or two. The question can only be answered by having better information concerning reliability of the test."

Henderson was then asked whether his answer would change if the difference between the two raw scores were 5 points. He stated:

"Well, obviously I would be more confident that there would be a difference in quality. But again I have no yardstick to go by without that reliability coefficient [i.e., an additional reliability study]."

The court has reviewed the individual raw scores that would be determinative of an individual's position on the promotional list for the limited number of appointments that this court is ordering under Part IV. The maximum point spread between the highest scoring and lowest scoring individuals who may yet be promoted from the list could be no more than four points. Although limited use of the eligibility list will be permitted there is no basis to believe that the exam which generated the eligibility

list is so precise that a difference of four raw score points realistically can determine which patrolman is better qualified for the position of sergeant. Nor has any witness attempted to assess whether the examination tested for the "value of experience."

For all these reasons it is concluded and determined that the limited additional appointments that this court will permit in Part IV of this opinion are to be made from the CSC Eligibility List for Sergeants prepared and published on February 14, 1973, and are to be based upon a patrolman's raw score plus seniority points as awarded pursuant to the present CSC rule.

[Supervisory Positions]

III. There remains to be considered the plaintiffs' attack on the 1972 promotional examinations for the supervisory positions above the rank of sergeant; namely, lieutenant, captain, deputy inspector, and inspector of police.

The only statistical evidence presented by the plaintiffs with reference to the 1972 promotional examinations for these positions concerned the position of lieutenant of police. One hundred and forty persons took the examination on October 6, 1972, of whom 136 were white and four were minorities. Undoubtedly the fact that only four minorities took this exam is attributable to the small percentage of the 209 sergeants who are minorities. Because the number of minorities taking the examination is too small a sample, no statistically reliable percentage comparisons with the whites who took the examination are possible. *Mayor of Philadelphia v. Education Equality League*, 42 U.S.L.W. 4405, 4409-10 (U.S. March 25, 1974).

Accordingly, in the absence of an adequate statistical sample with which to make a reliable percentage comparison with the whites who took the lieutenant's promotional examination, no prima facie case of a racially disproportionate impact as to the examination has been made. Of course none has been made for the higher ranks either. Hence the burden to prove the job relatedness of any of these examinations has not shifted to the City defendants. In passing, however, it is noted that in the formulation of new promotional examinations for positions above the rank of sergeant, ordered in Part IV, the Civil Service Commission will surely apply the teachings of this memorandum and order, and the lessons the parties, counsel, and the court have learned throughout this litigation.

[Remedial Relief]

*24 IV. In determining the relief that is appropriate in view of the foregoing findings of fact and conclusions of law it is essential to differentiate between permanent orders that relate to the formulation of new promotional examinations and related matters; and orders that are justified by this record to govern sergeants' promotions during the time remaining until the CSC certifies a new sergeant's promotional list based on a new sergeant's promotional examination.

As the testimony of Dr. Henderson disclosed, the 1972 Sergeant's Promotional Examination has utility as a predictor of successful job performance only within the top 10% of those who took the sergeant's exam. Thus within this limited range of highest scores there is a likelihood that persons appointed will be qualified. However, as he stated, "starting from the top, each succeeding person has a slightly higher probability of being inappropriate for the job." Dr. Henderson's conclusions provide guidelines for the interim orders that are determined to be justified by the entire record. The permanent orders likewise are based upon the entire record.

The following orders are entered contemporaneously with the separate memorandum and order of this court that relates to the screening of candidates for appointment as new patrolmen, and the assignment of police officers to the various units of the Cleveland Police Department.

1. The City defendants are enjoined from any further use of the February 14, 1973 "Sergeant's Promotional Eligibility List for the position of Sergeant of Police (Promo)," except as permitted in Paragraph 2.

2. In the order in which the names appear on said eligibility list, not more than 10 officers may be appointed from said list on or before September 15, 1974; not more than six additional officers may be appointed from said list between September 16 and October 30, 1974, inclusive; and not more than six additional officers may be appointed from said list after October 30, 1974 and until the replacement of said list by a new sergeant's promotional eligibility list.

3. The members of the CSC are directed to accomplish their previously announced intention of preparing new promotional examinations for the position of sergeant of police. Because of the interrelationship of the supervisory positions in the CPD, the members of the CSC are

directed also to prepare new promotional examinations for all ranks above that of sergeant.

4. Consistent with the foregoing findings and conclusions, in the formulation of any written examinations under its rule permitting "an objective type examination," the CSC may continue to use multiple-choice questions, but the members of the Commission are directed to give full and careful consideration to whatever other means of testing or selection may reasonably conform to the highest standards and practices applicable to the selection and promotion of police supervisory personnel.

5. Consistent with the foregoing findings and conclusions, in the formulation of any written examinations the members of the CSC are directed to produce job-related examinations validated in accordance with standards of educational and psychological testing set forth under

effective EEOC guidelines.

***25** 6. The members of the CSC are directed to submit any proposed promotional examinations to this court for its approval following consideration at an appropriate hearing.

7. The within order of this court and the separate order entered contemporaneously are subject to the further order of this court, and shall be administered consistently with the injunctive order entered by the Honorable Ben C. Green in the case of *Jean Clayton v. City of Cleveland*, No. C 73-784 (N.D. Ohio, May 9, 1974).
It is so Ordered.

Footnotes

- 1 Although not directly relevant to supporting plaintiffs' initial burden of demonstrating that the 1972 Sergeant's Promotional Examination had a racially disproportionate impact the statistics on the 1970 Sergeant's examination bear mention.
In 1970, 742 persons took the Sergeant's Promotional Examination. Sixty (8.1%) were minorities and 68 [682] (91.9%) were whites. Forty-four (73.3%) of the minorities passed the 1970 examination and 492 (72.1%) of the whites passed the examination. White patrolmen had an average raw score of 74.7 on the examination and minority patrolmen averaged 72.2.
Computation shows that during the two-year life of the 1970 Sergeant's Promotional Examination, a minimum raw score of 82 was required for promotion. Of those minorities taking the examination, 4 (6.7%) scored 82 or higher and of those whites taking the examination, 152 (22.3%) scored 82 or higher. Based upon their raw scores on the 1970 Sergeant's Examination, supplemented by seniority points provided by CSC rule, four (6.7%) of the minorities taking the exam and 105 (15.4%) of the whites taking the exam were promoted. Dr. Norman Henderson was asked to comment on the foregoing statistics. He testified:
". . . [I]n the black pool one would expect roughly 12 of those applicants to have scored 82 or above, and 48, 81 or less, and the actual number was 4 and 56.
"Again that is a statistically significant result, i.e. there is a difference in the distribution of scores, blacks and whites, with respect to this cutting point 82."
The court asked what that means. He answered:
"It tells us that when we look at the breakdown in percentages and we see that using Exhibit 93 as an example, where we see 22% of the whites taking the exam and scoring 82; and 6.7% of the minorities scoring 82; that that is not due to just normal fluctuation, and that if we had regiven that test to a similar population, they wouldn't have reversed themselves. . . ."
- 2 The passing score of 70 is established by rule of the Civil Service Commission.
- 3 The two-year life expectancy of the examination may be shortened if the CSC, as anticipated, prepares a new test to be administered in late 1974 or early 1975.
- 4 Since January, 1974 five more patrolmen have been promoted to the sergeant's rank. None of those persons promoted have been minority patrolmen.
- 5 Dr. Henderson has a Master's degree and received in 1961 a doctorate degree in psychology. He has been a professor at Oberlin College since 1960. His specialty is psychometrics, or statistics as related to psychological problems.
Dr. Henderson was hired by the Cleveland Civil Service Commission in early 1973 to prepare a job validated entrance examination for the position of patrolman in the City of Cleveland. This examination was administered by the Civil Service Commission on February 23, 1974.

Dr. Henderson was called by plaintiffs on April 18, 1974, to testify as to his analysis of performance statistics of minorities and whites on the 1970 and 1972 Sergeant's Promotional Examinations.

On July 2, 1974, at the conclusion of all the evidence and following the submission of briefs and oral argument, Dr. Henderson was recalled as a court witness, with no party objecting, each party given the right of cross examination. Prior to his recall he was provided with certain materials relevant to the 1972 Sergeant's Promotional Examination and the questioning was substantially limited to that area.

- 6 Descriptions of the procedures of criteria (predictive or concurrent), construct and content validation appear in Standards for Education and Psychological Tests and Manuals of the American Psychological Association, EEOC Guidelines, and recent cases, *e.g.*, *Vulcan, supra* at 394.
- 7 Dr. Richard S. Barrett has his Ph.D. in industrial psychology. He presently directs the laboratory of psychological studies and the applied psychology division of the Department of Management Sciences at Stevens Institute of Technology. He submitted a list of 32 fair employment practice cases in which, as of April 1974, he was given testimony (on deposition, in trial, or by affidavit).
- 8 Before going to work for McCann Associates in August 1965, Mr. Howeth had received a Bachelor's degree in industrial psychology at Penn State University in 1960, and between 1960 and 1965 he had held several positions in personnel work for the City of Fort Lauderdale, Florida, his last job being that of assistant to the personnel director. During his last two years he was "responsible for all police and fire testing for the city," including the preparation of examinations.
- 9 Personnel administrator since January 1972, Mr. Loeb had previously served for two and a half years as assistant director of law in the City Law Department, assigned to Civil Service matters and the Safety Department.
- 10 Howeth stated that after joining McCann and prior to his preparation of the Cleveland police and fire promotional examinations he "conducted numerous studies of the rank of police sergeant." This involved interviewing sergeants, ascertaining their duties and responsibilities, and observing them in the performance of their duties.
- 11 George W. O'Connor has spent his entire adult life in the area of criminology and police administration. Highlighting these past 20 years, he was employed as a police officer in Oakland, California; he was director of training, Chicago Police Department, Chicago, Illinois, from January 1961 to August 1, 1962. From August 1962 to September 1969 he was Director, Professional Standards Division, International Association of Chiefs of Police, Washington, D.C. After two years with the police programs division of LEAA, he served as Director of Public Safety, City of Cleveland, from December 7, 1970 to September 1971. Prior to becoming Commissioner of Public Safety, City of Troy, New York, he was director of the criminal justice program, Center of Urban Studies, University of Miami, Coral Gables, Florida.
- 12 The complete roster of sergeants and lieutenants was divided according to the units of the department to which they were assigned. The names of the sergeants and lieutenants in each unit were placed in envelopes. By random, names were then drawn from each of the envelopes. By telephone Mr. Faragher, FOP president, contacted each person chosen and asked if he would agree to be interviewed by George W. O'Connor. Each was told that the FOP was conducting a validation study of either the 1972 Sergeant's or Lieutenant's Promotional Examination. Each was aware that the exams were being challenged.
- 13 The Cleveland item analysis was never asked for nor seen by the CSC. During Howeth's testimony the court ordered its production. McCann prepared the item analysis by taking the examinations of every fifth test taker (test scores of 140 test takers) and by computing the individual performances upon each item. Test takers were divided into quartiles based upon overall performance on the exam. The item analysis identified the manner in which the members of each quartile responded to each item.
- 14 The discriminating quality of each item—that is, the ability to discriminate among the performances of individuals on each test item—was rated by McCann. Howeth explained that if everyone got the correct answer for a particular item then that item would not be "discriminating between the better and poorer candidates."
- 15 Though not done before the Cleveland examination, McCann is not performing racially differential item analyses of test results. When asked whether such an item analysis could be done of the Cleveland exams "if you were supplied with racial identification" of "examinees," Howeth answered, "Certainly." The City defendants have not to date requested or hired Howeth to perform such an item analysis.
- 16 Test reliability is calibrated by social statisticians on a 0 to 1. scale.

- 17 The three-year requirement for the sergeant's examination and the one-year requirement for each promotional test for the higher officer ranks in the CPD are specifically set forth in the CSC announcements for the respective examinations under "Minimum Qualifications for Entrance to the Examination." The CSC and Safety Department have apparently enforced these requirements for many years, even though the CSC has no express rule so providing.
- 18 Plaintiffs do not challenge the job relatedness of the time-in-grade requirements.
- 19 The Court's examination of the eligibility list for the rank of sergeant indicates a point spread of 53 points, however for purposes of considering the conclusions reached by Dr. Henderson, this difference of five points is not particularly meaningful.
-