"The decision of the Court of Appeal in Bauer is problematic from an equality perspective. By adopting an equal burden analysis, the traditional focus in pre-employment examination claims has shifted from the necessity of a test to its potentially discriminatory impact (…)

The difficulty of the Court of Appeal's ruling is that it applies at the initial stage of enquiry: where gendered fitness norms are applied equally within gendered groups no case of discrimination can be made. Whether such a test is justifiable in and of itself is of no concern. And whilst this may, in some cases, lead to better representation of women, it is also laden with difficulties, with the potential to perpetuate, rather than correct, inequality."

Sam Barnes

# *Physical Fitness and Gender Discrimination:* Entrenching Stereotypes

## Case Note: *Bauer v Lynch*, No. 14-2323 (4th Cir. 2016)

Sam Barnes[1]

The decision of the Court of Appeal for the Fourth Circuit in *Bauer* represents the latest solution to a complex issue in equality law: how best to accommodate physical difference. Overturning a judgment of the District Court, the Court of Appeal found that where men and women are held to comparable standards of fitness, no Title VII[2] discrimination claim arises. In this author's analysis, the approach of the Court operates against countervailing trends in American jurisprudence: serving to entrench, rather than remove, discriminatory practices in the field of employment. Where a particular standard of physical fitness is required, that standard ought to be applied equally irrespective of age, sex or race. Where a level of fitness is required as a business necessity, any test of physical capacity ought to be tailored towards the specific demands of the job.

### 1. Facts

To become a special agent with the Federal Bureau of Investigation (FBI), candidates are required to undergo an intensive 22-week training programme at the FBI Training Academy in Quantico, Virginia. Graduation from the Academy is dependent on the successful completion of a physical fitness test (PFT) consisting of timed runs, push-ups and other markers of physical performance. Several cases have been brought to American courts challenging PFTs on account of their disparate impact on women.[3] The FBI sought to address this problem by normalising testing standards between genders. Following an extensive study, minimum requirements were identified and implemented. As accepted by the Court, passing rates between men and women were statistically insignificant under the new test.

1 Sam Barnes is a Legal Research Intern at the Equal Rights Trust. The comment in this note is the author's own and does not necessarily represent the views of the Trust.

2 Specifically, Bauer alleged that the Federal Bureau of Intelligence's gender-normed physical fitness test (PFT) contravened the prohibition on discrimination contained in §2000e-16(a) and §2000e-2(1) of the Civil Rights Act of 1964, (Title VII) 78 Stat. 253, 42 U. S. C..

3 See Hollar, D., Physical Ability Tests and Title VII, *The University of Chicago Law Review*, Vol. 67, No. 3, 2000, pp. 777–803.

| FBI Minimum Physical Fitness Standards | | |
|---|---|---|
| **Event** | **Men** | **Women** |
| Sit-ups | 38 | 35 |
| 300-Metre Sprint | 52.4 Seconds | 64.9 Seconds |
| Push-ups | 30 | 14 |
| 1.5 Mile Sprint | 12 Minutes, 42 Seconds | 13 Minutes, 59 Seconds |

Despite passing the initial fitness screening required in order to access the Academy training programme, in July 2009, Jay Bauer was forced to resign his position having failed to meet the 30 push-up minimum requirement. He filed a civil action against the FBI, arguing that the gender-normed fitness test contravened §2000e-16(a)[4] and §2000e-2(1) of the Civil Rights Act 1964. Under §2000e-2(l) different cut-off scores in employment related tests on the basis of race, colour, religion, sex, or national origin are barred. Under §2000e-16(a), discrimination by federal employees on the grounds of race, colour, religion, sex, or national origin is similarly prohibited.[5] The District Court granted summary judgment in Bauer's favour, and the case was appealed to the Court of Appeal for the Fourth Circuit.

## 2. Decision

According to *Manhart*,[6] "facial" sex discrimination will be apparent "where the evidence shows treatment of a person in a manner which, but for that person's sex, would be different."[7] Ascribing § 2000e-2(a)(l)[8] its plain, unambiguous meaning, it was clear to the District Court that the differential fitness standards imposed on men and women on account of their sex were unlawful: there is no exception for innate physiological difference. Congress "was clearly aware of any such average physiological differences" but chose not to accommodate them.[9]

The Court of Appeal rejected this analysis. The "but for" test, set out in *Manhart*, does not address whether gender-normalised standards treat men differently from women. In *Gerdom*,[10] the Court of Appeal for the Ninth Circuit recognised that where "no significantly greater bur-

4 In the District Court, Bauer incorrectly relied upon § 2000e-2(a) which concerns discrimination in the private sector. The Court of Appeal, correctly applying §2000e-16(a) treated the two provisions as comparable "with the liability standards governing the former being applicable to the latter".

5 See above, note 2.

6 *City of Los Angeles, Dep't of Water & Power v Manhart*, 435 U.S. 702, 711 (1978).

7 *Ibid*.

8 See above, note 4.

9 *Bauer v Holder*, 25 F. Supp. 3d 842 (E.D. Va. 2014).

10 *Gerdom v Continental Airlines, Inc*., 692 F.2d 602 (9th Cir. 1982).

den" was imposed on either sex, a higher maximum weight limit for men than for women could be justified. Likewise, decisions of the Supreme Court would support the view that gender-normed fitness tests may be compatible with Title VII. In *VMI*,[11] whilst stressing that generalisations could not be used to exclude female candidates from the Virginia Military Institute, the Supreme Court recognised that some differences between men and women "were real, not perceived" and could, therefore, "require accommodations".

Of the few Title VII cases directly addressing the issue of gender-normed fitness standards, none have found a difference in treatment unlawful. In *Powell*,[12] the District Court was required to assess whether an FBI physical fitness test unlawfully discriminated against male applicants. Recognising that physiological differences result in males and females performing differently in PFTs, the Court concluded it did not. The PFT in question merely took account of existing biological differences. In *Hale*,[13] the Equal Employment Opportunity Commission made a similar finding. Hale complained that the FBI "held females to less rigorous physical requirements than males", violating the prohibition on sex discrimination. The administrative court judge found that distinctions based on physical differences between men and women are not necessarily discriminatory in nature.

Having analysed the above case law, the Court of Appeal vacated the previous judgment and remanded the case back to the District Court. According to the Court of Appeal, men and women "are not physiologically the same" for the purposes of physical fitness tests. Whether a PFT discriminates, therefore, will depend on whether it requires men and women to demonstrate different levels of fitness. Accordingly, gendered fitness standards will not contravene Title VII provided that "an equal burden of compliance" is imposed on men and women "requiring the same level of physical fitness of each". As the FBI test purported to meet this standard, the Court concluded that the District Court had erred in its disposition of Bauer's motion for summary judgment.[14]

## 3. Comment

The decision of the Court of Appeal in *Bauer* is problematic from an equality perspective. By adopting an equal burden analysis, the traditional focus in pre-employment examination claims has shifted from the necessity of a test to its potentially discriminatory impact. Whilst this may, in some cases, increase female participation in certain fields of employment, there are inherent dangers with accepting biological difference as a legitimate ground for differential treatment. This note will concern itself with three: the risk of direct discrimination; the negative impact of sex-based generalisation; and the wide discriminatory ambit of physical fitness tests.

11 *United States v Virginia ("VMI")*, 518 U.S. 515, 540 (1996).

12 *Powell v Reno*, No. 96-2743, 1997 U.S. Dist. LEXIS 24169 (D.D.C. July 24, 1999).

13 *Hale v Holder*, EEOC Dec. No. 570-2007-00423X (Sept. 20, 2010).

14 *Bauer v Lynch*, No. 14-2323, (4th Cir. 2016), pp. 25–6.

### *a. Physical Fitness Tests*

Pre-employment exams are used by employers to determine whether a candidate can meet basic job requirements. Where it can be proven that a test has a disproportionate impact on members of a protected class, the legality of that test will be subject to review. If the claim concerns disparate treatment (as in *Bauer*), an employment test may be justified as a bona fide occupational qualification (BFOQ).[15] If the test has a disparate impact on members of a protected class, it may be justified as a business necessity. These defences have been subject to different interpretations by American courts; however, both are materially concerned with the necessity of a test itself.[16] In the absence of a valid defence, a test must be removed or altered in order to remove its discriminatory impact. If, however, no group suffers harm, a case of discrimination cannot be made. The value of the test, from an equality perspective, is immaterial.

Two types of test have been developed by employers to evaluate physical ability. The first, known as a job simulation test,[17] examines a prospective employee's ability to complete tasks that they would be required to carry out in the course of their job.[18] The second, known as a physical fitness test, aims to assess an overall level of fitness. Tests of this nature are less focused on actual job tasks, requiring a candidate to demonstrate a high level of endurance or strength. The test is designed to determine whether a candidate is physically fit, which may be beneficial for the performance of some specific duties.[19]

In *Bauer* it was emphasised by counsel that the test was adopted in order to assess a general standard of fitness – it was not designed as a job simulation test.[20] Other tasks, required as part of the training programme, would measure applicants' ability to perform special agent functions. Whilst a high standard of physical fitness would support these functions, physical

15 Alternatively, disparate treatment may be justified where there is "a strong basis in evidence to believe it will be subject to disparate impact liability". See *Ricci v DeStefano* 557 U.S. 557, 585 (2009). Under the bona fide occupational qualification (BFOQ) defence, a required job qualification must be reasonably necessary to the "essence of the business". See, Manley, K., "The BFOQ Defence: Title VII's Concession to Gender Discrimination", *Duke Journal of Gender Law & Policy*, 2009, pp. 169–210.

16 Although this defence has garnered multiple interpretations (see above, note 3), recent case law would suggest circumscribing its application to those "minimum qualifications [required] for successful performance of the job". See *Meditz v City of Newark*, No. 10-2442 (3d Cir. Sept. 28, 2011), citing a number of Supreme Court judgments. The same standard was adopted in *Lanning*, and tacitly approved by the Court of Appeal in *Bauer*. See *Lanning v Se. Pa. Transp. Auth*., 181 F.3d 478 (3d Cir. 1999).

17 This type of test is also known a physical ability/agility test.

18 Society for Human Resource Management Foundation, *Selection Assessment Methods: A Guide to Implementing Formal Assessments to Build a High-Quality Workforce*, 2005, p. 15, available at: https://www.shrm.org/about/foundation/research/documents/assessment_methods.pdf.

19 *Ibid*., p. 12.

20 *Jay Bauer v Loretta Lynch*, *Oral Argument*, 15 September 2015, available at: https://www.courtlistener.com/audio/13469/jay-bauer-v-loretta-lynch.

fitness was a benefit in and of itself: candidates in good physical shape are less frequently injured and better prepared for training exercises.

The District Court, having already established that the PFT was facially discriminatory, sought to establish whether it could be justified as a BFOQ. The Court found it could not. Bauer's final (and failed) attempt to complete the test took place in the final week of the training programme. By that point his training had effectively concluded – he would graduate injury free. Regarding the general justification provided by counsel, the District Court noted that FBI officers are not required to undertake physical fitness tests once they have graduated from the academy. Although such tests are encouraged (albeit at a lower required passing rate than for Academy graduates), failure – either to undertake or successfully complete such a test – is inconsequential. The Court therefore concluded that the FBI had failed to link the PFT to any "qualifications that affect an employee's ability to do the job".[21]

### *b. The Limits of "Equal Burden"*

The decision of the Court of Appeal was ultimately concerned with gender parity. Wherever a minimum level of physical fitness or ability is required, certain groups will be more likely to pass the test, leading to a higher representation of those groups within the workforce. By accommodating biological differences, the field is equalised: men and women enjoy an equal prospect of success. There are, however, important limits to this line of reasoning:

#### *i. Direct Discrimination*

The most obvious problem with the Court of Appeal's approach is that it directly discriminates against male candidates. Men are held to a more onerous standard of fitness than women. When considering the serious doubt raised by the District Court as to the necessity of the PFT, this difference in treatment is difficult to justify. The FBI had decided against imposing the same fitness standards on serving FBI agents and other parts of the training programme would test applicant's abilities to perform actual job tasks. Without the PFT, men and women would graduate the academy with an equal rate of success, irrespective of their gender, and without the need to impose a more onerous standard on male applicants. Had the Court of Appeal decided in the alternative, it is likely that the test would have been abandoned – removing its discriminatory impact in whole. Instead, by accommodating biological difference, the necessity of the test becomes irrelevant. As long as men and women are held to comparable levels of fitness, differential treatment is justified.

Of course, not every case involving physical fitness testing will correspond to the facts of *Bauer*. Even the District Court seemed to recognise that in certain situations the adop-

21 As required under the BFOQ defence.

tion of differential fitness standards may be justified.[22] If the physical fitness of different groups directly corresponds to an ability to perform a particular role, gender normalised testing may be permissible.[23] The difficulty of the PFTs is that they rarely correspond directly to job tasks. As such, they require a high degree of scrutiny, absent under an equal burden analysis.

Many of the arguments made in defence of gendered fitness tests centre on the need for equal representation between groups in professions where women have historically been excluded. But biological difference is distinct from other arguments of gender inclusivity. International human rights law permits (and even obliges) positive action measures[24] in order to accelerate *de facto* equality between men and women. However, such measures are temporal. Differences due to biological difference are permanent in nature.[25] Similarly, a reasonable accommodation approach, (adopted by the Canadian Court),[26] does not sit easily with gendered-fitness norms.[27] Reasonable accommodation is designed to meet the needs of a *particular* individual with a particular set of personal characteristics.[28] Unlike the present case, such an accommodation will not usually give rise to a separate claim of disparate treatment.[29]

22 See above, note 9.

23 See *British Columbia (Public Service Employee Relations Commission) v. BCGSEU, [1999] 3 S.C.R. 3*, Para 76. Here, gendered aerobic standards in fire-service testing were advocated by the Canadian Supreme Court. Where a minimum aerobic standard is necessary to perform a job safely, an employer should ask "whether members of all groups require the same minimum aerobic capacity" and if not, "to reflect that disparity in the employment qualifications".

24 See Convention on the Elimination of All Forms of Discrimination against Women (CEDAW), G.A. Res. 34/180, 1979. Under Article 4(1) of CEDAW, "temporary special measures aimed at accelerating de facto equality between men and women shall not be considered discrimination".

25 And thus prohibited. See *ibid.* Interestingly, the CEDAW Committee has recognised that in certain circumstances, biological differences may require the "*non-identical treatment of women and men".* However, this is limited to measures aimed at protecting maternity. See the Committee on the Elimination of Discrimination against Women, *General Recommendation No. 25: On article 4, paragraph 1, on temporary special measures*, UN Doc. HRI/GEN/1/Rev.7, 2004, Para 8.

26 See above, note 23. On a number of occasions, the Court of Appeal justified its approach by reference to the need to "accommodate" women. Yet, even if a reasonable accommodation analysis was applied, many of the difficulties identified below, especially in relation to age discrimination, would still be present.

27 Outside of the Canadian context, few courts have recognised a duty to accommodate on the grounds of sex. See European Commission, *Reasonable Accommodation beyond Disability in Europe?*, 2013, available at: http://ec.europa.eu/justice/discrimination/files/reasonable_accommodation_beyond_disability_in_europe_en.pdf.

28 *Ibid*., p. 39.

29 Hence, whilst it may be unreasonable to prevent an employer from having their employees work on Sundays due to the religious requirements of a particular staff member, it may be appropriate to accommodate that individual's religious belief.

*ii. Gender Stereotypes*

When the issue of female representation is set aside, arguments made in support of an equal burden analysis lose their potency. Stereotypes and generalisations are salient in employment policies because they are utilised in order to discriminate between groups. Statistical evidence of difference, as a basis for differential treatment, has been rejected in a number of discrimination cases in Europe.[30] In *Lindorfer*,[31] the Court of Justice of the European Union was required to determine whether a decision of the Court of First Instance breached the principle of non-discrimination by finding that the use of variable factors accounting for differences between the male and female life expectancy were objectively justified. The Court found that it had. In the opinion of Advocate General Jacobs, it was observed:

[D]iscrimination of the kind in issue involves ascribing to individuals average characteristics of a class to which they belong (…) What is objectionable (and thus prohibited) in such discrimination is the reliance on characteristics extrapolated from the class to the individual, as opposed to the use of characteristics which genuinely distinguish the individual from others and which may justify a difference in treatment.[32]

Generalisations drawn from class characterisations (even where true) do little to evidence the ability of a particular member of that class to perform a specific task. If, as in *Bauer*, the necessity of an employment test itself is not sufficiently validated, whether individuals are held to comparatively equal standards is irrelevant. In accounting for biological difference, the use of such generalisations is justified. And this may have consequences beyond the realm of physical fitness.

Take the following example. The role of a flight attendant is one traditionally held by women. An airline introduces a maximum weight policy for airline staff. Under an ordinary claim of disparate treatment (where gendered standards were introduced), or disparate impact (where no gendered standards were introduced), the policy would fail for lack of proximity to actual job requirements. Under an equal burden test, however, a weight restriction bearing no relationship to actual job requirements could be adopted provided that the burden imposed was equally applied. In an area where men (even with equal selection policies) may be less inclined to apply, stereotyped roles such as the "skinny-female flight attendant" are entrenched, damaging long-term opportunities of gender equality even where a similar burden is placed on other groups. Equal burden, in this scenario, would contradict the established jurisprudence of the American Supreme Court,[33] who have consistently warned against the

30 See, for example, *X*, C-318/13, 3 September 2014.

31 *Lindorfer v Council of the European Union*, C-227/04, 11 September 2007.

32 *Ibid*., 27 October 2005, Decision of the Advocate General, Para 59.

33 The Supreme Court has accepted the existence of height and weight disparities between men and women, rejecting both as a marker of physical capacity due to their discriminatory impact and lack of proximity to actual job requirements. *Dothard v Rawlinson*, 433 U.S. 321 (1977).

use of "generalisations" or "tendencies" as a means to justify discriminatory treatment on the grounds of sex[34] and have rejected the use of discriminatory gender stereotypes as a legitimate ground for disparate treatment.[35]

*iii. Affected Groups*

One final issue, not yet discussed in this note, concerns the ambit of employment tests. Women are unlikely the only group adversely impacted by physical fitness requirements.[36] When the equal burden analysis is applied to other affected classes (such as older applicants), the difficulty of the *Bauer* judgment becomes clear. Although Title VII does not extend to discrimination on the grounds of age, a claim could be brought under the Age Discrimination in Employment Act (ADEA). It is important to note, under the ADEA, an employment practice that has a disparate impact on individuals of a particular age may be justified where it is "based on reasonable factors other than age".[37] This is a lower threshold than the business necessity defence established under Title VII and could, in theory, be utilised to defend a claim of disparate impact age discrimination where gender-normed fitness standards are introduced. Alternatively, applying a biological difference model to age, employers would be required to impose a bi-partite test: one capable of accounting for both age and gender disparities. In *Bauer*, no such account was given.

Aside from the practical difficulties of designing an employment test that is sufficiently measured to provide for all potentially affected groups,[38] there is a final difficulty with the *Bauer* decision. As discussed above, where a test is not necessary, an adverse impact or treatment claim will usually result in the removal (or modification) of that test in order to eliminate its discriminatory impact. Under an equal burden approach, however, any employment policy would only need to be modified to the extent that a statistical discrepancy within groups is accounted for. Other groups, similarly affected, would be required to bring separate claims in order to remove the discriminatory standard. When considering the intersection between age and gender in physical fitness, this seems unnecessarily burdensome – perpetuating inequality where a test does not relate to actual job requirements.

34 See above, note 11.

35 Even where the stereotype in question is "unquestionably true". See above, note 6.

36 Courts in Europe have accepted the potential limitations imposed by age on physical ability and fitness. See, *Colin Wolf v Stadt Frankfurt am Main*, C-229/08, 12 January 2010.

37 Age Discrimination in Employment Act of 1967, 29 U.S.C. § 623(f)(1).

38 It is difficult to envisage any PFT sufficiently developed to withstand multiple disparate impact claims. Where biological difference is accounted for, a test may be required to impose multiple standards (a younger woman; an older man; etc.) The case becomes more interesting when applied to race. Researchers have noted a positive correlation between race and performance in certain sports. See, Bejan, A. and others, "The Evolution Of Speed In Athletics: Why The Fastest Runners Are Black And Swimmers White", *International Journal of Design & Nature and Ecodynamics*, Vol. 5, 2010, pp. 199–211.

## Conclusion

Although the case has been remitted back to the District Court, it seems unlikely, following the Court of Appeal's decision, that Bauer will be successful in his claim. The use of physical fitness tests will continue to prove controversial. Irrespective of the approach taken, in certain situations it is likely that some group will be adversely affected. The difficulty of the Court of Appeal's ruling is that it applies at the initial stage of enquiry: where gendered fitness norms are applied equally within gendered groups no case of discrimination can be made. Whether such a test is justifiable in and of itself is of no concern. And whilst this may, in some cases, lead to better representation of women, it also laden with difficulties, with the potential to perpetuate, rather than correct, inequality.