



AGE- AND GENDER- ADJUSTED
PHYSICAL FITNESS STANDARDS:

Misperceptions, Myths, and Reality

Statement of the problem:

There continues to be confusion in the law enforcement profession regarding the legality of age- and gender-based fitness standards. Specifically, the controversy revolves around the question, “Are the Cooper Institute for Aerobics Research (CIAR) age- and gender-based fitness norms acceptable for application as standards for law enforcement agencies?” **The short answer is no.**

This position paper summarizes how this controversy developed, and presents our conclusions on this issue. This position is a consensus of FitForce™ and its associates – Fitness Intervention Technologies and Hoffman & Associates..

Assumptions:

1. There are essential functions inherent to the law enforcement profession that require some level of physical fitness/readiness.
2. In general, law enforcement agencies desire that their officers can perform those essential functions at a minimum level of safety and effectiveness.
3. Likewise, agencies desire to reduce their exposure to legal liability.
4. Therefore, agencies strive to strike a balance between doing the right thing and being fair to their employees.

Facts bearing on the problem:

1. Former CIAR position.

Beginning in 1979, the CIAR collected physical fitness test data on thousands of men and women. The CIAR organized the data into normative tables for each test (1.5 mile run, one-repetition maximum bench press ratio, one-repetition maximum leg press ratio, sit up, push up, sit and reach, and body composition). Test scores were displayed from the best on each event (99th percentile) to the worst (1st percentile). Those tables enabled men and women to compare their performance on each test to others of their age and gender.

At the time the CIAR began training law enforcement Fitness Instructors across the country, agencies were struggling to develop fitness tests and standards acceptable to the courts. Officers attending the Physical Fitness

Specialist courses learned about the CIAR norms, and believed they were a good solution to the problem.

Agencies asked the CIAR for a recommendation on what point on the normative tables to apply as a fitness standard. The CIAR considered the following points in formulating a recommendation:

- Fitness is important for job performance, health, and as a lifestyle area to prevent disability.
- Just how much fitness is necessary to do the job was an unknown.

In the absence of data documenting the specific amount of fitness necessary to do the job, the CIAR staff concluded it was reasonable to expect officers to be as fit as the general population. The CIAR suggested using the 50th percentile (half of the people tested scored above that point, and half scored below), based on each officer's age and gender. The CIAR recognized that using those norms as standards was **not** job related. There are six additional points to make about the CIAR norms:

- First, they represent a summary of how people tested by the CIAR scored on each of those tests. Those tests had nothing to do with law enforcement fitness, nor were the scores ever validated as being predictive of the ability to perform the essential physical functions of law enforcement jobs at a minimum level of safety and effectiveness. The CIAR norms are based on patients coming to the Cooper Clinic for medical exams. There are no law enforcement officers in their sample.
- Secondly, the data in the 1.5 mile run tables did not actually come from performances on the 1.5 mile run. The majority was "reverse engineered" from the results of VO2 max treadmill tests. The 1.5 mile run is a field test used to predict VO2 max. Researchers developed a regression equation to estimate a person's VO2 max using their 1.5 mile run time. Staffers at the CIAR plugged the measured VO2 maxes into that equation to estimate the 1.5 mile run times.
- When the CIAR staff came to realize that agencies were putting themselves in jeopardy of litigation by using age- and gender-based standards, they collapsed all of the data into a single database. That was interesting, but using the 50th percentile on those norms as a standard is no more valid than randomly picking scores to serve as standards.
- Dr. Tom Collingwood, using test results from validation studies he performed, compiled a table of single, law enforcement norms. Again, while interesting, there was no study conducted to validate the 50th or any other percentile as predictive of the fitness levels necessary to perform law enforcement essential tasks.

- The CIAR continued to receive inquiries from law enforcement agencies who thought the CIAR could provide validated fitness standards. The CIAR staff asked Dr. Collingwood what he was finding in his validation studies. Dr. Collingwood provided the CIAR with ranges of standards he had validated for several agencies to use as “examples”, not as recommended standards. The CIAR has included those examples in a position paper on their web site. Many agencies have incorrectly concluded that the CIAR had validated those scores as standards. That is not true. For an agency to use any of those standards they would have to conduct a transportability validation study
- Prior to 1991, there were no legal restrictions to applying different standards for groups of people performing the same job. It was illogical, but not illegal.

2. Current FitForce™ and Associates’ position

We take the position that having different standards for groups of people performing the same job is illogical. We further believe that Section 106 of the Civil Rights Act of 1991 and the Americans with Disabilities Act make such standards illegal.

Our position changed in 1991 based upon both scientific and legal requirements (Civil Rights Act of 1991 and the ADA). FitForce™, Fitness Intervention Technologies, and Hoffman & Associates have performed numerous validation studies to determine the physical fitness/ readiness levels necessary to perform the strenuous physical tasks of the job. We base our current position on what we have learned through those validation studies, as well as our interpretation of the current legal landscape:

- We have developed accepted research methodologies to determine the minimum amount of fitness necessary to perform strenuous physical tasks. The data used for analysis is specific to the agency.
- The issue is not physical fitness but “physical readiness” to do the job.
- Absolute (single) standards predict who can and who cannot perform the strenuous physical tasks of the job. **Age- and gender-based norms do not.**
- The literal interpretation of Section 106 of the Civil Rights Act of 1991 prohibits the use of gender-based standards.
- The ADA emphasizes that all employment standards must be job related. Therefore, criterion validated absolute standards are in compliance (because they predict job task performance) and age- and

gender-based standards are not (because they do not predict job task performance).

- Unless agencies develop job descriptions based on age and gender, the most rational way to conceptualize the job is "same job- same standard". Therefore absolute standards are the most logical approach.

3. Test validity requirements

The purpose of any test is to discriminate. In the case of a physical fitness/readiness, the agency wants to discriminate between those (applicants, recruits and incumbents) who can and those who cannot perform essential job functions at a minimum level of safety and effectiveness.

Test validity based on data is the best way to ensure the integrity of a test. It is highly unlikely that standards based on opinion ("we think officers should be able to do at least 25 pushups"), convention ("we have always used the CIAR norms"), or a political agenda ("we must lower the standards to include more women in the work force") will be legally defensible. None of those approaches will result in tests and standards that are job related.

The ability to defend a fitness test score as a job-related physical fitness/readiness standard hinges on the integrity of the entire validation. There are three priorities for that validation process:

- **Practical** - Standards must relate to the accomplishment of the agency's mission. In other words the standard must predict with some degree of accuracy who can and who cannot perform the job or complete training.
- **Scientific** – The EEOC Uniform Guidelines of 1978 identify three acceptable methods for validating test standards – content, construct, and criterion. Opinion, convention, and political agenda don't meet the requirements of any of those validation methods.
- **Legal** - The standard must meet the requirements of Title VII of the Civil Rights Act of 1964 (be job related if it produces disparate impact on a protected class), the Americans with Disabilities Act (represent the minimum job-related requirement), and Section 106 of the Civil Rights Act of 1991 (cannot be different based on gender). The key considerations are avoiding discrimination by gender (CRA64, CRA91) and disability (Americans with Disabilities Act). But even if there is disparate impact, the standard can be enforced if it is job related and consistent with business necessity -(*USDOJ v. SEPTA, U.S. Court of Appeals, 3rd Cir. 2000*).

From a test validity perspective the first priority is the most important. The second priority provides a methodology, and the third serves as a "check" in the system. In other words, once a "mission accomplishment" standard has been developed using acceptable methods of validation, it must be reviewed to be sure it is in compliance with the law.

Most exercise scientists who perform validation studies in public safety agree that "same job-same standard" is the correct approach. The dissenting opinions come either from those who are not exercise scientists or those in the field who have little or no experience in validation testing to predict job-performance capability.

If an agency using age- and gender-based standards is challenged, there is strong data based evidence that those standards are not job related. The test and standard would not have any validity and, as a consequence, would not be defensible.

5. Former U.S. Department of Justice position

A major contributor to this controversy has been an interpretation of statements made by members of the DOJ's Civil Rights Division. Prior to 2001, the DOJ expressed an opinion that the use of age- and gender-based fitness standards may be acceptable under certain circumstances.

DOI proposed that, for selection purposes only, age- and gender- based standards may be an acceptable alternative to validated absolute standards. Their rationale was that age- and gender-based standards produce less disparate impact and allow more females into the job, which was their stated objective. In other words it was an agenda based position rather than one based on test validity.

The DOJ agenda based position appeared to rest on their interpretation of the term "employment related tests" and a specific interpretation of Section 106 of the Civil Rights Act of 1991. At the time their opinion was that the law only applied to "job related" standards. They contended that standards used only as predictors of fitness, with no claim of job relatedness, could be age- and gender-based. But in reality any test that is used to include or exclude individuals is a de facto "employment related test".

Their interpretation additionally contended that age- and gender-based standards, which are relative, are the same standards for all groups because the percentiles are the same. The same **relative** percentile rank for a norm group

(e.g. the 50th percentile for a 20-29 year old male versus the 50th percentile for a 30-39 year old female) is not the same test score. That is illogical in that the raw scores (actual test performance) represented by those percentile scores are different.

The issue had been crafted by the DOJ as a gender issue. However, in our opinion, it **is a training issue**. The differences between males and females are irrelevant for doing the job if one applies the "same job - same standard" rationale. Within law enforcement there are not separate standards for firearms, driving, typing, conducting breathalyzers, using a baton etc. Physical fitness/readiness to do the job should not be any different. If a job requires a certain level of physical ability, then one's gender or age is immaterial. What is of concern is whether the physical ability required is out of the range for a class of employees (such as females) to achieve. Research clearly has shown that the absolute standards validated as being predictive of the minimum levels of fitness necessary to perform essential functions are attainable by anyone healthy enough to train, regardless of age or gender.

The other problem with DOJ's interpretation, as we saw it, is that even if an agency claimed that an applicant standard was not "job related" it would function as a "job related standard" because some people would be excluded from the job. If challenged, the criteria for exclusion or inclusion would not have any data to support its job relatedness. Section 106 of the Civil Rights Act clearly states different standards are not allowed for employment related standards:

"It shall be unlawful employment practice for a respondent, in the connection with the selection or referral of applicants or candidates for employment or promotion, to adjust the scores of, **use different cutoff scores for**, or otherwise alter the results of, employment related tests on the basis of race, color, religion, sex or national origin."

The following scenario illustrates the problem with age and gender-based standards that functionally act as job related selection standards.

An agency has validated 27 pushups as an absolute standard for recruits and incumbents. The agency has decided to use relative standards for applicants, and requires applicants to meet the 50th percentile for their age and gender. A 25 year old male does 29 push ups, which does not meet the applicant standard of 33 (the 50th percentile for 20 – 29 year old males). That individual would be denied employment even though he would have been able to pass the job-related fitness level validated for recruit academy graduation and incumbents. On the other hand, a 25 year old female who did 20 push ups would be accepted, even though she scored less than the job related standard. Needless to say, this scenario presents huge potential for litigation problems.

The other laws affecting this issue (ADA, Civil Rights of 1964) clearly state that employee standards must be job related.

6. Current DOJ position

Since 2001 DOJ has reversed its position, causing even more confusion. Based upon several conversations with DOJ attorneys over the past five years, we believe they now contend that validated absolute standards are not only acceptable, but are in fact the **only ones that are defensible for applicant, training or incumbent standards.**

In conversations with the Section Chief of the DOJ's Civil Rights Division, he made clear that the DOJ is focused on agencies having standards that are job related. That is their priority, not any agenda based purpose such as diversity. He reviewed a booklet we had prepared called "Absolute vs. Relative Fitness Standards for Law Enforcement", and pronounced it to be "right on the money."

At a State Police Hiring Summit in San Antonio in 2005, the DOJ Section Chief was asked about the confusion regarding absolute versus relative standards. He answered that to be compliant with Title VII of the Civil Rights Act of 1964, an agency must apply the "same job, same standard" approach.

Conclusions

1. The **purpose** of a physical fitness/readiness standard is to provide reasonable assurance that applicants, recruits and incumbents are physically able to train and perform the strenuous essential tasks of the job.
2. There are three considerations when validating fitness/readiness standards: practical, scientific, and legal.
3. Opinion, convention, and agenda are not enough to defend a fitness/readiness standard as being job related. There must be data to document job relatedness.
4. To qualify as a standard, a test score must predict the ability to perform the job at a minimum level of safety and effectiveness. The standard must discriminate between who can and who cannot perform at that minimum level. However, a "more is better" approach cannot be used when setting fitness/readiness standards.
5. **The bottom line is that only absolute fitness/readiness standards that have been scientifically validated as predictive of the levels of fitness required by the job are job related, legally defensible and logical.**

About FitForce™

FitForce™ is committed to finding Total Fitness Solutions for the public safety community we serve. Our pledge is to provide our clients with the very best scientific, legal and practical training and education, validation of selection and retention standards, policy and procedure analysis and development, as well as ongoing administration, arbitration and litigation support.

FitForce™ and its consultant team together have over 60 years of public safety physical fitness experience. This includes: over 100 articles, columns, chapters, books and technical reports, experience with over 200 agencies and their representatives, over 130 validation studies and a database of over 4000 randomly selected law enforcement officers at agencies for whom we've developed standards. *If you would like to discuss how we could be of assistance to your agency, please call us at 978.745.3629.*