
Appeal No. 22-2433

In the
United States Court of Appeals
For the Seventh Circuit
Chicago, Illinois 60604

CATHERINE ERDMAN,

Plaintiff - Appellant,

v.

CITY OF MADISON,

Defendant - Appellee.

On Appeal from the United States District Court
For the Western District of Wisconsin,
The Honorable Judge William M. Conley Presiding,
District Court Case No: 3:16-cv-00786-wmc

Brief and Short Appendix of Plaintiff-Appellant

Jeff Scott Olson
THE JEFF SCOTT OLSON LAW FIRM, S.C.
1025 Quinn Drive, Suite 500
Waunakee, WI 53597-2502
608-283-6001
jsolson@scofflaw.com

Counsel of Record for Plaintiff-Appellant

CIRCUIT RULE 26.1 DISCLOSURE STATEMENT

Pursuant to Circuit Rule 26.1, counsel of record for the Plaintiff-Appellant, Jeff Scott Olson, hereby makes the following disclosures.

1. The full name of every party that the attorney represents in the case (if the party is a corporation, you must provide the corporate disclosure information required by Fed. R. A-P. 26.1 by completing item #3):

Catherine Erdman

2. The names of all law firms whose partners or associates have appeared for the party in the case (including proceedings in the district court or before an administrative agency) or are expected to appear in this Court.

The Jeff Scott Olson Law Firm, S.C.

3. If the party or amicus is a corporation: Identify all of its parent corporations, if any; and list any publicly held company that owns 10% or more of the party's or amicus's stock.

Catherine Erdman is not a corporation.

4. Provide information required by FRAP 26.1(b) – Organizational Victims in Criminal Cases:

Not applicable.

5. Provide Debtor information required by FRAP 26.1 (c) 1 & 2:

Not applicable.

Dated this Wednesday, March 15, 2023.

Respectfully submitted,

THE JEFF SCOTT OLSON LAW FIRM, S.C.

/s/Jeff Scott Olson

Jeff Scott Olson

ATTORNEYS FOR PLAINTIFF-APPELLANT

Appeal No. 22-2433

In the
United States Court of Appeals
For the Seventh Circuit
Chicago, Illinois 60604

CATHERINE ERDMAN,

Plaintiff - Appellant,

v.

CITY OF MADISON,

Defendant - Appellee.

On Appeal from the United States District Court
For the Western District of Wisconsin,
The Honorable Judge William M. Conley Presiding,
District Court Case No: 3:16-cv-00786-wmc

Brief of Plaintiff-Appellant

Jeff Scott Olson
THE JEFF SCOTT OLSON LAW FIRM, S.C.
1025 Quinn Drive, Suite 500
Waunakee, WI 53597-2502
608-283-6001
jsolson@scofflaw.com

Counsel of Record for Plaintiff-Appellant

TABLE OF CONTENTS

TABLE OF CONTENTS.....i

TABLE OF AUTHORITIESiv

JURISDICTIONAL STATEMENT 1

ISSUES PRESENTED FOR REVIEW..... 2

STATEMENT OF THE CASE..... 3

Facts 3

I. Most of the Facts Were Stipulated by the Parties in the Joint Pretrial Statement..... 3

 A. The 2014 Recruitment and Testing Process. 3

 B. Background and Validation of the Madison Fire Department Selection Process..... 5

 C. Plaintiff Catherine Erdman’s Application. 12

 D. The Candidate Physical Ability Test. 13

 E. The Disparate Impact of the Madison Physical Abilities Test. 15

II. The District Judge Made Additional Findings After the Trial. 19

 A. The Madison Fire Department 2014 Selection Process. 19

 B. The Candidate Physical Ability Test 20

 C. Comparison Between the CPAT and the Madison Fire Department Physical Abilities Test in terms of Adverse Impact on Women..... 20

III. Still Additional Facts were Established at Trial by Uncontradicted Evidence. 21

 A. Males and Females. 21

B. Catherine Erdman.	22
C. The Madison Fire Department.....	23
IV. The Present Litigation	24
SUMMARY OF ARGUMENT.....	26
ARGUMENT	29
I. The Standards of Review	29
A. The Standard of Review of a Summary Judgment Decision is <i>De Novo</i> . 29	
B. The Standard of Review of a Trial Court’s Decision after a Bench trial Depends on what is Being Reviewed.	30
1. The “Clearly Erroneous” Standard of Review Applies to Findings Entered after a Bench Trial.	30
2. Conclusions of Law are Reviewed <i>De Novo</i>	31
3. Findings on Mixed Questions of Fact and Law are Reviewed with an Appropriate Degree of Deference.	32
II. A reasonable trier of fact could conclude that but for the unlawful disparate impact of the PAT, the Plaintiff would have been hired, and if such a finding is made, she will be entitled to back pay, front pay, and reinstatement into the position at issue.	33
A. A reasonable trier of fact could find that Ms. Erdman is entitled to back pay and reinstatement into the job at issue.....	33
1. This Court Should Apply the Ordinary Civil “Preponderance of the Evidence” Standard of Proof.	34
2. If Ms. Erdman Proves that the PAT had an Unlawful Disparate Impact, She will be Entitled to a Trial on her Claim for Make-Whole Relief. 36	
B. Ms. Erdman will be entitled to injunctive relief.....	39
III. Even if the City showed that the test was a job-related business necessity, the Plaintiff carried her third-stage burden of showing that the City had	

considered and rejected an alternative test with substantially equal validity and a much smaller adverse impact on women. 40

IV. The nature of the Plaintiff’s stage-three burden of proving that an available alternative selection practice (1) results in a less disparate impact, and (2) serves the employer’s legitimate needs is to identify a substantially equally valid alternative selection procedure with a smaller gap between passing rates. 42

 A. Relevant Parts of the Uniform Guidelines..... 44

 B. A Plaintiff succeeds in her stage three burden if she identifies an alternative test that serves the employer's legitimate interest in efficient and trustworthy workmanship. 44

V. Ms. Erdman Proved that the CPAT has a Much Smaller Disparate Impact on Women than the PAT and is Substantially Equally Valid. 46

 A. The City Considered and Rejected the CPAT. 46

 B. The CPAT has a much smaller disparity in passage rates by sex. 48

 C. The CPAT is at least substantially equally valid. 54

 D. The District Court’s Reasons for Finding the CPAT not Substantially Equally as Valid as the Madison Test do not Withstand Scrutiny..... 61

 1. A Test’s Exact Replication of Job Tasks Cannot Support a Finding of Greater Validity if those Skills Can be Learned on the Job. 61

 2. The need to perform a transferability study. 67

 3. The PAT having been a good predictor of outcome historically, as defined by a high passage rate out of the academy 70

 4. The Department’s comparatively high percentage of female firefighters..... 71

CONCLUSION..... 73

Rule 32(a)(7) CERTIFICATION 75

FORMER CIRCUIT RULE 31(e) CERTIFICATE 75

CERTIFICATE OF SERVICE..... 76

TABLE OF AUTHORITIES

CASES

Albemarle Paper Co. v. Moody, 422 U.S. 405 (1975)..... 45

Allen v. City of Chicago, 351 F.3d 306 (7th Cir. 2003)..... 44

Anderson v. City of Bessemer City, N.C., 470 U.S. 564 (1985)..... 19

Anderson v. Liberty Lobby, Inc., 477 U.S. 242 (1986)..... 30

Bose Corp. v. Consumers Union of United States, Inc., 466 U.S. 485 (1984)..... 32

Celotex Corp. v. Catrett, 477 U.S. 317 (1986)..... 30

City of Los Angeles v. Lyons, 461 U.S. 95 (1983)..... 40

Ernst v. City of Chicago, 837 F.3d 788 (7th Cir. 2016) 55, 59, 71, 72

Evans v. City of Evanston, 881 F.2d 382 (7th Cir. 1989) 34, 35

G. M. Enterprises, Inc. v. Town of St. Joseph, 350 F.3d 631 (7th Cir. 2003) 29

Griggs v. Duke Power Co., 401 U.S. 424 (1971)..... 43

Horan v. City of Chicago, No. 98 C 2850, 2003 WL 22284090 (N.D. Ill. Sept. 30, 2003)
..... 59

Lanning v. Se. Pa. Transp. Auth., 308 F.3d 286 (3d Cir. 2002) 55, 59, 72

McDonnell Douglas Corp. v. Green, 411 U.S. 792 (1973) 45

Moriarty v. Glueckert Funeral Home, Ltd., 155 F.3d 859 (7th Cir. 1998) 32

Outlaw v. Newkirk, 259 Fed.3d 833 (7th Cir. 2001) 30

Palmerton v. Wal-Mart Stores, Inc., No. CV 17-30-H-CCL, 2017 WL 4031432 (D.
Mont. Sept. 13, 2017)..... 35

Ricci v. DeStefano, 557 U.S. 557 (2009) 43

Tolan v. Cotton, 572 U.S. 650 (2014)..... 30

U.S. Bank Nat. Ass'n ex rel. CWC Capital Asset Mgmt. LLC v. Vill. at Lakeridge, LLC,
138 S. Ct. 960 (2018)..... 32

United States v. United States Gypsum Co., 333 U.S. 364 (1948)..... 31

Walton v. Jennings Cmty. Hosp., Inc., 999 F.2d 277 (7th Cir. 1993) 32

Wards Cove Packing Co. v. Atonio, 490 U.S. 642 (1989) 42

STATUTES

28 U.S.C. § 1291 1

42 U.S.C. § 2000e 1
 42 U.S.C. § 2000e-5(f)(3) 1
 42 U.S.C.A. § 2000e-2 (k)..... 42

RULES

Fed. R. Civ. P. 52 30
 Fed. R. Civ. P. 52(a)(6) 19
 Fed. R. Civ. P. 56(c)..... 30

TREATISES

Wright & A. Miller, Federal Practice & Procedure 32

REGULATIONS

29 C.F.R. § 1607.3(B) 25, 44

JURISDICTIONAL STATEMENT

The district court had jurisdiction of this action, which asserted claims under the Title VII of the Civil Rights Act of 1964, as amended, 42 U.S.C. § 2000e, et seq., pursuant to 42 U.S.C. § 2000e-5(f)(3). The Western District of Wisconsin was a proper venue for this action because it is a judicial district in the state in which the unlawful employment practice is alleged to have been committed within the meaning of 42 U.S.C. § 2000e-5(f)(3).

The Seventh Circuit Court of Appeals has jurisdiction of this appeal pursuant to 28 U.S.C. § 1291.

The case was tried to the court without a jury on October 15 and 16, 2018, and the district court entered an Opinion and Order (dkt. # 81) dismissing the Plaintiff's claim on July 20, 2022. The court entered judgment dismissing Plaintiff's claim on its merits (dkt. # 82) on the same day, July 20, 2022.

No motion for a new trial, nor for alteration of the judgment, nor any motion claimed to toll the time within which to appeal, has

been filed since the entry of the July 20, 2022, judgment. The Plaintiff timely filed this appeal by electronically filing a Notice of Appeal in the district court on August 11, 2022. There have been no prior or related appellate proceedings.

ISSUES PRESENTED FOR REVIEW

Was there a material issue of fact for trial as to whether Plaintiff Erdman would more likely than not have been selected for a City of Madison firefighter job if she had passed its physical ability test for firefighter candidates?

Did the Plaintiff prove that, in declining to adopt the Candidate Physical Ability Test promulgated by the International Association of Firefighters and the International Association of Fire Chiefs and modified as a result of negotiations between those entities and the Equal Employment Opportunity Commission, the City had rejected an available alternative test that (1) results in a lesser disparate impact, and (2) serves the City's legitimate needs?

STATEMENT OF THE CASE

Facts

I. Most of the Facts Were Stipulated by the Parties in the Joint Pretrial Statement.

Although the district court's decision does not mention it, in their Joint Pretrial Statement (dkt. # 53), the parties stipulated to most of the operant facts in the case. In material part, these stipulated facts are:

The City of Madison, Wisconsin, operates a Fire Department that employs approximately 365 firefighters. It annually engages in a recruitment process to fill vacancies. (Dkt. #53, ¶ 3.)

A. The 2014 Recruitment and Testing Process.¹

In the Job Bulletin for the Department's 2014 recruitment for firefighter positions, the physical requirements of the job were identified as follows:

¹ The headings in this section were not part of the parties' stipulation; they are inserted for the convenience of the court.

While not an exclusive list, the following examples are meant to illustrate some of the extreme physical demands and working conditions inherent in the role of a firefighter.

Physical Demands

1. Pick up and advance charged fire hoses.*
2. Force entry with axe/battering ram.*
3. Rescue/extricate victim(s).*
4. Perform CPR; apply bandages.
5. Climb stairs carrying heavy equipment, while wearing firefighter protective clothing that weighs in excess of 50 pounds.*
6. Strip and vent roofs, breach walls, overhaul burned buildings.*
7. Lift and climb/descend ladders (with victims).*
8. Visually determine fire status/hazards; assess patient conditions.
9. Hear calls for help; identify fire noise, etc.
10. Walk on roof tops under adverse conditions.
11. Operate power tools and extrication equipment; tie knots.
12. Stoop, crawl, crouch, and kneel in confined spaces.*

13. Reach, twist, balance, grapple, bend and lift under emergency conditions.
14. Run, dodge, jump and maneuver with equipment.*
15. All of the above may be performed wearing heavy and restrictive protective clothing/gear in excess of 50 pounds.

The items noted with an asterisk are tasks which the City believed were assessed in the City's Physical Abilities Test, consistently referred to as the PAT.² (Dkt. #53, ¶ 4.)

B. Background and Validation of the Madison Fire Department Selection Process.

The Department first developed and had its application process validated by Landy, Jacobs & Associates, Inc., in 1997. (Dkt. #53, ¶ 5.) In 1997, Landy, Jacobs and Associates, Inc., provided the Department with a Firefighter Job Analysis, Test Development,

² Sometimes this is called the "Physical Ability Test" and sometimes it is called the "Physical Abilities Test," even in the selfsame official document. See, e.g., Exhibit 9, at 3 of 181. Former Fire Chief Debra Amesqua calls it the "physical agility test." (Tr., dkt. # 67, 208:8, 19.) For consistency, this brief will use "Abilities."

Administration and Scoring Report, on all phases of its hiring process, including the Physical Abilities Test. (Dkt. #53, ¶ 6.)

The Department had its hiring process revalidated in 1999 by SHL Landy, Jacobs & Associates. (Dkt. #53, ¶ 7.)

The Department had its hiring process revalidated in 2003 by SHL USA, Inc./EB Jacobs, LLC. (Dkt. #53, ¶ 8.)

The Department had its hiring process revalidated in 2013-2014 by Ergometrics & Applied Personnel Research, Inc. (Dkt. #53, ¶ 9.)

The hiring selection process consisted of several stages, at each of which an applicant could be eliminated from further consideration:

1. Application screening for minimum qualifications;
2. Written Test;
3. Physical Abilities Test;
4. Oral Board; and,
5. Chief's Interview.

(Dkt. #53, ¶ 10.)

The PAT consists of seven events which the City believed simulated activities related to firefighting, including: 1) Equipment Shuttle; 2) Ladder Event; 3) Hose Drag; 4) Sledgehammer Event; 5) Search; 6) Rescue (through a doorway); and 7) Pike Pole. (Dkt. #53, ¶ 11.)

Each event in the test was assigned both a cut-score, also sometimes referred to as a passing score, and a minimally acceptable performance score, which represented a specific performance level below the cut-score. (Dkt. #53, ¶ 12.) If the candidate met or exceeded the cut-score for an event he or she received a score of 1 on that event. (Dkt. #53, ¶ 13.) If the candidate completed the event within the minimally acceptable performance standard, but did not meet or exceed the cut-score, he or she received a score of 0 on that event. (Dkt. #53, ¶ 14.) If the candidate did not complete an event within the minimally acceptable performance standard, he or she was disqualified and eliminated from the selection process. (Dkt. #53, ¶ 15.)

In order to progress from the PAT to oral boards, candidates had to a) complete all seven events within the minimally acceptable performance standards and b) receive an overall score of at least 5 (e.g., he or she had to meet or exceed the cut-scores for 5 of the 7 events). (Dkt. #53, ¶ 16.)

In its 1997 validation study, ninety-four incumbents (currently employed firefighters) participated in the PAT to establish cut-scores for firefighter candidates. (Dkt. #53, ¶ 17.)

In 1999, some of the PAT events were modified and new minimally acceptable performance standards and cut-scores were developed. In order to do this, 102 incumbent firefighters completed the modified PAT. (Dkt. #53, ¶ 19.)

The PAT component of the selection procedure was not altered from 1999 to 2011. (Dkt. #53, ¶ 20.)

In 2013-2014, Ergometrics performed a content validity study with respect to the test content and job demands. This involved obtaining data from agencies from around the country which were compared with a similar study of the Madison Department's

firefighters' job duties. This study confirmed that the job duties of a Madison Department firefighter were highly correlated with those of other departments. (Dkt. #53, ¶ 21.)

Based on Ergometrics 2014 validation report, the 1999 cut-scores were maintained and two changes in testing procedure were made: 1) increased standardized recovery times between events; and 2) the use of weighted vests. (Dkt. #53, ¶ 26.)

The 2014 PAT contained following events:

1. Stair climb/stepmill: This event required an applicant to step continuously on a step mill for four minutes and twenty-eight seconds with weighted vests. The disqualification time was one minute and thirty seconds.
2. Ladder: This event required an applicant to remove a ladder from the wall, rotate it as the applicant crosses the room, and place it flat on the floor within a marked area. The applicant then was required to go to a wall mounted ladder, stand in a marked box and raise, then lower, the ladder using a hand-over-hand motion. The applicant then had to go back to the other ladder and return it. The passing time was ninety seconds and the disqualification time was two minutes and five seconds.

3. Hose Drag: This event required an applicant to pick up the end of a hoseline and move through a course. An applicant was required to complete the course pulling the hose in forty-nine seconds or less. The disqualification time was ninety seconds.
4. Sledgehammer: This event required an applicant to successfully strike a raised target with a sledgehammer twelve times. To pass, an applicant was required to do this in fourteen seconds or less. The disqualification time was thirty seconds.
5. Search: This event required an applicant to move through a maze in a crouched position in darkness. To pass, an applicant was required to complete the event in one minute and thirteen seconds. The disqualification time was one minute and forty-five seconds.
6. Rescue: This event required an application to drag a mannequin through a lighted maze. To pass, an applicant was required to complete this task within fifty seconds or less. The disqualification time was ninety seconds.
7. Pike Pole: This event required an applicant to pick up a pike pole and push the door (fifty pounds) of the ceiling up until the indicator light went on and return the pole to its sleeve. The applicant then had to take a second pole, stand back 18 inches, and pull the pole five times so

that the indicator light went on and a weight (eighty pounds) hit the bottom each time. An applicant had to do this repetition as many times as possible within one minute. The applicant then had a rest time of thirty seconds at which time the cycle occurred again. This was done for a total of four cycles. To pass, an applicant had to complete twenty repetitions. The disqualification standard was sixteen repetitions.

(Dkt. #53, ¶ 27)

Each applicant had to wear a forty-pound weighted vest throughout all seven events. (Dkt. #53, ¶ 28.)

In the 2014 hiring process, applicants were provided with access to a video of these tests and were allowed to practice prior to the actual test on the course for one hour. In addition the Department's web site contained detailed information about all of the steps of the testing process, including a video of the PAT. (Dkt. #53, ¶ 29.)

Applicants who availed themselves of the opportunity to practice on the test course were not permitted to wear the forty-pound weighted vest during their practice session. (Dkt. #53, ¶ 30.)

C. Plaintiff Catherine Erdman's Application.

In 2007, Ms. Erdman applied for and was hired into a position as firefighter/EMT-B for the City of Janesville, Wisconsin. Ms. Erdman has worked in this position since then. (Dkt. #53, ¶ 31.)

Ms. Erdman applied for a position as a Madison firefighter in 2013, seeking one of the positions that would be filled beginning in 2014. (Dkt. #53, ¶ 32.) She passed the application screening and the written test but was eliminated from the selection process when she was disqualified for failure to complete a sufficient number of repetitions in the allotted time in the last phase of the Physical Abilities Test. (Dkt. #53, ¶ 33.) Ms. Erdman passed five of the seven components of the PAT. (Dkt. #53, ¶ 34.) Ms. Erdman did not receive a passing score on the ladder test because, while she was able to complete the test within the allotted time to avoid disqualification, she did not meet the cut score to "pass." (Dkt. #53, ¶ 35.) Ms. Erdman was eliminated from the hiring process because she failed to complete the minimum number of repetitions in the allotted time in the Pike Pole test. (Dkt. #53, ¶ 36.)

D. The Candidate Physical Ability Test.

In 2013, the Madison Fire Department was aware of the existence of another physical abilities test utilized by fire departments around the country, the Candidate Physical Ability Test (CPAT) licensed by the International Association of Fire Fighters. (Dkt. #53, ¶ 37.)³ This test was created by the International Association of Fire Chiefs' (IAFC) and the International Association of Fire Fighters' (IAFF) Joint Labor Management Task Force, and is used as a screening tool for many fire departments across the nation. (Dkt. #53, ¶ 38.)

The Candidate Physical Ability Test (CPAT) was validated as a content-based test designed for screening potential firefighter candidates. (Dkt. #53, ¶ 39.)

The CPAT contains eight component parts, seven of which are the same or similar to events in the PAT. Among the differences

³ The CPAT Manual, describing its origins and history, its elements, and the required procedure for fire departments to be licensed to use it, is in the record as Exhibit 37.

between the tests is that the CPAT requires applicants to complete all components within a designated time (10 minutes, 20 seconds) rather than individually timing the events as is done with the PAT. (Dkt. #53, ¶ 40.) Another difference in the tests involves the pike pole (referred to the CPAT as the “breach and pull”). In the CPAT, applicants are allowed to stand directly under the point to which they have hooked their pole in accomplishing the pull portion of that test. In the PAT, applicants can stand under the hook point to push up, but must move 18 inches back when completing the pull portion of the test. (Dkt. #53, ¶ 41.)

The CPAT was revalidated in 2013. (Dkt. #53, ¶ 42.)

In 2013, the Madison Fire Department was aware of the CPAT, considered it, and decided to continue to use the PAT. (Dkt. #53, ¶ 43.)

Ms. Erdman passed, twice, the Candidate Physical Ability Test (CPAT) for firefighter candidates. Ms. Erdman passed the CPAT once when she was first hired as a Janesville firefighter and once again, two weeks after failing the Madison Physical Abilities

Test. Ms. Erdman did not take part in training sessions before either CPAT. (Dkt. #53, ¶ 44.)

E. The Disparate Impact of the Madison Physical Abilities Test.

In the 2014 hiring process, a total of 1887 applicants applied (146 women and 1723 men). (City statistics do not indicate the gender of 18 of the applicants.) (Dkt. #53, ¶ 46.) As a result of the hiring process, four females and thirteen males were hired. (Dkt. #53, ¶ 47.)

471 males and 28 females appeared to take the 2014 Madison Fire Department PAT. Of these candidates, thirteen males and three females quit during the test. Thus, females quit at a somewhat higher rate, 10.71%, than males, 2.76%. (Dkt. #53, ¶ 45.)

The 2014 Madison Fire Department PAT passage rate for women who appeared to take the test was 4 out of 28 or 14%. The 2014 Madison Fire Department passage rate for men who appeared to take the test was 395 out of 471, or 84%. (Dkt. #53, ¶ 49.) Of the candidates who did not quit during the test, 49 males and 20 females

were disqualified during the test for failing to meet the minimally acceptable performance standard on a particular event. The rate at which females who did not quit during the test were disqualified during the test for failing to meet the minimally acceptable performance standard on a particular event, 80%, exceeded the rate at which males who did not quit during the test were disqualified, 10.7 %. (Dkt. #53, ¶ 50.)

Of the candidates who were not disqualified, one female and 14 males failed the test by virtue of failing to achieve passing scores on at least five of the seven events. The rate at which females who did not quit during the test failed in this manner was 4% and the rate at which males who did not quit during the test failed in this manner was 3.06%. (Dkt. #53, ¶ 51.)

The gender difference in the percentage of disqualifications among those who actually showed up for the test was as follows: 71.4% of women who showed up for the test were disqualified, compared to only 10.4% of men who showed up for the test. This difference is statistically significant. (Dkt. #53, ¶ 52.)

The numbers of men and women who appeared for the test, who quit during the test, who did not quit during the test, who were disqualified during the test for failing to meet the minimum performance standard on one event, who failed the test by failing to receive a passing score, and who passed the test, are set forth in this table:

2014 Test by Sex						
Category	Males			Females		
	Percentage against Males Who Appeared	Percentage against Males Who Appeared and Did Not Quit	Number	Percentage against Females Who Appeared	Percentage against Females Who Appeared and Did Not Quit	Number
Appeared to Take Test			471			28
Quit during Test	2.76%		13	10.71%		3
Did not Quit during Test	97.24%		458	89.29%		25
Disqualified for Poor Performance on an Event	10.40%	10.70%	49	71.43%	80.00%	20
Failed Test	2.97%	3.06%	14	3.57%	4.00%	1
Passed Test	83.86%	86.24%	395	14.29%	16.00%	4

(Dkt. #53, ¶ 53.)

Theodore Gerber, a professor of sociology at the University of Wisconsin – Madison, was retained by the Plaintiff to analyze the statistical significance of the observed differences by gender in

successfully passing the Madison fire Department's 2014 Physical

Abilities Test. (Dkt. #53, ¶ 54.) He concluded, accurately:

. . . among the total number of original applicants, female applicants were disproportionately and adversely affected by each phase of attrition from the testing process: they were more likely to withdraw prior to the test (7.9% vs. 1.1%), more likely to fail to appear (18.4% vs. 10.3%), far more likely to be disqualified (52.6% vs. 9.2%), and more likely to quit during the test (7.9% vs. 2.4%). The difference between women and men with respect to failing to appear is not statistically significant. . . All three of the other comparisons reveal statistically significant differences by gender. . . Thus, the data provide statistical evidence that among the pool of original applicants, women are not only less likely than men to pass the test, but also more likely to withdraw prior to the test, to be disqualified during the test, and to quit during the test.

(Dkt. #53, ¶ 55.)

The 2014 PAT as a whole had a statistically significant disparate impact on female applicants. (Dkt. #53, ¶ 56.)

II. The District Judge Made Additional Findings After the Trial.

Since this case was ultimately decided after a trial to the court without a jury, any additional findings by the district judge must be accepted by a reviewing court unless they are “clearly erroneous.” Fed. R. Civ. P. 52(a)(6), *Anderson v. City of Bessemer City, N.C.*, 470 U.S. 564 (1985). In setting down his findings of fact after the trial in this case, Judge Conley began by noting that, in his summary judgment ruling, he had set out a large number of facts that he regarded as undisputed, and he incorporated these by reference into his post-trial decision. (A-23-24, n. 1.)

The following facts were either incorporated into the district judge’s final decision from the summary judgment decision or found by the district judge after hearing the evidence at trial:

A. The Madison Fire Department 2014 Selection Process.

In 2014, the Fire Department ultimately hired every one of the four women who passed the Physical Abilities Test but hired only

thirteen of the 395 men who successfully completed the Physical Abilities test. (A-3.)

B. The Candidate Physical Ability Test

This test was developed in conjunction with ten fire departments in North America, including Austin, Texas, Los Angeles County, California, and New York City. (A-29-30, ¶ 26.) The CPAT has not been locally validated. (A-30. ¶ 30.) The CPAT is valid generally. (A-38.) The City's expert, Charles Swander, in his role as a co-owner of a testing business, has administered the CPAT over 10,000 times. (A-38.)

C. Comparison Between the CPAT and the Madison Fire Department Physical Abilities Test in terms of Adverse Impact on Women.

According to a validation study conducted by an exercise physiologist at the University of Texas-Austin, examining the Austin Fire Department, approximately 48% of female participants passed the CPAT. (A-30, ¶ 27.) In a broader study of fire departments using the CPAT, the pass rate for women was 68.0% as compared to 49.0%

in departments using other physical ability tests. (A-38.) The CPAT has less of an adverse impact on female applicants than other tests. (A-38.)

Erdman again participated in the Fire Department's 2016 recruitment. (A-7.) Erdman still desires to work for the Department, and she intends to apply during future recruitments. (A-7.)

III. Still Additional Facts were Established at Trial by Uncontradicted Evidence.

A. Males and Females.

There are physiological, biomechanical, and fitness differences between men and women that might affect their relative performance on physical ability candidate screening tests. (Tr., dkt. # 67, 22:11-15.) On average, men are taller, weigh more, have more muscle mass, have less body fat, and have more hemoglobin in their blood so are able to deliver more oxygen to their bodies. (Tr., dkt. # 67, 22:17-20.) All of these differences give males an advantage and

are probably one reason why pass rates for men on physical tests are higher than pass rates for women. (Tr., dkt. # 67, 22:20-23.)

Shorter persons, and women in particular, may tend to use compensatory biomechanical procedures in order to perform a task. (Tr., dkt. # 67, 22:20-23.) As Professor Weltman put it, this “[d]oesn't mean that they can't do the job; it just means they do it differently.” (Tr., dkt. # 67, 23:6-7.)

B. Catherine Erdman.

After a college career playing Division 1 soccer on a full athletic scholarship for four years, earning a masters' degree in international commerce and policy, and a short career in sales, Ms. Erdman entered firefighting, first as a volunteer in Poynette, Wisconsin, and then, in 2007, as a full-time firefighter in Janesville, Wisconsin. (Tr., dkt. # 67, 103.) At Janesville, in a 90-person fire department (Tr., dkt. # 67, 105:14-17), she moved up through a succession of assignments. (Tr., dkt. # 67, 104:22 - 105:4.) As a result of being nominated by her peers and chosen by the fire chief, she

received the Janesville Firefighter of the Year award during the time she was applying for a position in Madison. (Tr., dkt. # 67, 105:5-11.) At the time of the trial in 2018, she had been deployed to about 230 fires, about 60-65 of which were structure fires. (Tr., dkt. # 67, 105:15-19.)

C. The Madison Fire Department.

At the time of the trial, the Madison Fire Department had a joint labor-management recruitment committee. Their only authority was to recommend things relating to recruitment. They were not allowed to make other recommendations that would be supportive of promoting diversity unless it relates to recruiting itself. (Tr., dkt. # 67, 131:9-12.) There was no committee with responsibility to make sure the selection process was fair to women and minorities. (Tr., dkt. # 67, 131:16-19.)

IV. The Present Litigation

Ms. Erdman filed this lawsuit on November 29, 2016, against a single defendant, the City of Madison, making a single claim, that the Physical Abilities Test in the City's 2014 selection process for new firefighters had an unlawful disparate impact on women. (Dkt. # 1.) The City moved for summary judgment. (Dkt. # 12, *et seq.*) Its motion was denied in all respects but one: the district court held that no reasonable trier of fact could find that, if Ms. Erdman has passed the 2014 Physical Abilities Test, she would have been hired as a Madison firefighter. (A-2022, Dkt. # 47.) That decision is challenged on this appeal.

The case was tried to District Judge William M. Conley without a jury on October 15 and 16, 2018 (dkt. ## 64-65), and post-trial briefing was completed by December 21, 2018 (dkt. ## 75-76).

Three and a half years later, on July 20, 2022, the district court issued its Opinion and Order. (A-23, Dkt. # 81.) The court found that the Plaintiff had proven that the 2014 Physical Abilities Test had

“a disparate impact on female applicants, sufficient to satisfy her burden on the first element of a disparate impact claim.” (A-32.)

The court went on to find that the 2014 Physical Abilities Test “was job-related for the position and consistent with business necessity.” (A-37.) The court added that “the Department met its obligation to investigate ‘alternative selection procedures with evidence of less adverse impact . . . to determine the appropriateness of using or validating it in accord with [the Uniform] guidelines.’ 29 C.F.R. § 1607.3(B).” (A-35, ellipses in original.)

Finally the court ruled that the Candidate Physical Ability Test “Was Not Adequate.” (A-37). While the court found that it is “more probably true than not true that the CPAT has less of an adverse impact on female applicants” than the Madison test (A-38), and “that the CPAT is valid generally” (A-38), it concluded that “plaintiff has not demonstrated by a preponderance of the evidence that the CPAT meets the Department’s legitimate needs as an alternative to the 2014 PAT,” (A-40). This finding led to the

judgment of dismissal issued the same day (A-42) and it is the central target of this appeal.

SUMMARY OF ARGUMENT

The disparate impact branch of Title VII jurisprudence permits a plaintiff to prove unlawful discrimination via a burden-shifting proof paradigm. The plaintiff's initial burden is to prove that an employment practice that is neutral on its face has a statistically significant adverse impact on members of her protected group. If she carries that burden, the employer must prove that the challenged practice is a business necessity for the position in question, or the plaintiff prevails at this stage. If the employer meets that burden, the plaintiff may still prove the practice to be unlawful by showing that the employer has refused to adopt an available alternative employment practice that serves the employer's legitimate needs with less disparate impact.

In this case, the district court correctly found that the 2014 Madison Fire Department's Physical Abilities Test had a large

disparate impact on women. It also found, though, that the test was a job-related business necessity, meaning that it fulfilled its purpose of eliminating candidates from the selection process who did not have the physical abilities to perform the duties of a Madison firefighter. The second-stage burden is a low bar for employers and this finding is not challenged on appeal.

The court below also correctly held that the national physical ability test for firefighter applicants promulgated by the International Association of Firefighters and the International Association of Fire Chiefs and modified as a result of negotiations between those entities and the Equal Employment Opportunity Commission, the Candidate Physical Ability Test, which had been considered and rejected by the Madison department, has less of a disparate impact on female applicants.

The court erred, though in finding that the Candidate Physical Ability Test would not have served the Department's legitimate needs. It reached this conclusion by failing to correctly apply two

established legal principles and some poor logic. The legal principles it neglected were:

- A court may not use a test's more exact replication of particular job tasks to conclude that it is a superior selection device if the skills necessary to perform those tasks can be learned in training or on the job;
- The superior test is the test that does a better job of identifying applicants meeting the minimum qualifications for success on the job, not the test that weeds out all but the most qualified applicants;

The poor logic was the district court's reasoning that because applicants who passed the Madison test had good success rates in the Department's training academy, applicants who passed the Candidate Physical Ability Test would be unlikely to do as well there.

If this Court applies correct legal rules and sound reasoning, it will decide the issue of liability for the Plaintiff.

If it does so, it should also reverse the district court's decision on summary judgment that Ms. Erdman is not entitled to make-whole relief because she could not prove at trial that if Madison had used the Candidate Physical Ability Test, which she passed both before and after failing the Madison test, she would probably have survived the Fire Chief interview and been hired. The district judge reached this erroneous conclusion by failing to give Ms. Erdman the benefit of available and reasonable inference from the evidence.

ARGUMENT

I. The Standards of Review

A. The Standard of Review of a Summary Judgment Decision is *De Novo*.

This Court's review of a grant of summary judgment is *de novo*. *G. M. Enterprises, Inc. v. Town of St. Joseph*, 350 F.3d 631, 636 (7th Cir. 2003).

Summary judgment is proper if the evidence of record shows that there is no issue as to any material fact, and the moving party is

entitled to judgment as a matter of law. Fed. R. Civ. P. 56(c); *Celotex Corp. v. Catrett*, 477 U.S. 317, 322 (1986). The burden of establishing the lack of any genuine issue of material fact is always upon the movant. *Outlaw v. Newkirk*, 259 Fed.3d 833, 837 (7th Cir. 2001).

The record must be examined with all factual disputes resolved in favor of the nonmoving party. *Anderson v. Liberty Lobby, Inc.*, 477 U.S. 242, 255 (1986). The Court must observe the “fundamental principle that at the summary judgment stage, reasonable inferences should be drawn in favor of the nonmoving party.” *Tolan v. Cotton*, 572 U.S. 650, 660 (2014).

B. The Standard of Review of a Trial Court’s Decision after a Bench trial Depends on what is Being Reviewed.

1. The “Clearly Erroneous” Standard of Review Applies to Findings Entered after a Bench Trial.

Fed. R. Civ. P. 52 provides: “Findings of fact, whether based on oral or other evidence, must not be set aside unless clearly erroneous, and the reviewing court must give due regard to the trial

court's opportunity to judge the witnesses' credibility." "A finding is 'clearly erroneous' when, although there is evidence to support it, the reviewing court on the entire evidence is left with the definite and firm conviction that a mistake has been committed." *Anderson v. City of Bessemer City, N.C.*, 470 U.S. 564, 573 (1985), quoting *United States v. United States Gypsum Co.*, 333 U.S. 364, 395 (1948).

"Documents or objective evidence may contradict the witness' story; or the story itself may be so internally inconsistent or implausible on its face that a reasonable factfinder would not credit it. Where such factors are present, the court of appeals may well find clear error even in a finding purportedly based on a credibility determination." *Anderson v. City of Bessemer City, N.C.*, 470 U.S. 564, 575, (1985).

2. Conclusions of Law are Reviewed *De Novo*.

While this Court reviews a district court's factfinding under a clearly erroneous standard, it reviews the legal analysis of the court below *de novo*. *Walton v. Jennings Cmty. Hosp., Inc.*, 999 F.2d 277, 282

(7th Cir. 1993). When “the district court employs the wrong legal standard in assessing the facts, its findings are clearly erroneous.” *Moriarty v. Glueckert Funeral Home, Ltd.*, 155 F.3d 859, 864 (7th Cir. 1998). See *Bose Corp. v. Consumers Union of United States, Inc.*, 466 U.S. 485, 501 (1984) (“An appellate cour[t has] power to correct errors of law, including those that ... infect ... a finding of fact that is predicated on a misunderstanding of the governing rule of law”); see also 9C C. Wright & A. Miller, *Federal Practice & Procedure* § 2585, p. 392 (3d ed. 2008) (“[I]t is axiomatic that the conclusions of law of the trial judge are not protected by the ‘clearly erroneous’ test”)

3. Findings on Mixed Questions of Fact and Law are Reviewed with an Appropriate Degree of Deference.

Mixed questions of fact and law are reviewed differently. “[T]he standard of review for a mixed question all depends – on whether answering it entails primarily legal or factual work. *U.S. Bank Nat. Ass'n ex rel. CWCapital Asset Mgmt. LLC v. Vill. at Lakeridge, LLC*, 138 S. Ct. 960, 967 (2018). Some mixed questions of fact and law

“require courts to expound on the law, particularly by amplifying or elaborating on a broad legal standard. When that is so – when applying the law involves developing auxiliary legal principles of use in other cases – appellate courts should typically review a decision *de novo*.” *Id.* But where “mixed questions immerse courts in case-specific factual issues – compelling them to marshal and weigh evidence, [and] make credibility judgments . . . appellate courts should usually review a decision with deference.” *Id.*

II. A reasonable trier of fact could conclude that but for the unlawful disparate impact of the PAT, the Plaintiff would have been hired, and if such a finding is made, she will be entitled to back pay, front pay, and reinstatement into the position at issue.

A. A reasonable trier of fact could find that Ms. Erdman is entitled to back pay and reinstatement into the job at issue.

In 2014, the Fire Department ultimately hired every one of the four women who passed the Physical Abilities Test but hired only thirteen of the 395 men who successfully completed the Physical Abilities test. (A-3.) Thus, in the 2014 hiring process, the City’s

emphasis on hiring women (Affidavit of Steven Davis, 2/23/2018, dkt. # 15, ¶¶7-8) was such that a woman who made it through the Physical Abilities Test had a one hundred percent chance of being hired. The Plaintiff's burden is only to prove that it is more probable than not that she would have been hired if she had passed the test.

1. This Court Should Apply the Ordinary Civil "Preponderance of the Evidence" Standard of Proof.

The district court held that no reasonable trier of fact could hold that Ms. Erdman would have been hired if she had passed the Physical Abilities Test. (A-21-22.) The court adopted the standard for awarding make-whole relief from *Evans v. City of Evanston*, 881 F.2d 382 (7th Cir. 1989), where this Court said that make-whole relief was appropriate "where it is reasonably clear that, had it not been for the discriminatory behavior, the plaintiff would have got (or retained) the job or other employment benefit in issue, and where making the plaintiff whole would not unduly injure innocent third parties." (A-20, quoting *Evans*, 881 F.2d at 386.)

Neither this Court, in *Evans* or elsewhere, nor the district court, has discussed the relationship between that case's "reasonably clear" standard of proof and the ordinary civil "preponderance of the evidence" standard.⁴ The standard of proof applicable to damages claims in civil actions is "preponderance of the evidence." Federal Civil Jury Instructions of the Seventh Circuit, 3.09, 3.10. The preponderance standard applies to the equitable remedy of back pay. *Id.*, 3.10. Neither this Court, in *Evans*, nor the district court, suggested that the "reasonably clear" standard should be higher than the preponderance standard, and neither has offered any explanation as to why this should be so. In its opinion in this case, this Court ought to explicitly state that *Evans* did not intend to suggest a standard of proof for make-whole relief different from the

⁴ There is some authority for the proposition that, at least as used in Montana state law, "reasonably clear" is a more stringent standard of proof. See, e.g., *Palmerton v. Wal-Mart Stores, Inc.*, No. CV 17-30-H-CCL, 2017 WL 4031432, at *1 (D. Mont. Sept. 13, 2017) ("The test for reasonably clear liability is an objective test. It is a greater degree of certainty than preponderance-of-the-evidence and more akin to the 'clear and convincing evidence' standard." (internal quotation marks and citations omitted)).

preponderance standard, and that a plaintiff who proves by a preponderance of the evidence that, but for an employer's rejection of her candidacy because she failed an unlawfully discriminatory test, she would have gotten the job at issue, should receive full make whole relief.

2. If Ms. Erdman Proves that the PAT had an Unlawful Disparate Impact, She will be Entitled to a Trial on her Claim for Make-Whole Relief.

It seems clear that women candidates who passed the Physical Abilities Test were not selected for firefighter jobs using the same standards applied to male candidates. The City hired only thirteen of the 395 men who successfully completed the Physical Abilities test. (A-3.) That is a success rate of 3.3%. It hired all four women who successfully completed the Physical Abilities test. (A-3.)

If the first woman who passed the Physical Abilities Test would have been assessed using the same standard applied to the male candidates, she would have had a 3.3% chance of being hired. The chance that, applying the same standard, the City would have

hired both the first and second women who passed the test would have been 0.1%. (The calculation is exactly like calculating the odds of flipping a coin and getting heads twice in a row: $50\% \times 50\% = 25\%$.)

If all four women who passed the Physical Abilities Test had been selected or rejected using the same standard applied to the men, the chance of all four of them getting hired would be a miniscule 0.0001% ($3\% \times 3\% \times 3\% \times 3\%$). So clearly, a reasonable trier of fact could find that the women candidates were not assessed using the same standard applied to the men, but were assessed under a much easier standard.

What that means is that a reasonable trier of fact could conclude that, if Ms. Erdman had passed the Physical Abilities Test, she would have had a much better chance of being hired than the average man who passed the test, because she was female. The success rate of other female candidates may not be enough by itself to prove that Ms. Erdman would have been hired, but it is not the only evidence Ms. Erdman has of her probable success. At the time

she was tested in 2014, Ms. Erdman had been a serving firefighter for the City of Janesville for seven years. (A-6.) Considered in conjunction with the 100% success rate of the female candidates who passed the 2014 Physical Abilities Test, and the City's emphatic public claims that it wanted more female firefighters (Affidavit of Steven Davis, 2/23/2018, dkt. # 15, ¶¶7-8), the Plaintiff should have been at least entitled to a trial on the issue of whether she would have enjoyed the same fate as the passing female applicants or been the rejected exception. After all, she does not have to prove that the City "would have hired any woman who passed the 2014 PAT" (A-21). She just has to prove that it is more probable than not that the City would have hired *her*.

The district judge was influenced by the fact that the "plaintiff was not hired in 2016 even though she had progressed to the chief's interview." (A-21.)⁵ Maybe a reasonable trier of fact could infer

⁵ Progressing to the chief's interview in 2016 did not require passing the Physical Abilities Test, because the order of events in the selection process was changed so that the chief's interview came before the physical test. (Tr., dkt. # 67. 120:3-6.)

from this that there was something disqualifying about her candidacy, in addition to her having failed the Physical Abilities Test, but that is a defense-favorable inference to which the City is not entitled at summary judgment. On the contrary, Ms. Erdman, who had been required to file her charge of discrimination with the EEOC less than a year after her 2014 rejection to preserve her Title VII rights, should be entitled to an inference that the City's disposition of her candidacy *after* she had initiated litigation against it has no predictive value as to what its decision would have been had she passed the physical test *before* she filed her case. With no evidence about *why* Ms. Erdman did not survive the chief's interview in 2016, the naked fact that she did not provides no support for that defense contention that it would be impossible for her to prove at a trial that she likely would have been hired in 2014.

B. Ms. Erdman will be entitled to injunctive relief.

Even if she were to fall short of proving that she would have gotten a job in 2014 but for the unlawful physical ability test, Ms.

Erdman is entitled to injunctive relief. She still wants to be a Madison firefighter. She has applied twice since she was washed out of the 2014 hiring process, and, especially if the test is modified to be fairer to women, she expects to apply again. (Erdman Declaration, ¶ 20.) These facts give her a concrete interest in the future lawfulness of the test sufficient to permit her to seek injunctive relief under the strictures laid down by the U.S. Supreme Court in *City of Los Angeles v. Lyons*, 461 U.S. 95 (1983).

III. Even if the City showed that the test was a job-related business necessity, the Plaintiff carried her third-stage burden of showing that the City had considered and rejected an alternative test with substantially equal validity and a much smaller adverse impact on women.

The district court found that the Physical Abilities Test had a disparate impact on female applicants significant enough to shift the burden of proof to the defense. (A-32-33.) The City did not dispute that, when the passages rates of males and females on the test were compared, there was a large, and highly statistically significant,

difference. The City's defense on this point was to contend that the comparison ought not be between male and female passage rates on the test as a whole, but between passage rates on just the one event that washed Ms. Erdman out of the selection process. The court rejected this unusual argument. (*Id.*)

The trial court then went on to hold that the City had carried its second-stage burden of proving that the test was "job-related for the position and consistent with business necessity." (A-33-37.)

Although Ms. Erdman argued against such a conclusion in the trial court, she does not ask this Court to reverse this holding. Rather, she prays the Court to reverse the district court's ruling that the Candidate Physical Ability Test was not shown to be a substantially equally valid selection instrument, which emerged from the district court's failure to apply the correct legal standards.

IV. The nature of the Plaintiff's stage-three burden of proving that an available alternative selection practice (1) results in a less disparate impact, and (2) serves the employer's legitimate needs is to identify a substantially equally valid alternative selection procedure with a smaller gap between passing rates.

The statute, 42 U.S.C.A. § 2000e-2 (k) entitled "Burden of proof in disparate impact cases," was crafted by Congress as part of the Civil Rights Act of 1991 in order to repair the perceived damage to disparate-impact doctrine done by *Wards Cove Packing Co. v. Atonio*, 490 U.S. 642 (1989).⁶ It is not a model of clarity. A plaintiff can win her case either if the defendant employer "fails to demonstrate that the challenged practice is job related for the position in question and consistent with business necessity," or, assuming the employer makes this showing, if the plaintiff makes a demonstration "in accordance with the law as it existed on June 4, 1989," the day before *Wards Cove* was handed down, "with respect

⁶ "The Civil Rights Act of 1991," website of the United States Equal Employment Opportunities Commission, <https://www.eeoc.gov/eeoc/history/35th/1990s/civilrights.html>, last accessed Dec. 7, 2018.

to an alternative employment practice and the respondent refuses to adopt such alternative employment practice.” 42 U.S.C.A. § 2000e-2 (k).

The Supreme Court’s translation of this language is:

Twenty years after *Griggs [v. Duke Power Co., 401 U.S. 424 (1971)]*, the Civil Rights Act of 1991, 105 Stat. 1071, was enacted. The Act included a provision codifying the prohibition on disparate-impact discrimination. That provision is now in force . . . Under the disparate-impact statute, a plaintiff establishes a prima facie violation by showing that an employer uses “a particular employment practice that causes a disparate impact on the basis of race, color, religion, sex, or national origin.” 42 U.S.C. § 2000e-2(k)(1)(A)(i). An employer may defend against liability by demonstrating that the practice is “job related for the position in question and consistent with business necessity.” *Ibid.* Even if the employer meets that burden, however, a plaintiff may still succeed by showing that the employer refuses to adopt an available alternative employment practice that has less disparate impact and serves the employer's legitimate needs. §§ 2000e-2(k)(1)(A)(ii) and (C).

Ricci v. DeStefano, 557 U.S. 557, 578 (2009).

A. Relevant Parts of the Uniform Guidelines.

The Uniform Guidelines⁷ speak to the issue of when an employer is required to use an alternative selection procedure:

§1607.3(B). Consideration of suitable alternative selection procedures. Where two or more selection procedures are available which serve the user's legitimate interest in efficient and trustworthy workmanship, and which are substantially equally valid for a given purpose, the user should use the procedure which has been demonstrated to have the lesser adverse impact. . . .

(Exhibit 51, p. 2.)

B. A Plaintiff succeeds in her stage three burden if she identifies an alternative test that serves the employer's legitimate interest in efficient and trustworthy workmanship.

The issue is whether the alternative test is “substantially equally valid.” *Allen v. City of Chicago*, 351 F.3d 306, 312 (7th Cir. 2003), citing 42 U.S.C. § 2000e-2(k)(1)(A); 29 C.F.R. § 1607.3(B).

⁷ This Court has looked to the EEOC’s Uniform Guidelines on Employee Selection Procedures (in the records as Exhibit 51) as authority in disparate impact cases. *Gillespie v. State of Wis.*, 771 F.2d 1035, 1041 (7th Cir. 1985).

How equal is “substantially equal”?

As a result of the 1991 Amendments to the Civil Rights Act, Pub.L. 102-166, § 105(a), 105 Stat. 1071, 1074 (1991) (codified at 42 U.S.C. § 2000e-2(k)), plaintiffs seeking to demonstrate a less discriminatory alternative must do so under the law that existed prior to the Supreme Court's decision in *Wards Cove Packing Co. v. Atonio*, 490 U.S. 642 (1989). See 42 U.S.C. § 2000e-2(k)(1)(C); see also *Price v. City of Chicago*, 251 F.3d 656, 660 (7th Cir.2001). Prior to *Wards Cove*, the Supreme Court expressed the controlling principle in *Albemarle [Paper Co. v. Moody*, 422 U.S. 405 (1975)]. “If an employer does then meet the burden of proving that its tests are ‘job related,’ it remains open to the complaining party to show that other tests or selection devices, without a similarly undesirable racial effect, *would also serve the employer's legitimate interest in ‘efficient and trustworthy workmanship.’*” *Albemarle*, 422 U.S. at 425 (quoting *McDonnell Douglas Corp. v. Green*, 411 U.S. 792, 801 (1973)).

Allen v. City of Chicago, 351 F.3d 306, 312 (7th Cir. 2003)(emphasis supplied).

V. Ms. Erdman Proved that the CPAT has a Much Smaller Disparate Impact on Women than the PAT and is Substantially Equally Valid.

A. The City Considered and Rejected the CPAT.

In this case, the record shows that the City's decision-makers considered and consciously rejected the CPAT. Then firefighter Francis Tatar, whose job assignment was to address the recruitment and staffing needs of the Department (Tr., dkt. # 67, 179:23-24), and who sat on the recruitment committee, investigated the CPAT to a certain extent. On January 7, 2013, she reported to Fire Chief Steven Davis and others that she had contacted Chief Brenda Berkman and learned that, before Berkman had retired from the Fire Department of New York in about 2007, the organization Women in the Fire Service had originally opposed the CPAT for various reasons (Tr., dkt. # 67, 185:19 - 187:1, Exhibit 25), including "Failure to provide orientation and pre-exam prep for candidates." About this deficiency, Berkman wrote, "This omission was/is especially critical

for the success of women candidates as women's scores improve with practice." (Tr., dkt. # 67, 187:21 - 187:1, Exhibit 25.)

On January 9, 2013, Tatar emailed a number of recruitment committee members and others about the CPAT. (Tr., dkt. # 67, 187:21 - 188:12; Exhibit 26.) She reported that she had learned that Women in the Fire Service (which by then had changed its name to I-Women (Tr., dkt. # 67, 188:20-22) by then *did* support the use of the CPAT. (Exhibit 26.)

On September 5, 2013, Tatar emailed a large number of Department personnel including Chief Davis on behalf of what she referred to as the Recruitment and Hiring Committee. (Exhibit 29.) She reported on a number of concerns about the validity of the Madison test. (*Id.*) Under the heading, "Early access to training," she informed these persons,

At least 2 training sessions, no less than 8 weeks prior to the CPAT, are offered to every candidate. At this training session, candidates receive help from trainers to become familiar and prepare for the test. Official testing equipment is used. No less than 2 timed practice

runs are offered to every candidate, within 30 days of the test.

(Id.)

Tatar had gone so far as to obtain and read the 2013 Austin revalidation study of the CPAT. (Tr., dkt. # 67, 198:6-18; Exhibit 31.) She is also sure she talked about the CPAT with Chief Davis and she forwarded the information she had learned in her investigation of the CPAT to him as well. (Tr., dkt. # 67, 200:156-21.)

B. The CPAT has a much smaller disparity in passage rates by sex.

The district court correctly found “that the CPAT has less of an adverse impact on female applicants.” (A-38.)

Among departments surveyed for a 2008 report, the average CPAT pass rate for women was 68.0% compared to 49.0% in departments using other physical ability tests. The CPAT ratio of pass rates for women compared to men was 77.4% which is higher than the 61.7% ratio reported for all other physical ability tests. (A-38; Exhibit 40, pp. 6-7.)

The 2014 Madison Fire Department PAT pass rate for women who appeared to take the test was 4 out of 28 or 14%. (Dkt. #53, ¶ 53.) The 2014 Madison Fire Department pass rate for men who appeared to take the test was 395 out of 471, or 84%. (*Id.*) The female pass rate as a percentage of the male pass rate was thus 16%, a much greater disparity than the 77.4% ratio observed with the CPAT.

It is not difficult to determine some of the reasons why the CPAT screens out fewer women. Professor Weltman discussed some of the differences between the two tests in his testimony. (Tr., dkt. # 67, 14:12 – 15:8.) The CPAT does not include a requirement that an applicant remove a ladder from an elevated rack and replace it there. The ladder event is:

4. Ladder Raise & Extension

This task is intended to test the participant's ability to place a ground ladder from a fire truck and extend it to a roof or window. The subject walks to the top rung of a grounded ladder, lifts the unhinged end from the ground, and raises it up hand over hand until it is secured against the wall. He/she must then immediately advance to the next ladder and, standing

with both feet in the marked box, extend the fly section until it stops, then lower" it back to the starting point,

(Exhibit 6, CPAT Revalidation Study, dkt. # 15-2, 3.)

When the PAT was being considered in 2013, Fire Lieutenant Francis Tatar sent an email noting that the test required candidates to remove a 20-foot ladder of an elevated rack at a time when Madison fire trucks no longer had 20-foot ladders mounted on their sides. (Exhibit 29, Tr., dkt. # 67, 140:14-16.) The district judge found that "the 20-foot ladder was used when at least some of the Department's rigs still used a 24-foot extension ladder." (A-28) He believed that, "The choice of ladder, as well as the required manipulation of the ladder, reflected the job requirements at that time," (A-28), but, in her testimony, Ms. Erdman explained that a 24-foot extension ladder is only 14 feet long when collapsed and mounted on a fire truck (Tr., dkt. # 67, 113:18-20), and she explained how the longer 20-foot ladder is more awkward to maneuver for shorter candidates than the collapsed extension ladder (Tr., dkt. # 67, 114:5-25). The ladder test was changed after 2014 so that now the

applicants pull a ladder out of a sleeve that simulates the ladder storage compartments in the back of fire trucks. (Tr., dkt. # 67, 143:23 - 144:5.) The sleeve is not as high off the ground as the wall-mounted ladder rack used in the 2014 test. (Tr., dkt. # 67, 144:6-8.) One of the reasons for this change was to make the event fairer to persons of shorter stature. (Tr., dkt. # 67, 144:9-11.) Ergometrics had stated in a 2013 email that changing the ladder event “to simulate pulling ladder from back of a truck or modifying the event (i.e., lowering the height of the ladder or eliminating the final step of putting the ladder back on the wall) . . . would likely result in less adverse impact.” (Exhibit 30, p. 3, Tr., dkt. # 67, 162:21-24; 167:3-13.)

In addition, the CPAT is easier for shorter candidates in the pike pole event, because they are not required to stand so far off to the side when pulling down against resistance. (Tr., dkt. 67, 110: 10-24.) Confining individuals to a given space to perform the pike pole event may not allow for compensation in mechanics that would allow Ms. Erdman to successfully complete the event and is not consistent with what would happen during an active fire where she

could maneuver to achieve optimal performance. (See, Tr., dkt. # 67, 109.) Ms. Erdman has fought fires in the field and often had to pull down ceilings. Her practice has been to stand close enough to do the job without pulling burning material down onto herself or others. This goal is achieved as much through technique as through the firefighter's positioning of himself or herself. The effective techniques include perforating the burning ceiling so it can be pulled down in smaller pieces that can be more easily controlled than larger sections, and these are techniques that can be taught. (Tr., dkt. # 67, 109.)

Finally, the CPAT requires departments to provide more meaningful preparation to prospective test-takers than the Madison test. As a result of the test creators' negotiations with the Equal Employment Opportunity Commission in 2006,⁸ Departments were required to offer at least two practice sessions within the eight weeks

⁸ The Conciliation Agreement between the International Association of Firefighters and the EEOC is appended to the CPAT Manual as Appendix F, and is in the record as Exhibit 37, pp. 89-94.

prior to the actual CPAT test with trained personnel available to take the candidates through each part of the test and explain its nuances. (Tr., dkt. # 67, 15:19 – 16:5.) In addition to these orientation sessions, the departments are required to provide timed run-throughs as well. (Tr., dkt. # 67, 16:6-12.) The reason for these modifications was that, as Professor Weltman put it, “there is a very strong learning effect” from repetitions of the test. (Tr., dkt. # 67, 16:4-5.)⁹

This strong learning effect from test repetitions is believed to be more pronounced for women than for men because sheer brute strength will often allow men to perform the test adequately while “women may need to employ different biomechanical strategies.” (Tr., dkt. # 67, 27:17-18.) Professor Weltman acknowledged, however, that he was aware of no studies of the degree to which this difference can be observed. (Tr., dkt. # 67, 31:23 – 32:1.) The Austin CPAT Revalidation Study, however, reported that the test results

⁹ In one of the original validation studies of the CPAT, candidate performance improved by a minute and a half from the first run-through to the second. Tr., dkt. # 67, 18:5-10.)

gave “credence to the observation that familiarity with the test is a definite advantage for women.” (Exhibit 6, p. 10; Tr., dkt. # 67, 37:21-23.) That kind of familiarity could be obtained by the orientations and timed run-throughs that the CPAT has required departments to offer since 2006, as it can be acquired on the job. (Tr., dkt. # 67, 37:25 – 38:6.)

C. The CPAT is at least substantially equally valid.

The Plaintiff need not show that the CPAT is a *better* measure of the qualities conducive to effective firefighting than the City’s test, or even that it is just as good as the City’s test in identifying the very best candidates and screening out all others. The Plaintiff need only show that the CPAT “would also serve the employer’s legitimate interest in efficient and trustworthy workmanship.” *Allen v. City of Chicago*, 351 F.3d 306, 312 (7th Cir. 2003), quoting *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 425 (1975).¹⁰

¹⁰ Professor Weltman testified that “in terms of creating a broad-based applicant pool that’s capable of being trained in the academy to do the job, in my opinion the CPAT is a much more valid test. (Tr., dkt. # 67, 19:1-3.)

This formulation fits in well with the principle that “a discriminatory cutoff score on an entry level employment examination must be shown to measure the *minimum* qualifications necessary for successful performance of the job.” *Ernst v. City of Chicago*, 837 F.3d 788, 804 (7th Cir. 2016), quoting, *Lanning v. Se. Pa. Transp. Auth.*, 308 F.3d 286, 287 (3d Cir. 2002) (emphasis supplied).

As Professor Weltman explained in his report:

The Candidate Physical Ability Test (CPAT) was developed as a content-based test designed for screening potential firefighter candidates. It was a joint collaboration between the International Association of Fire Chiefs (IAFC) and the International Association of Fire Fighters (IAFF) (IAFF/IAFC Joint Labor Management Task Force) and is widely accepted as a bona fide occupational qualification screening instrument to determine which candidate recruits are able and which are not able to perform the physically demanding job of firefighting. It is important to note that this test was developed in conjunction with 10 leading fire departments across North America: Austin TX, Los Angeles Co., CA, Calgary, ALB, Metro Dade Co., FL, Charlotte NC, New York City. NY, Fairfax Co., VA, Phoenix, AZ, Indianapolis, IN, and Seattle, WA. . . . The Task Force directed the Technical Committee to develop a performance/screening test for the 10

departments that would measure critical physical abilities associated with firefighting.

(Exhibit 45,¹¹ 13; *see also*, Exhibit 6, 13, Tr., dkt. # 67, 12:17 - 13:25.)

“Great effort was made to adhere to” the Uniform Guidelines “in the development and validation of CPAT.” (Exhibit 6, 13.) The CPAT was revalidated in 2013. (Dkt. #53, ¶ 42.)¹²

Carl Swander’s company, NTN, has administered the CPAT over 10,000 times. (Tr., dkt. # 68, 2-108:13-18.) His company has done the validity/transportability studies required by the IAFF and IAFC (see Exhibit 37, p. 10) for each department where it has administered the CPAT and they have found it to be valid for each of these “many, many” (Tr., dkt. # 68, 2-114:19) departments. (Tr., dkt. # 68, 2-108:19 - 2-109:8.) There is no reason to believe it would not be found valid for the Madison Fire Department. (Tr., dkt. # 68,

¹¹ The parties’ expert witnesses’ reports were formally received in evidence without objection (Tr., dkt. # 67, 40:1-20), and the district court relied on Professor Weltman’s report in making its post-trial findings. (A-38.) Judge Conley explained during the trial that in bench trials he often reads the expert reports and proceeds directly to cross-examination. (Tr., dkt. # 67, 17:11-13.)

¹² The report of the 2013 revalidation of the CPAT is in the record as Exhibit 6.)

2-114:21 - 2-115:1.) After all, “the demands of a firefighter are highly similar, regardless of location, setting or individual differences.”

(Exhibit 9, p. 99 of 181.)

Given its pedigree and the evidence of its validation and revalidation, it is unsurprising that the district court said, “The court further finds that the CPAT is valid generally, which the defendant does not really challenge.” (A-38.) The defense that won the case was that, as good as the CPAT is, the Madison test is better. In order to reach that conclusion, the district court had to ignore important legal constraints, and applying those principles properly should lead this Court to reverse.

It bears mentioning that, while the trial court found that the City’s test was sufficiently valid to carry its second-stage burden of proof and shift the burden back to the Plaintiff, when the validity of the Madison test and the CPAT are compared, as they must be in the third stage of the analysis, the Madison test does not set a very high bar. First, the cut scores, meaning the pass/fail scores, and the disqualification scores, in the 2014 test came from the 1999 test (Tr.,

dkt. # 67, 12:7-12). The 1999 validation report says that the disqualification times were set at the time of the poorest performing incumbent firefighter who took the test for purposes of the validation process. (Exhibit 4, 16; Tr., dkt. # 67, 38:7-17).

The disqualification scores were more important than the pass-fail scores, as least for the women: Eighty percent of the women who appeared to take the test and did not quit were disqualified. Only one (4%) of the women who (Dkt. #53, ¶ 53) appeared to take the test and did not quit failed by virtue of not achieving passing scores on at least five events. (*Id.*) Yet establishing the disqualification times at the level of the poorest performing incumbent is a method not supported in any refereed journal. (Tr., dkt. # 67, 38:7-17.)

In contrast, the manner in which the disqualification time for the CPAT was determined is well-documented, and involved trained subject-matter experts from ten different fire departments

observing thousands of firefighters performing the test at 13 different simulated speeds. (Tr., dkt. # 67, 11:13-25.)¹³

Secondly, the City's evidence failed to make any attempt to "correlate" performance on the PAT with "actual performance on the job." *Horan v. City of Chicago*, No. 98 C 2850, 2003 WL 22284090, at *71 (N.D. Ill. Sept. 30, 2003). As in *Horan*, the City hoped to slide by on an argument that the excellence of its test is obvious. But the issue is whether "performance on a test . . . fairly predicts relative performance as a firefighter," *Horan, id.*, at *72, and there was no evidence that it does. "[A] discriminatory cutoff score on an entry level employment examination must be shown to measure the minimum qualifications necessary for successful performance of the job." *Ernst v. City of Chicago*, 837 F.3d 788, 804 (7th Cir. 2016), quoting, *Lanning v. Se. Pa. Transp. Auth.*, 308 F.3d 286, 287 (3d Cir.2002) (emphasis supplied). As the district court noted, while this evidence is important, it is not strictly required. (A-11.)

¹³ "33" at the cited point in the transcript appears to be a typographical error. See, Tr., dkt. # 67, 76:11-14.)

Third, the absence of opportunities for repetition of the test also impairs the test's validity. As Professor Weltman explained:

Q Are you saying that the reliability of a test can increase with the number of repetitions?

A Absolutely. As you practice and you learn how to do the procedure, the test becomes more reliable, particularly if it's a test that has a learning effect. Now, there may be some tests where a learning effect is minimal so doing the test from day one to day two is going to produce the same score. There are other tests where there's a considerable learning effect, and so there are a number of practice trials that need to be performed before the test can be reproduced reliably, and that's one of the reasons why the CPAT incorporates these two mandatory training sessions with coaching, so that candidates can learn how to do the procedures and develop what we call in our field a true score, a score that's actually reflective of their true ability.

(Tr., dkt. # 67, 24:12 - 25:1.)

With these glaring deficiencies in the PAT's validity evidence, it is not so daunting a task to show that the CPAT is substantially equally valid.

D. The District Court's Reasons for Finding the CPAT not Substantially Equally as Valid as the Madison Test do not Withstand Scrutiny.

1. A Test's Exact Replication of Job Tasks Cannot Support a Finding of Greater Validity if those Skills Can be Learned on the Job.

The City's argument at the third stage of the disparate impact paradigm was that the CPAT would not serve the City's legitimate business needs, because the Madison PAT, and especially the two events in which Ms. Erdman failed to achieve a passing score, the ladder event and the pike pole event "were specifically designed to replicate the tasks Madison firefighters would be expected to execute in light of the equipment available to the Department at that time and of concerns about safety." (A-39.) The district court referred to its "crediting" of this testimony (A-39), and this appears to be one of the reasons the trial judge held that the CPAT was not substantially equally as valid as the Madison test.

The City argued that its test is self-evidently more valid than the CPAT because it more precisely replicates the exact job duties of

Madison firefighters, right down to the maximum number of floors permitted by law¹⁴ in Madison buildings.¹⁵

If it did, in fact, find that the City's test was more valid than the CPAT because it more precisely replicated the physical tasks performed by Madison firefighters (without any evidence that its test takers' results in fact correlated better with actual job performance than the scores of CPAT test-takers), it failed to apply a crucial legal principle: tests that favor majority test-takers and exactly replicate job tasks are not favored where the skills necessary to perform those tasks can be learned on the job.

An exact replication of the job duties by the test components is frowned upon by the Uniform Guidelines because it is a poor guarantor of validity where the skills tested can be learned. The

¹⁴ https://madison.com/ct/news/local/writers/mike_ivey/by-design-madison-has-no-tall-buildings/article_64e73570-b5e1-11e3-9e53-0019bb2963f4.html, last accessed Dec. 6, 2018.

¹⁵ When, however, the Plaintiff argued that the ladder component of the test was not a realistic replication, Mr. Swander had no difficulty in justifying the use of test components that do not exactly replicate real tasks in the field as long as they "measure ability to handle material in a safe manner and using the same muscle groups as are utilized on the job." (Exhibit 39, Swander Report, 18.)

performance of applicants with no prior experience in performing the tested tasks can not only be improved by learning on the job, it can be improved by coaching and instruction in specialized techniques, by practice, and by physical fitness training. (Exhibit 6, 2013 CPAT Revalidation Report, pp. 5-6 of 42.)

The evidence is unanimous that the sorts of instruction, coaching, and practice that the CPAT must provide because of the 2006 conciliation agreement with the EEOC probably help all applicants do better on that test, and would probably help them do better on the Madison PAT. The evidence is also unanimous that these performance enhancers would probably help women and smaller men more than larger and stronger men who have the option of solving many mechanical problems – though not all – through the application of sheer brute force. (Tr., dkt. # 67, 237:18 – 238:24.) Even the City's staunchest witnesses defended the City's failure to provide these things based on budgetary and logistical factors rather than by claiming they do not work to narrow the

passage-rate gap between men and women. (Tr., dkt. # 67, 231:8-16. 238:24 – 239:4.)

But, as Chief Amesqua acknowledged, the instruction, coaching, and practice that help women perform these firefighting tasks better can also be gained on the job (Tr., dkt. # 67, 236::16-21), and they are certainly covered in the training of new firefighters before they assume their duties. The significance of this training cannot be overlooked. As even Dr. Swander said:

While it is certainly not required to account for training, it is another component that can be used to justify lower minimum standards. This helps to ensure optimum opportunity to hire a diverse workgroup, as firefighters do go through a rigorous training period that includes physical fitness. Although true improvement of all candidates is unknown because prior training is not evaluated, there may be expected improvement.

(Swander Report, Exhibit 47, 9).

The Uniform Guidelines prohibit using the inability to perform specific tasks as a disqualifier where these tasks can be learned.

F. Caution against selection on basis of knowledges, skills, or ability learned in brief orientation period. In general, users should avoid making employment decisions on the basis of measures of knowledges, skills, or abilities which are normally learned in a brief orientation period, and which have an adverse impact.

(Uniform Guidelines, § 1607.5(F), Exhibit 51, 4.)

In the technical standards for content validity studies, the Uniform Guidelines provide:

Content validity is also not an appropriate strategy when the selection procedure involves knowledges, skills, or abilities which an employee will be expected to learn on the job.

(Uniform Guidelines, § 1607.14(C)(1), Exhibit 51, 9.)

City expert Jacobs acknowledged that he tries to avoid testing skills that can be learned on the job, and that there is a tension between trying to select, for testing, activities that will be *used* on the job without testing performance in activities that an employee will *learn* on the job. (Tr., dkt. # 68, 2-66:6-10.) Erdman, a very experienced firefighter, testified that there are techniques that can be learned on the job for pulling down a burning ceiling that do not

require the firefighter to stand a full foot and a half from the point of attack. (Tr., dkt. # 67, 1-108 - 1-110.)

There is only one conclusion. The Guidelines prohibit making the components of a physical abilities test replicate exactly the critical physical tasks associated with a job when such replication comes at the cost of creating a test with a stupendous adverse impact on women, and where subjects who pass the test and are hired will be able to learn these abilities “in a brief orientation period,” or “on the job.” This is especially so where there is no evidence that test-takers who do better on the test will also do better in their actual performance on the job.

In failing to apply these principles, the district court committed the sort of error of law that this Court may reverse.

The other bases for the district court’s rejection of the CPAT were also infected with legal error.

The portion of the district court’s post-trial Opinion and Order rejecting the CPAT as substantially equally valid is very short. It sets forth its evidentiary grounds for its conclusion as follows:

[D]efendant proffered credible evidence of numerous burdens associated with adopting the CPAT as an alternative test, including: (1) the need to perform a transferability study; (2) the PAT having been a good predictor of outcome historically, as defined by a high passage rate out of the academy; (3) the Department's comparatively high percentage of female firefighters, leading to a possible inference that the CPAT may have a favorable disparate impact on women but results in the washing out of ultimately unsuccessful applicants after the additional expenditure of time and money at the academy phase; and (4) certain elements of the PAT were designed specifically for Madison, in light of characteristics of the city, the Department's equipment or other considerations, including safety.

(A-40.)

Of these grounds, the last one is a repetition of the one discussed just above. The others will be discussed here:

2. The need to perform a transferability study.

For a Department to administer the CPAT, it must receive a license. (Exhibit 37, 12.) To receive a license, it must do a transportability study. (Exhibit 37, 4.) Such a study requires a survey of current firefighters to ascertain their job duties (*id.*, at 10, et seq.)

and an equipment survey (*id.*, at 12, et seq.). If the duties and the equipment at a given department are sufficiently similar to those of the original ten departments, the duties and equipment of which were considered in designing the CPAT, transportability is demonstrated and licensure is granted. (*Id.*, at 12, “Evaluation of Job Analysis and Equipment Survey.”)

In 2013 and 2014, Ergometrics performed a content validity study of the PAT test against Madison firefighters' job demands. (Dkt. #53, ¶ 21.) That study found that the job duties of a Madison firefighter were highly correlated with the job duties of firefighters in other departments. (*Id.*)

Professor Weltman testified that there is no reason to believe that, if the CPAT were adopted in Madison, the disparity between female passing rates and male passing rates would be any different than the disparity reported in the “National Report Card” study (Exhibit 40, at 6-7: CPAT female passing rate of 68.0% was 77.4% of the male passing rate), because the CPAT had higher female passing

rates than other tests “[i]n all the cases where it’s been transported.”

(Tr., dkt. # 67, 95:24 – 96:11.)

So, the CPAT would be unlikely to be found invalid in Madison, and it would be likely to be successful in significantly raising the passing rate of females.

If it was the cost of a transportability study that the district court found justified retaining the Madison test with its much greater disparity in passing rates, the Court should consider whether this makes any sense considering the fact that the City was regularly shelling out money to consultants to validate its own more discriminatory test. It paid for the test in 1997, had it revalidated in 1999, and 2003, and had it revalidated again in 2013-2014. (Dkt. #53, ¶¶ 5-9.) Especially in 2013-2014, this money could have been better spent by doing the modest work necessary to locally validate the CPAT.

It is somewhat ironic that the City justifies its refusal to adopt the CPAT with its much smaller disparity between male and female passing rates because of the expense of doing a study to confirm its

local validity, which is the sort of study fire departments do largely to position themselves to defend against disparate impact lawsuits. (See, *e.g.*, testimony of former Fire Chief Debra Amesqua, Tr., dkt. # 67, 205:14-19.) So it keeps a test with a vastly greater disparity in passing rates because it does not want to incur the expense of preparing to defend a test with a much smaller disparity in passing rates.

3. The PAT having been a good predictor of outcome historically, as defined by a high passage rate out of the academy

The City's central justification for keeping its highly discriminatory PAT is that it believes it has an elite fire service and wants to keep it that way by hiring the most physically capable people it can find. (City Brief, dkt. # 70, 17.)¹⁶ This approach

¹⁶ The City's belief is well expressed in a question posed by its counsel to Professor Weltman: "Do you agree that diversity should never come by lowering validated entry standards?" (Tr., dkt. # 67, 75:12-13.) If the City's "entry standards" are having a disparate impact on females because they are set substantially above "the minimum acceptable level," *Ernst*, 837 F.3d 788 at 803, they are unlawful. When pressed to agree that the standards for passing the CPAT were lower than the standards for passing the Madison test -- a fact that would have supported the City's theory that

conflicts with *Ernst v. City of Chicago*, 837 F.3d 788 (7th Cir. 2016), which held, “physical test cutoff scores should be set at the minimum acceptable level, not at the maximum level, because the goal is to identify people with sufficient physical skills.” *Id.*, 837 F.3d at 803 (internal quotation marks omitted). As the *Ernst* court said later,

The minimum requirement is adequacy, not superiority. See *Lanning v. Se. Pa. Transp. Auth.*, 308 F.3d 286, 287 (3d Cir.2002) (affirming that, in a disparate-impact claim, “a discriminatory cutoff score on an entry level employment examination must be shown to measure the minimum qualifications necessary for successful performance of the job”)

Ernst, 837 F.3d 788 at 804.

4. The Department’s comparatively high percentage of female firefighters.

In a 2008 report (Exhibit 40) Madison’s Fire Department was reported to be 15% female, significantly higher than the national

its test was more valid because it selected more capable candidates – Professor Weltman rightly said that since they were two different tests, the performance standards required to pass them could not be compared. (Tr., dkt. # 67, 73:5-9.)

average. (Tr., dkt. # 67, 91:4-11.) The district court believed that this fact hurt the Plaintiff's ability to prove that the CPAT was substantially equally as valid as the Madison test because it supported "a possible inference that the CPAT may have a favorable disparate impact on women but results in the washing out of ultimately unsuccessful applicants after the additional expenditure of time and money at the academy phase." (A-40.)

This inference is simply not logically available. Is it *possible* that some fire departments have low percentages of female firefighters because a significant number of women who pass their physical entry tests wash out during training? Yes, it's possible, but it's just a guess, and it is not based on any evidence, much less on evidence that Madison has achieved a relatively high female representation percentage. Is it possible that fewer Madison female firefighters wash out during training because they have passed a harder test to get in than female firefighters in other departments? Yes, it's possible, but it is also just a guess. In fact, Professor Weltman testified that it would be unlikely that more females would

wash out of the academy after passing the CPAT than after passing the Madison test. (Tr., 96:14-23.)

The most logical inference regarding a reason for Madison's relatively high percentage of female firefighters is not that they have to pass a harder test to get into the academy. It is that Madison has achieved a high percentage of females by hiring every single female who passes its Physical Abilities Test, while it hires only about 3% of the men who pass. (A-3.)

CONCLUSION

This Court should hold that the district court's conclusion that the CPAT is not substantially equally as valid as the Madison Physical Abilities Test is clearly erroneous because it reflected a failure to apply crucial legal principles, reverse it, and remand for the entry of judgment for the Plaintiff. If it does this, the Court should also reverse the district court's decision granting summary judgment to the City on the issue of whether, if she had passed a lawful physical test, Ms. Erdman would more probably than not

have been offered a position as a Madison firefighter, and remand the case for a trial on this issue.

Dated this Wednesday, March 15, 2023.

Respectfully submitted,

Catherine Erdman,

Plaintiff-Appellant,

By

THE JEFF SCOTT OLSON LAW FIRM, S.C.

Jeff Scott Olson

(Counsel of Record)

State Bar No. 1016284

131 W. Wilson Street, Suite 1200

Waunakee, WI 53597-2502

Phone: 608 283 6001

Facsimile: 608 283 0945

e-mail: jsolson@scofflaw.com

/s/Jeff Scott Olson

JEFF SCOTT OLSON

Attorneys for Plaintiff-Appellant

Rule 32(a)(7) CERTIFICATION

This brief complies with the type-volume limitation of Rule 32(a)(7) inasmuch as it does not exceed 14,000 words, from the Jurisdictional Statement through the Conclusion. The length of this brief is 13,152 words.

Dated this Wednesday, March 15, 2023.

/s/Jeff Scott Olson

Jeff Scott Olson

FORMER CIRCUIT RULE 31(e) CERTIFICATE

The undersigned counsel of record hereby certifies as follows:

1. This brief has been reproduced digitally in non-scanned pdf format and uploaded to the Court's website.

/s/Jeff Scott Olson

Jeff Scott Olson

CERTIFICATE OF SERVICE

I hereby certify that on Wednesday, March 15, 2023, I electronically filed the foregoing with the Clerk of the Court for the United States Court of Appeals for the Seventh Circuit by using the CM/ECF system. I certify that all participants in the case are registered CM/ECF users and that service will be accomplished by the CM/ECF system.

/s/Jeff Scott Olson

Jeff Scott Olson

In the
United States Court of Appeals
For the Seventh Circuit
Chicago, Illinois 60604

CATHERINE ERDMAN,

Plaintiff - Appellant,

v.

CITY OF MADISON,

Defendant - Appellee.

On Appeal from the United States District Court
For the Western District of Wisconsin,
The Honorable Judge William M. Conley Presiding,
District Court Case No: 3:16-cv-00786-wmc

Short Appendix of Plaintiff-Appellant

Jeff Scott Olson
THE JEFF SCOTT OLSON LAW FIRM, S.C.
1025 Quinn Drive, Suite 500
Waunakee, WI 53597-2502
608-283-6001
jsolson@scofflaw.com

Counsel of Record for Plaintiff-Appellant

STATEMENT THAT ALL REQUIRED MATERIALS ARE IN APPENDIX

It is hereby certified that the Short Appendix includes all of the materials required by Circuit Court Rule 30 (a) and (b).

/s/Jeff Scott Olson

Jeff Scott Olson

CONTENTS OF SHORT APPENDIX

<u>Record Item</u>	<u>Page</u>
Opinion and Order on Summary Judgment (dkt. # 47).....	1
Opinion and Order after Trial to the Court (dkt. # 81).....	23
Judgment for Defendant (dkt. # 82).....	42

IN THE UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF WISCONSIN

CATHERINE ERDMAN,

Plaintiff,

v.

CITY OF MADISON,

Defendant.

OPINION AND ORDER

16-cv-786-wmc

Plaintiff Catherine Erdman filed this civil action against defendant City of Madison, claiming its fire department uses a physical abilities test in hiring that has a disparate impact on women in violation of Title VII of the Civil Rights Act of 1964, as amended, 42 U.S.C. § 2000e, *et seq.* Before the court is the City's motion for summary judgment, asserting that: plaintiff was not subject to disparate treatment on the basis of sex, and in any event, plaintiff is not entitled to front or back pay damages. (Dkt. #12.)¹ In response, plaintiff argues that she has made a *prima facie* case of disparate impact and that the City has failed to meet its burden of demonstrating that the physical abilities test is job-related for the position in question and consistent with business necessity.² Because the court agrees that plaintiff has put forth sufficient evidence from which a reasonable fact finder could find in favor of plaintiff on all three elements, this case will

¹ In its motion, the City also argues that: plaintiff is only entitled to equitable remedies; and plaintiff is not entitled to a jury trial. Plaintiff concedes both of these arguments.

² At the end of her opposition brief, plaintiff also seeks summary judgment as to the City's liability, albeit in cursory fashion. (Dkt. #19.) For the reasons explained below, there may be grounds to enter partial summary judgment in plaintiff's favor, finding that she has made out a *prima facie* case of disparate impact. Because this case will be tried to the bench, however, the court is inclined to review the statistics and law as part of the bench trial and make any determination as to liability at that time.

proceed to trial. As for the availability of front or back pay, the court agrees plaintiff has failed to put forth sufficient evidence at summary judgment to support a finding that it was “reasonably clear” Erdman would have been offered a firefighter position in 2014, but for the physical abilities test. Accordingly, the court will grant in part and deny in part defendant’s motion.

UNDISPUTED FACTS³

A. Department’s 2014 Recruitment Process

Defendant City of Madison (“City”) operates a fire department (“Fire Department” or “Department”) that employs approximately 365 firefighters. Each year, the Fire Department engages in a recruitment process to fill vacancies. The hiring process consists of multiple stages, including: (1) an application screening for minimum qualifications; (2) a written test; (3) the physical abilities test (“PAT”); (4) an oral board; and (5) the chief’s interview.

The posting for the 2014 recruitment identified the following physical requirements for firefighter positions:

While not an exclusive list, the following examples are meant to illustrate some of the extreme physical demands and working conditions inherent in the role of a firefighter.

Physical Demands

1. Pick up and advance charged fire hoses.*
2. Force entry with axe/battering ram.*
3. Rescue/extricate victim(s).*
4. Perform CPR; apply bandages.

³ Unless otherwise noted, the court finds the following facts undisputed and material for the purpose of deciding the present motion.

5. Climb stairs carrying heavy equipment, while wearing firefighter protective clothing that weighs in excess of 50 pounds.*
6. Strip and vent roofs, breach walls, overhaul burned buildings.*
7. Lift and climb/descend ladders (with victims).*
8. Visually determine fire status/hazards; assess patient conditions.
9. Hear calls for help; identify fire noise, etc.
10. Walk on roof tops under adverse conditions.
11. Operate power tools and extrication equipment; tie knots.
12. Stoop, crawl, crouch, and kneel in confined spaces.*
13. Reach, twist, balance, grapple, bend and lift under emergency conditions.
14. Run, dodge, jump and maneuver with equipment.*
15. All of the above may be performed wearing heavy and restrictive protective clothing/gear in excess of 50 pounds.*

Each task marked by an asterisk was assessed in the 2014 PAT.

A total of 1887 applicants participated in the 2014 recruitment. Of these, 1723 were men, 146 were women, and 18 were not clearly identified by gender. Four hundred and ninety-nine applicants appeared to take the PAT -- 471 men and 28 women. Of these, 404 applicants -- 395 men, four women, and five not clearly identified -- successfully completed the PAT. Ultimately, the Fire Department hired four women and thirteen men after its 2014 recruitment.

B. Development and Implementation of the PAT

From 1997 through 2014, the Fire Department contracted with the following groups to develop and implement a hiring process for candidates: Landy Jacobs and Associates; SHL Landy Jacobs, Inc.; SHL USA; EB Jacobs, LLC; and Ergometrics & Applied Personnel Research. The 2014 PAT emerged out of this process.

In 1997, Landy Jacobs and Associates, Inc., provided the Department with a Firefighter Job Analysis, Test Development, Administration and Scoring Report. The 1997 PAT included seven events: (1) equipment shuttle; (2) ladder event; (3) hose drag; (4) sledgehammer event; (5) search; (6) rescue; and (7) pike pole.⁴ Each PAT component was assigned a so-called “cut-score” and a minimally acceptable score, the latter representing a performance level below the cut-score. If a candidate met or exceeded the cut-score, the candidate received a score of one for that event. If the candidate met or exceeded the minimally acceptable score, but fell short of the cut-score, then the candidate received a score of zero for that event. However, if the candidate failed to meet the minimally acceptable score for any event, the candidate was disqualified and eliminated from the selection process.⁵ To proceed from the PAT to the next stage -- the oral boards -- a candidate not only had to complete all seven events with at least a minimally acceptable score of zero, but also had to receive an overall score of at least five points (*i.e.*, meet or exceed the cut score for at least five events).

To determine cut-scores for firefighter candidates participating in the 1997 PAT, Landy Jacobs and Associates drew on PAT scores from 94 incumbent firefighters. The demographic characteristics of the incumbent sample were not indicated in Landy’s report. The cut-score for each event was set at one standard deviation off the mean of

⁴ The parties dispute whether the tested events correlate directly to critical tasks performed by Department firefighters. (Def.’s Corr. Resp. to Pl.’s PFOF (dkt. #41) ¶ 19.)

⁵ For example, a candidate had to complete at least 20 repetitions to meet the cut score and receive one point for the pike pole, and 16 repetitions to meet the minimal acceptable score, while any score lower than 16 repetitions would result in outright disqualification from the recruitment process.

incumbent scores.⁶ For a timed event, the cut-score would be the mean time of incumbent scores plus one standard deviation; for an event involving repetitions, the cut-score would be the mean number of repetitions performed by incumbents minus one standard deviation. The parties dispute how the 1997 PAT minimally acceptable performance standards were determined. (*See* Pl.’s Corr. Resp. to Def.’s PFOFs (dkt. #41) ¶ 25.)

In 1999, the Fire Department contracted with SHL Landy Jacobs, Inc., to administer essentially the same selection system that had been used in 1997, with minor modifications. SHL Landy Jacobs modified some of the events and developed new minimally acceptable performance standards and cut-scores by drawing on PAT scores from 102 incumbent firefighters. To determine cut-scores, SHL Landy Jacobs eliminated the top and bottom five percent of scores from each event before calculating the mean and standard deviation. The parties dispute whether SHL Landy Jacobs’ report provided the demographic characteristics of the incumbent sample, but the report does indicate that the sample closely matched the Department’s overall gender balance. (*See* Olson Decl., Ex. C (dkt. #23-3) 14-15.) As with the 1997 PAT, the parties dispute how the 1999 PAT minimally acceptable performance standards were determined. (*See* Pl.’s Corr. Resp. to Def.’s PFOFs (dkt. #41) ¶ 29.)

From 1999 to 2011, the Fire Department did not alter the PAT. In 2013, the

⁶ In statistics, the standard deviation is a measure that is used to quantify the amount of variation or dispersion of a set of data values, calculated as:

$$SD = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

“Standard deviation,” Wikipedia, https://en.wikipedia.org/wiki/Standard_deviation.

Department contracted with Ergometrics & Applied Personnel Research, Inc. (“Ergometrics”) to develop a new “validation report” for the hiring process. As a result, Ergometrics did not alter the 1999 PAT, but rather worked with the Department to validate the PAT before its 2014 recruitment. The validation study was based on the performance of nineteen incumbent firefighters -- 16 males and three females. Relying on this 2014 validation report, the Department maintained the cut-scores adopted in 1999.

The parties dispute whether the 2014 PAT was reliable, and whether it validly measured the skills that a firefighter needs to successfully perform her duties. The court will touch on these disputed facts further in its opinion below.

C. Plaintiff’s Background and Applications

Plaintiff Catherine Erdman is a woman and has been a firefighter/EMT-B for the City of Janesville, Wisconsin, since 2007. She applied for a position with the Department in 2013 and 2015, and she participated in the 2014 and 2016 recruitment process.

During the 2014 recruitment process, Erdman met the minimum qualifications and passed the application screening and the written test, but failed the PAT. Erdman passed the following PAT components: equipment shuttle, hose drag, sledgehammer event, search and rescue to achieve five total points. While Erdman did not receive a passing score on the ladder event, she nevertheless attained the minimum acceptable score required to avoid disqualification. However, Erdman failed to complete the minimum acceptable score of 16 repetitions in the pike pole event. Because Erdman

completed only 12 repetitions on that event, she was not hired.

Erdman again participated in the Fire Department’s 2016 recruitment. This time, she successfully completed the PAT and advanced to the chief’s interview stage. Unfortunately, she was not hired once again. Erdman still desires to work for the Department, and she intends to apply during future recruitments.

D. Statistical Evidence of Disparate Impact⁷

The following chart illustrates the results of the 2014 PAT by sex:

2014 PAT by Sex						
Category	Males			Females		
	Percentage Against Males Who Appeared	Percentage Against Males Who Appeared and Did Not Quit	Number	Percentage Against Females Who Appeared	Percentage Against Females Who Appeared and Did Not Quit	Number
Appeared to Take Test			471			28
Quit During Test	2.76%		13	10.71%		3
Did Not Quit During Test	97.24%		458	89.29%		25
Disqualified for Failing to Meet Minimally Acceptable Score	10.40%	10.70%	49	71.43%	80.0%	20
Failed Test	2.97%	3.06%	14	3.57%	4.00%	1
Passed Test	83.86%	86.24%	395	14.29%	16.00%	4

⁷ Any statistical analysis and terminology used in this opinion is drawn from the parties’ briefs, the parties’ experts’ submissions or relevant case law, acknowledging “that the law mandates statistical discussion in this case.” *Ernst v. City of Chi.*, 837 F.3d 788, 805 n.2 (7th Cir. 2016).

The overall pass rate for women who appeared to take the test (4/28) was about 17 percent of the pass rate for men who appeared to take the test (395/471). Women's failure rate (1/28) for the test was about 120% that of men's failure rate (14/471).⁸ The disqualification rate for women who appeared to take the test and did not quit (20/25) was 748 percent that of men who appeared to take the test and did not quit (49/458). Plaintiff's expert stated that these differences were highly significant statistically, and defendant does not dispute that finding for purposes of summary judgment.

The parties have not provided a detailed breakdown the performance of the 2014 candidates in each PAT event by sex, except for how candidates performed on the pike pole -- the event Erdman failed. Only seven female candidates made it to the pike pole, and all but one -- the plaintiff -- passed, for a pass rate of 85.7 percent. Of the 395 males who made it to the pike pole component, fourteen were disqualified, for a pass rate of 96.5 percent. Furthermore, it is clear that by the time of the pike pole event, 72 percent (18/25) of the female candidates who had appeared to take the PAT and did not quit had already been disqualified.⁹

OPINION

Federal Rule of Civil Procedure 56(a) requires that the court grant summary judgment "if the movant shows that there is no genuine dispute as to any material fact

⁸ As used here, "failure" means an applicant met the minimally acceptable score for each of the seven events, but failed to meet the cut-score for at least five of the seven events.

⁹ While the parties submit proposed findings of facts concerning a proposed alternative test, they are largely disputed as described below. Accordingly, resolution of those facts will have to await trial.

and the movant is entitled to judgment as a matter of law.” Fed. R. Civ. P. 56(a). In considering defendant’s motion, “[t]he evidence of [plaintiff as] the non-movant is to be believed, and all justifiable inferences are to be drawn in [her] favor.” *Anderson v. Liberty Lobby, Inc.*, 477 U.S. 242, 255 (1986). Summary judgment is appropriate only if “the record taken as a whole could not lead a rational trier of fact to find for the non-moving party.” *Sarver v. Experian Info. Solutions*, 390 F.3d 969 (7th Cir. 2004) (quoting *Matsushita Elec. Indus. Co. v. Zenith Radio Corp.*, 475 U.S. 574 (1986)).

Because both parties here ultimately seek summary judgment, the court will “look to the burden of proof that each party would bear on an issue of trial then require that party go beyond the pleadings and affirmatively to establish a genuine issue of material fact.” *Santaella v. Metro. Life Ins. Co.*, 123 F.3d 456, 461 (7th Cir. 1997). If either party “fails to make a showing sufficient to establish the existence of an element essential to that party’s case, and on which that party will bear the burden at trial,” summary judgment against that party is appropriate. *Mid. Am. Title Co. v. Kirk*, 59 F.3d 719, 721 (7th Cir. 1995) (quoting *Tatlovich v. City of Superior*, 904 F.2d 1135, 1139 (7th Cir. 1990)). A summary judgment determination is essentially an inquiry as to “whether the evidence presents a sufficient disagreement to require submission to a jury or whether it is so one-sided that one party must prevail as a matter of law.” *Liberty Lobby*, 477 U.S. at 251–52.

Plaintiff claims that defendant’s hiring process uses a physical abilities test that has a disparate impact on women in violation of Title VII. 42 U.S.C. § 2000e-2(k). For the reasons that follow, the court finds that plaintiff has produced sufficient evidence to

permit the trier of fact to find a *prima facie* showing of disparate impact. However, genuine disputes as to material facts preclude entry of summary judgment.

I. Remedies and Right to Jury Trial

As an initial matter, plaintiff is entitled only to equitable remedies and is not entitled to a jury trial. Under 42 U.S.C. § 1981a, a plaintiff seeking to recover compensatory damages must show that the defendant “engaged in unlawful intentional discrimination (not an employment practice that is unlawful because of its disparate impact) prohibited under . . . the Act.” 42 U.S.C. § 1981a(a)(1). Additionally, under 42 U.S.C. § 1981, a party may demand a jury trial only when seeking compensatory or punitive damages. 42 U.S.C. § 1981a(c). In the present case, plaintiff brings only a disparate impact claim. As plaintiff concedes, this means that she is entitled only to equitable remedies and, therefore, she is not entitled to trial before a jury. (Pl.’s Corr. Opp’n (dkt. #32) 19.) As a result, the court previously modified the scheduling order to reflect that the case will be tried to the bench, rather than to a jury.

II. Disparate Impact

Title VII prohibits hiring practices that have a disproportionately adverse impact on employees with protected characteristics, such as sex, even if there is no intent to discriminate. *Ernst v. City of Chi.*, 837 F.3d 788, 794 (7th Cir. 2016). To prove such a “disparate impact,” plaintiff must show that a particular hiring practice had an adverse impact on applicants with a protected characteristic, such as sex. 42 U.S.C. § 2000e-2(k)(1)(A)(i); *Ernst*, 837 F.3d at 796. As part of the *prima facie* case, the plaintiff must

also show causation, typically by “offer[ing] statistical evidence of a kind and degree sufficient to show that the practice in question has caused the exclusion of applicants for jobs or promotions because of their membership in a protected group.” *Watson v. Fort Worth Bank & Tr.*, 487 U.S. 977, 994 (1988).

Where an employee has made a *prima facie* showing, an employer can defend by showing that: (1) the challenged practice does not cause the disparate impact, 42 U.S.C. § 2000e-2(k)(B)(ii); or (2) the practice is job-related for the position and consistent with business necessity, 42 U.S.C. § 2000e-2(k)(1)(A)(i). *See Griggs v. Duke Power Co.*, 401 U.S. 424, 432 (1971). Although employers have the burden of proof as to each defense, they “are not required, even when defending standardized or objective tests, to introduce formal ‘validation studies’ showing that particular criteria predict actual on-the-job performance.” *Watson*, 487 U.S. at 998. Assuming the defendant is able to make a sufficient showing as to either defense, the burden shifts back to the applicant to prove that the employer refuses to adopt an alternative hiring practice resulting in less disparate impact and serving the employer’s legitimate needs. 42 U.S.C. §§ 2000e-2(k)(1)(A)(ii), (C); *Ernst*, 837 F.3d at 794; *see also* 29 C.F.R. § 1607.3B. (“[T]he user should use the procedure which has been demonstrated to have the lesser adverse impact.”).

Defendant argues that it is entitled to summary judgment on several grounds. Principally, defendant asserts that plaintiff has not made out a *prima facie* case of disparate impact, and more particularly, that the PAT did not have a disparate impact with respect to the pike pole event on which this particular plaintiff was disqualified. Next, defendant asserts that even if there was a disparate impact, the PAT was “job

related” and “consistent with business necessity.” Finally, defendant asserts that plaintiff has failed to put forth evidence of an alternative employment practice with a less adverse impact. The court addresses each of these arguments separately below.

A. Relevance of the PAT as a Whole and the Disqualifying Component in Particular

According to defendant, the disparate impact analysis must focus on the part of the PAT that disqualified plaintiff -- here, the pike pole event -- rather than on the entire PAT. In support, defendant cites to the Supreme Court’s decisions in *Connecticut v. Teal*, 457 U.S. 440 (1982), and *Watson v. Fort Worth Bank & Tr.*, 487 U.S. 977 (1988), for the proposition that “the proper focus is on the employment requirement that created the bar to opportunity.” (Def.’s Reply (dkt. #40) 5.) Defendant further argues that there is no proof of disparate impact as to the pike pole event, which disqualified plaintiff, and that, therefore, plaintiff has not met her initial burden.

In *Teal*, the Supreme Court rejected a so-called “bottom line” defense to disparate impact claims, which would have shielded employers who used discriminatory practices so long as, at bottom, the workforce was balanced. 457 U.S. at 442. In no way, however, does *Teal*’s rejection of the “bottom line” defense limit plaintiff’s disparate impact claim to the specific part of the PAT that disqualified her.

In *Watson*, the Court introduced the requirement that a plaintiff asserting a disparate impact do so by “isolating and identifying the specific employment practices that are allegedly responsible for any observed statistical disparities.” 487 U.S. at 994; *see also* 42 U.S.C. 2000e-2(k)(1)(A)(i) (complaining party must demonstrate that

respondent “uses a particular employment practice that causes a disparate impact” on an impermissible basis). Certainly, *Watson* stands for the proposition that a plaintiff must identify and target a particular employment practice. The policy behind this requirement is also clear enough: a disparate impact claim based on a discretionary system as a whole or a general policy “could lead to employers being held liable for the ‘myriad of innocent cases that may lead to statistical imbalances.’” *Puffer v. Allstate Ins. Co.*, 675 F.3d 709, 717 (7th Cir. 2012) (quoting *Watson*, 487 U.S. at 994). Here, plaintiff asserts just that: the PAT as a whole is the “particular employment practice that causes a disparate impact on the basis of . . . sex.” 42 U.S.C. § 2000e-2(k)(1)(A)(i).

Although it is relatively simple in this case to separate out the particular PAT event that ultimately disqualified this plaintiff in 2014, she disputes that her challenge is limited to proving the disparate impact of the pike pole event alone -- as opposed to the entire PAT -- even if the impact of each component, including the pike pole event, can be measured individually,. On this, plaintiff appears to have the better of the law to date. First, nothing about the Court’s analysis, reasoning or ultimate holding in *Watson* appears to *require* plaintiff to isolate a component of the PAT in asserting a disparate impact claim. Second, neither the parties nor this court in its own research could find a case analyzing whether a “particular employment practice” must be limited to the particular component of a physical abilities test or of a similar multi-component test. Third, courts evaluating similar cases involving PATs appear to have uniformly considered the entire PAT as an indivisible hiring practice. *See, e.g., Pietras v. Bd. of Fire Comm’rs of Farmingville Fire Dist.*, 180 F.3d 468, 475 (2d Cir. 1999) (comparing pass rate

for women and pass rate for men on PAT as a whole); *Arndt v. City of Colorado Springs*, 263 F. Supp. 3d 1071, 1075 (D. Colo. 2017) (requirement that all sworn officers pass PAT annually was a specific employment practice); *cf. Ernst*, 837 F.3d at 796 (“[P]hysical-skills entrance test has an adverse impact on women[.]”). While defendant argues that cases like *Ernst* involve multiple plaintiffs, and thereby warrant a broader impact analysis, this argument misunderstands the nature of a disparate impact claim: a disparate impact claim challenges a facially neutral practice that has a disproportionately adverse impact on applicants who share a protected characteristic. *See Watson*, 487 U.S. at 986–87; *Bennett v. Roberts*, 295 F.3d 687, 698 (7th Cir. 2002).¹⁰

In fairness, none of these past decisions expressly considered the question of whether the particular employment practice can or should be limited to a component of a PAT, perhaps because the facts and arguments never presented itself so starkly. From the court’s understanding of a disparate impact claim, however, if multiple successive components of the PAT are collectively discriminatory, as plaintiff claims they are, she need not demonstrate a disparate impact as to the last component simply because she succeeded in overcoming earlier discriminatory components. Similarly, if an employer uses a hiring practice that discriminatorily eliminates members of the protected class at each event level, the employer may not be shielded from liability because of the smaller sample size at each successive event level. Even if such a burden could befall on this

¹⁰ Defendant’s argument that a greater number of plaintiffs widens the scope of the disparate impact analysis ignores the fact that even when *one* individual brings a disparate impact case, the analysis focuses on the *class* of persons who share the protected characteristic. That is, the claim is that a facially neutral practice discriminates against a protected class of people, not against the plaintiff alone.

plaintiff, a factual dispute remains as to whether the combined demands of all of the components of the PAT, which are extremely physical and occur consecutively, support evaluating the PAT as a single hiring practice. *See* 42 U.S.C. § 2000e-2(k)(B)(i).

Moreover, the disparate impact provision in Title VII states that, if the plaintiff can show that elements of a hiring process are not capable of separation for analysis, the court may analyze them as a single employment practice. 42 U.S.C. § 2000e-2(k)(B)(i). One could argue that this provision simply requires the plaintiff to isolate one of the five hiring stages described above as she has with respect to the PAT, rather than a particular component of one of those stages. Still, because the individual components of the PAT are capable of separation perhaps, defendant's position is not meritless. As such, while the court will deny defendant's motion for summary judgment as to this argument, the court will require the parties to further brief this narrow issue in advance of the final pretrial conference and will consider the law and facts afresh at the close of trial.

Assuming the correct unit of measurement is the PAT as a whole, the remaining question is whether plaintiff has produced sufficient statistical evidence to allow a reasonable factfinder to conclude that the challenged hiring practice -- the PAT -- has a disproportionately negative effect upon members of the plaintiff's protected class. *See Noreuil v. Peabody Coal Co.*, 96 F.3d 254, 259 (7th Cir. 1996). As noted above, the overall pass rate for women who appeared to take the test (4/28) was about 17% of the pass rate for men who appeared to take the test (395/471). The disqualification rate for women who appeared to take the test and did not quit (20/25) was 748% that of men who appeared to take the test and did not quit (49/458). Plaintiff's expert has stated

that these differences were highly significant statistically, and defendant does not dispute those statistics. (See Pl.'s Corr. Resp. to Def.'s PFOFs (dkt. #41) ¶¶ 155-63.) As such, a reasonable factfinder *could* conclude that plaintiff has made out a *prima facie* case of disparate impact on the basis of sex caused by a specific employment practice (the PAT).

B. Relationship of PAT to Job

Assuming plaintiff makes out a *prima facie* case of disparate impact at least for the PAT as a whole, the burden shifts to the defendant to show that the practice is job-related for the position and consistent with business necessity. 42 U.S.C. § 2000e-2(k)(1)(A)(i); *Griggs*, 401 U.S. at 432. “A test is ‘job-related’ if it measures traits that are significantly related to the applicant’s ability to perform the job.” *Gillespie v. State of Wis.*, 771 F.2d 1035, 1039 (7th Cir. 1985). However, “[e]mployers are not required to support their physical-skills tests with formal validations studies, which show that particular criteria predict actual on-the-job performance.” *Ernst*, 837 F.3d at 796 (internal citation and quotation marks omitted). However, “[w]hen an employer relies on a validity study, federal regulations establish technical standards for these studies.” *Id.* (citing 29 C.F.R. § 1607.14(B)(4)).

In *Ernst v. City of Chicago*, 837 F.3d 788 (7th Cir. 2016), the Seventh Circuit recently considered a disparate impact challenge to a physical abilities test used by the City of Chicago in screening for paramedic positions, concluding that the City failed to show that its test was job-related and consistent with business necessity. Reviewing the technical requirements for validating a test at length, the court described the importance of “determining the extent to which a study accurately measures what it sets out to

measure.” As for evaluating the study in *Ernst*, the Seventh Circuit noted a number of flaws, including the sample population’s abnormally high-performance scores, and the attempt to normalize those results by including scores from New York paramedics, who presumably performed below than the population sample from among Chicago paramedics. 837 F.3d at 801. The court also questioned whether the “work samples actually test the skills that Chicago paramedics learn on the job,” specifically pointing to the test’s failure to replicate: (1) the distance, including set of stairs, required to carry equipment; (2) the weight of the patients; (3) the use of stair chairs; and (4) the method of carrying a stretcher, among other concerns. *Id.* at 802-04.

Largely through the testimony of her expert Arthur L. Weltman, Ph.D., FASCM, the plaintiff here calls into question the validity and reliability of the PAT, offering similar criticisms to those credited in *Ernst*. Specifically, plaintiff challenges the reliability of the PAT based on defendant’s failure to provide practice sessions as part of the test formulation and the lack of information about “test-retest” reliability of scores in the various PAT reports. Moreover, plaintiff faults defendant for failing to provide a rationale for adopting the cut-scores or the minimally acceptable performance standards.

Finally, with respect to whether the tested tasks replicate actual firefighting duties, plaintiff challenges two components of the PAT. First, with respect to the ladder event, plaintiff contends that it fails to map against the actual requirements of a City of Madison firefighter, pointing out that the length of the ladder used (20 feet) for testing is significantly longer than the longest ladder on the sides of vehicles used by the department (14 feet), and that under normal circumstances, placing the ladder back on a

truck would not be a time-sensitive part of the job. Therefore, plaintiff argues it should not even be included in the timed component of the event. Similarly, with respect to the pike pole event, plaintiff challenges whether the test's constraint as to the placement of the person vis-à-vis the piece of ceiling to be pulled down sufficiently replicates real firefighting work and in particular, whether it unfairly affects shorter people's ability to perform that element of the PAT.

The City principally responds to plaintiff's attempt to use *Ernst*, and more specifically, plaintiff's reliance on the Seventh Circuit's criticism of the paramedics physically abilities test on the basis that the study at issue was a "criterion" study, whereas defendant describes its study as a "content" one. *See generally* 29 C.F.R. 1607.14. While the specific validity requirements for these two tests differ somewhat, the regulations for a content study similarly require a selection procedure that ensures a "representative sample of the content of the job," and that the skill to be tested is a "necessary prerequisite to successful job performance." 29 C.F.R. § 1607.14(C)(1). Indeed, in determining whether the test is content valid, the Seventh Circuit directed district courts to consider:

(1) the degree to which the nature of the examination procedure approximates the job conditions; (2) whether the test measures abstract or concrete qualities; and (3) the combination of these factors, i.e. [sic] whether the test attempts to measure an abstract trait with a test that fails to closely approximate the working situation.

Bryant v. City of Chi., 200 F.3d 1092, 1099 (7th Cir. 2000) (quoting *Gillespie v. State of Wis.*, 771 F.2d 1035, 1043 (7th Cir. 1985)) (alteration in *Bryant*). While the PAT at issue here was presumably validated under a different method, the general requirements

overlap significantly with the criterion test criticized in *Ernst*.

Regardless of the specific applicability of *Ernst*, there is a genuine dispute as to whether the PAT was “job related” in light of plaintiff’s expert testimony, as well as “consistent with business necessity.” Although defendant may ultimately make a showing sufficient to establish both, the court cannot adequately weigh the evidence provided by the parties’ experts at summary judgment.

C. Alternative Employment Practices

Because the court will not enter summary judgment as to whether the 2014 PAT was “job related” and “consistent with business necessity,” it is unnecessary to consider whether there was an alternative employment practice with a less adverse impact. If defendant succeeds at trial in showing that the 2014 PAT was “job related” and “consistent with business necessity,” plaintiff will be given the opportunity of presenting evidence that there was an alternative with a less adverse impact.

At summary judgment, it is enough to observe that plaintiff appears to have put forth sufficient evidence for a trier of fact to find that the Candidate Physical Abilities Test (“CPAT”) -- a test created by the International Association of Fire Chiefs’ and the International Association of Fire Fighters’ Joint Labor Management Task Force that the City considered adopting but eventually rejected -- is an alternative employment practice with a less adverse impact. The City contends that plaintiff has failed to show that this test (1) would have a less adverse impact with respect to female firefighters or (2) is a valid one for the City of Madison Fire Department. In response, plaintiff proffers evidence that the CPAT pass rate for women was 68.0%, higher than other physical

abilities tests for women generally, and, of course, significantly higher than the PAT pass rate of 14% for women in the Department's 2014 recruitment. Moreover, the CPAT appears to address two of the core concerns raised about the PAT above: (1) modifying the ladder event to focus on the extension of the ladder, rather than the carrying component; and (b) modifying the pike pole event to allow shorter candidates to stand at a closer yet still safe location.

As such, there is a genuine dispute of material fact as to this third element -- whether there is an alternative employment practice with a less adverse impact -- precluding summary judgment.

III. Front and Back Pay Damages

Finally, defendant argues that plaintiff's request for front and back pay damages is overly speculative. First, defendant notes that only 17 of 404 applicants who successfully completed the 2014 PAT were hired. Second, defendant points out that plaintiff was not hired in 2016, even after successfully completing the PAT and making it to the chief's interview. Accordingly, defendant asserts that the Seventh Circuit's decision in *Evans v. City of Evanston*, 881 F.2d 382 (7th Cir. 1989), should control.

In *Evans*, a class of female applicants who failed a PAT requested, *inter alia*, a new PAT and back pay relief. 881 F.2d at 386. The court held that such make-whole relief was appropriate "where it is reasonably clear that, had it not been for the discriminatory behavior, the plaintiff would have got (or retained) the job or other employment benefit in issue, and where making the plaintiff whole would not unduly injure innocent third parties." *Id.* Because "only 1.2 percent of the applicants who passed the [PAT] . . . were

actually hired,” the court held “what the class members lost was not a job but a long-shot chance at a job.” *Id.* The appropriate make-whole relief under those circumstances was for the class members to “be restored to the place they would have occupied if they pass a new physical agility test approved by the district court.” *Id.*

In the face of *Evans*, plaintiff responds that all four female candidates who passed the PAT in 2014 were ultimately hired. Plaintiff argues, therefore, that this means that a woman who successfully completed the PAT in 2014 had a 100 percent chance of being hired. Under this view, plaintiff asserts, it is more probable than not, if not likely, that she would have been hired but for the discriminatory effect of the PAT.

In the end, plaintiff’s suggestion that defendant would have hired any woman who passed the 2014 PAT goes too far. True, defendant hired all four women who successfully completed the 2014 PAT, but four successful applicants constitutes a very small sample. Additionally, even if plaintiff had passed the 2014 PAT, she still had to complete two more stages: the oral board and the chief’s interview. As defendant points out, plaintiff was not hired in 2016 even though she had progressed to the chief’s interview. Like the *Evans* plaintiffs, what she lost out on in 2014 was not a job, but a chance at a job. If her claim succeeds and defendant is found liable, plaintiff will be placed in the position she would have occupied but for the discriminatory test. However, plaintiff has failed to put forth evidence from which a reasonable factfinder could find it was “reasonably clear” that she would have been hired in 2014 but for the PAT. Accordingly, the court will grant summary judgment to defendant on plaintiff’s request for an award of front or back pay damages, finding such an award would be overly

speculative.

ORDER

IT IS ORDERED that:

- 1) Defendant's motion for summary judgment (dkt. #12) is GRANTED IN PART and DENIED IN PART as set forth above. Specifically, defendant is granted summary judgment on (1) any of plaintiff's claims for compensatory damages; (2) plaintiff's request for a jury trial; and (3) plaintiff's request for front or back pay damages. In all other respects, defendant's motion is denied.
- 2) On or before October 5, 2018, the parties are directed to file briefs further addressing the specific question of whether an individual plaintiff who was not hired based on a failure to pass a specific component of a stage in the hiring process (whether a physical abilities test or otherwise) is required to show disparate impact with respect to that individual component or can relies on proof of the disparate impact of the stage as a whole.

Entered this 19th day of September, 2018.

BY THE COURT:

/s/

WILLIAM M. CONLEY
District Judge

IN THE UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF WISCONSIN

CATHERINE ERDMAN,

Plaintiff,

v.

CITY OF MADISON,

Defendant.

OPINION AND ORDER

16-cv-786-wmc

On October 15 and 16, 2018, the court held a trial to the bench on plaintiff Catherine Erdman's claim that the City of Madison, and more specifically its Fire Department, adopted a physical abilities test ("PAT") that has a disparate impact on women in violation of Title VII of the Civil Rights Act of 1964, as amended, 42 U.S.C. § 2000e, *et seq.* For the reasons explained below, the court now concludes that: (1) plaintiff met her burden of proving that the Fire Department's PAT has an adverse impact on female applicants; (2) defendant met its burden of proving that the PAT is job-related and consistent with business necessity; and (3) plaintiff did not meet her burden of proving that the alternative physical abilities test she identifies, the Candidate Physical Abilities Test ("CPAT"), will serve the Fire Department's legitimate needs. Accordingly, the court will find in defendant's favor.

ADDITIONAL FINDINGS OF FACTS¹

A. Disparate Impact of the PAT on Female Applicants

1. A total of 1887 applicants participated in the 2014 recruitment. Of these, 1723

¹ The court's opinion and order on summary judgment set forth a number of facts that were

were men, 146 were women, and 18 were not clearly identified by gender.

2. Four hundred and ninety-nine applicants appeared to take the PAT -- 471 men and 28 women. Of these, 404 applicants -- 395 men, four women, and five not clearly identified -- successfully completed the PAT.

3. Excluding those not self-identifying with either gender, the following chart illustrates the results of the 2014 PAT by sex:

2014 PAT by Sex						
Category	Males			Females		
	Percentage Against Males Who Appeared	Percentage Against Males Who Appeared and Did Not Quit	Number	Percentage Against Females Who Appeared	Percentage Against Females Who Appeared and Did Not Quit	Number
Appeared to Take Test			471			28
Quit During Test	2.76%		13	10.71%		3
Did Not Quit During Test	97.24%		458	89.29%		25
Disqualified for Failing to Meet Minimally Acceptable Score	10.40%	10.69%	49	71.43%	80.0%	20
Failed Test	2.97%	3.06%	14	3.57%	4.00%	1
Passed Test	83.86%	86.24%	395	14.29%	16.00%	4

undisputed, concluding that issues of material fact precluded entry of summary judgment in defendant's favor. (Dkt. #47.) Rather than repeat them, this opinion incorporates those undisputed facts and limits this discussion to the additional findings material to the court's ultimate legal conclusions.

4. As reflected in the chart, the overall pass rate for women who appeared to take the test (4/28 or 14.29%) was about 17% of the pass rate for men who appeared to take the test (395/471 or 83.86%).

5. Conversely -- the women's failure rate -- defined as applicants who met the minimally acceptable score for each of the seven events, but failed to meet the cut-score for at least five of the seven events -- of 1 out of 28 (3.57%), for the test was roughly 120% that of men's failure rate of 14 out of 471 (2.97%).

6. Finally, the women's *disqualification* rate -- defined as those who appeared to take the test and did not quit -- of 20 out of 25 (80%) was 748% that of men's disqualification rate of 49 out of 458 (10.69%).

B. Job-Relatedness and Business Necessity of PAT

7. Debra Amesqua became the Madison Fire Department Chief in 1996. Following her appointment, Chief Amesqua engaged Landy, Jacobs and Associates ("LJA"), to develop the Department's PAT in 1997.

8. Directed by Amesqua to develop a test that correlated with the tasks on the job, LJA developed the PAT under the Uniform Guidelines on Employee Selection Procedures (1978), 29 C.F.R. § 1607, *et seq.*

9. In particular, Rick R. Jacobs, Ph.D., an industrial psychologist, was one of the individuals who developed the PAT in conjunction with exercise physiologists. Jacobs had developed physical ability tests for a little more than a decade before taking on the task of developing the PAT at issue here.

10. In developing a PAT, LJA focuses on job simulation activities, rather than traditional exercise-based activities. Chief Amesqua and other fire department personnel were involved in studying the specific jobs as performed by Madison firefighters.

11. LJA completed content validity reports in 1997 and 1999, finding the PAT valid as required under 29 C.F.R. § 1607.14(B)(4).

12. LJA relied on incumbents to establish the cut and disqualifications scores. For the 1997 PAT in particular, LJA drew on PAT scores from 94 incumbent Madison firefighters randomly selected but controlled to reflect the Madison Fire Department's diversity as to race and gender. At that time, roughly 17% of the department were women.

13. In 1999, LJA modified some of the events and developed new minimally acceptable performance standards and cut-scores by drawing on PAT scores from 102 incumbent firefighters.²

14. The cut scores roughly eliminated the bottom 16% of incumbents, which Jacobs opined offered a good method for managing the two errors, passing applicants who are not able to perform the job and eliminating applications who could perform the job. The disqualification score reflected the lowest performing incumbent.

15. In both Jacobs' and Chief Amesqua's experience, applicants perform better than incumbents on the test because they are more motivated. Moreover, Amesqua wanted to use incumbents to set cut scores because of "washout" (attrition) concerns during the subsequent training academy. Specifically, for costs and other reasons, Amesqua wanted

² While there may have been some overlap in the incumbents participating in the test, LJA did not control for that, although it did control for race and gender.

to insure that the applicants entering the academy could perform the physical requirements of the job.

16. Following Chief Steven Davis's appointment to replace Chief Amesqua in late 2012 or early 2013, Ergometrics & Applied Personnel Research, Inc. ("EAPRI") was retained by the Madison Fire Department to validate the PAT again. EAPRI's President, Carl Swander, Ph.D., and his team then conducted a "content validation study." Like the earlier studies by LJA, this study was conducted under the Uniform Guidelines, among other professional publications.

17. As part of the study, EAPRI used Madison Fire Department "subject matter experts" ("SMEs") to determine what the required speed should be of firefighter candidates taking the examination. As was the case in LJA's studies, EAPRI again used incumbents to set the minimum standards requirements. In addition, Ergometrics also provided physiological measurements of incumbents to gain information as to the overall energy required to perform the PAT's discrete tasks and the PAT overall as a whole.

18. Working with the SMEs, EAPRI selected tasks or test events representative of the variety of physical demands Madison firefighter job applicants would be expected to perform on the job.

19. EAPRI next studied 19 incumbents to set the minimum level of performance requirements. In determining the cut off, EAPRI set the scores for the selected task or event at one standard deviation below the mean, which means that: (1) 82% of participants should do this well or better; and (2) the required max time for passing the test is often about 1.4 times as long as the average time for incumbents.

20. EAPRI's method for setting employment screening standards (or cutoffs) is a common practice for ergometric testing and is more objective than setting the time requirements based on observed speeds. EAPRI checked the cut score by considering observations from experienced command staff.

21. Focusing on the specific elements of the PAT, the "ladder test" was developed and changed to reflect the equipment used at the time of the test. In particular, the 20-foot ladder was used when at least some of the Department's rigs still used a 24-foot extension ladder. The choice of ladder, as well as the required manipulation of the ladder, reflected the job requirements at that time.

22. As for the "pike pole test," the PAT used a box or a line to require applicants to stand at least 18 inches back from the ceiling being pulled down with a pike pole to account for safety concerns with standing directly under the ceiling. Further, the pole was long enough to allow applicants of different heights to adjust the hold to account for that differential.

23. While the PAT does not include a practice requirement, in contrast to the CPAT suggested as an alternative by plaintiff, there are other methods for promoting familiarity with the test, including a manual, video and an exercise guide.

24. The content validity reports also did not include a test-retest for reliability, although again, the PAT offered other methods of testing for reliability, including an "interrater agreement," where different administrators of the test agree on the score for the same candidate. LJA confirmed that this test for reliability was completed in 1997 and 1999 while EAPRI, relied on the execution and administration of the PAT to ensure

reliability in 2013.³

25. The posting for the 2014 recruitment identified the following physical requirements for firefighter positions:

While not an exclusive list, the following examples are meant to illustrate some of the extreme physical demands and working conditions inherent in the role of a firefighter.

Physical Demands

1. Pick up and advance charged fire hoses.*
2. Force entry with axe/battering ram.*
3. Rescue/extricate victim(s).*
4. Perform CPR; apply bandages.
5. Climb stairs carrying heavy equipment, while wearing firefighter protective clothing that weighs in excess of 50 pounds.*
6. Strip and vent roofs, breach walls, overhaul burned buildings.*
7. Lift and climb/descend ladders (with victims).*
8. Visually determine fire status/hazards; assess patient conditions.
9. Hear calls for help; identify fire noise, etc.
10. Walk on roof tops under adverse conditions.
11. Operate power tools and extrication equipment; tie knots.
12. Stoop, crawl, crouch, and kneel in confined spaces.*
13. Reach, twist, balance, grapple, bend and lift under emergency conditions.
14. Run, dodge, jump and maneuver with equipment.*
15. All of the above may be performed wearing heavy and restrictive protective clothing/gear in excess of 50 pounds.*

Each task marked by an asterisk was assessed in the 2014 PAT.

C. Adequacy of CPAT as Alternative Test

26. The CPAT was developed in joint collaboration between the International

³ Swander also explained that test-retest reliability screen required at least 50 participants, which the Madison Fire Department could not provide.

Association of Fire Chiefs and the International Association of Firefighters, in conjunction with ten fire departments in North America, including Austin, Texas, Los Angeles County, California, and New York City.

27. According to a validation study conducted by an exercise physiologist at the University of Texas-Austin, examining the Austin Fire Department, approximately 48% of female participants passed the CPAT.

28. In 2013, the Department considered whether to adopt the CPAT, although the consideration was not because of, or solely because of, concerns about female applicants' performance on the PAT.

29. While there is significant overlap between the tasks in the CPAT and that in the PAT, plaintiff identified four differences between the CPAT and the PAT that her expert maintains are material: (1) the CPAT has an overall time requirement, rather than time requirements for the discrete tasks as is the case for the PAT; (2) the CPAT has two, required practice runs without coaching help, each of which applicants can opt to "run hot" and count as a passed test if the applicant successfully completes it; (3) the cut scores under the CPAT were developed differently using *applicants'* times, rather than incumbents'; and (4) the pike pole and ladder tests have different elements.

30. The CPAT has not been locally validated.

OPINION

Title VII prohibits hiring practices that have a disproportionately adverse impact, also referred to as a "disparate impact," on employees with protected characteristics, such as sex, even if there is no intent to discriminate. *Ernst v. City of Chi.*, 837 F.3d 788, 794

(7th Cir. 2016). To prove her claim here, therefore, plaintiff must show that a particular hiring practice had a disparate impact on female applicants. 42 U.S.C. § 2000e-2(k)(1)(A)(i); *Ernst*, 837 F.3d at 796. As part of her *prima facie* case, the plaintiff must also show causation, typically by “offer[ing] statistical evidence of a kind and degree sufficient to show that the practice in question has caused the exclusion of applicants for jobs or promotions because of their membership in a protected group.” *Watson v. Fort Worth Bank & Tr.*, 487 U.S. 977, 994 (1988). In turn, an employer can defend against a plaintiff’s *prima facie* case by: refuting proof that the challenged practice is a cause of the disparate impact, 42 U.S.C. § 2000e-2(k)(B)(ii); *or* by showing that the practice is job-related for the position and consistent with business necessity, 42 U.S.C. § 2000e-2(k)(1)(A)(i); *see Griggs v. Duke Power Co.*, 401 U.S. 424, 432 (1971).

Although employers have the burden of showing the relationship between a requirement and the employment in question, “employers are not required, even when defending standardized or objective tests, to introduce formal ‘validation studies’ showing that particular criteria predict actual on-the-job performance.” *Watson*, 487 U.S. at 998. Rather, the burden then shifts back to the applicant to prove the employer refuses to adopt an alternative hiring practice that results in a less disparate impact while still serving the employer’s legitimate needs. 42 U.S.C. §§ 2000e-2(k)(1)(A)(ii), (C); *Ernst*, 837 F.3d at 794; *see also* 29 C.F.R. § 1607.3B. (“[T]he user should use the procedure which has been demonstrated to have the lesser adverse impact”).

At summary judgment, defendant argues that plaintiff’s proof falls fatally short at three, independent stages of the disparate impact analysis. First, defendant argues that

plaintiff has not made out a *prima facie* case of disparate impact as to female applicants generally, or at minimum, as to plaintiff in particular given her disqualification was at the component stage of the PAT at issue. Second, defendant argues that even if there was a disparate impact, it successfully established the PAT was “job related” and “consistent with business necessity.” Finally, defendant argues that plaintiff failed to prove a viable alternative employment practice with a less adverse impact. The court addresses each argument in turn below.

I. The PAT Produced a Disparate Impact on Female Applicants

Based on the statistical analysis of the 2014 PAT results described at summary judgment, and summarized above, as supported by expert testimony proffered by plaintiff, the court found that the differences in the performance of male and female applicants during the PAT were highly significant statistically and that a reasonable factfinder could conclude that the PAT produced a disparate impact on female applicants, sufficient to satisfy her burden on the first element of a disparate impact claim. Moreover, at trial, the defendant again did not dispute that plaintiff had satisfied the first element of the claim if the court considered the performances of applicants attempting the PAT as a whole, rather than considering the discrete elements of the PAT that resulted in plaintiff’s failure in particular -- the ladder and pike pole elements -- which she also argues were particularly disadvantageous for females. In its decision on summary and judgment, the court also discussed at length its reasons for analyzing the PAT as a whole as the challenged hiring practice (9/19/18 Op. & Order (dkt. #47) 12-16), and invited additional briefing on this issue from both parties.

Nevertheless, defendant simply regurgitates the same arguments it previously made in its post-trial briefing. (Def.'s Br. (dkt. #54); Pl.'s Br. (dkt. #74) 5-9.) The court sees no reason to revisit this issue as a matter of law, and again concludes that the appropriate unit of analysis for plaintiff's claim is the PAT as a whole. Moreover, the court concludes based on the undisputed record that plaintiff has established by a preponderance of the evidence that the 2014 PAT had an adverse impact on female applicants.

II. The PAT was Job-Related and Consistent with Business Necessity

Given that plaintiff has satisfied her burden of demonstrating that the PAT as a whole produced a disparate impact on women, the burden shifts to defendant to show that the requirements of the PAT are job-related for the position and consistent with business necessity. 42 U.S.C. § 2000e-2(k)(1)(A)(i); *Griggs*, 401 U.S. at 432. “A test is ‘job-related’ if it measures traits that are significantly related to the applicant’s ability to perform the job.” *Gillespie v. State of Wis.*, 771 F.2d 1035, 1039 (7th Cir. 1985). “Employers are not required to support their physical-skills tests with formal validations studies, which show that particular criteria predict actual on-the-job performance.” *Ernst*, 837 F.3d at 796 (internal citation and quotation marks omitted). As is in this case, however, “[w]hen an employer relies on a validity study, federal regulations establish technical standards for these studies.” *Id.* (citing 29 C.F.R. § 1607.14(B)(4)).

While the parties appear to dispute whether § 1607.14 requires the defendant to prove the validity of the PAT itself or the study used to do so, the industrial psychologist firms that the Department employed to validate the PAT used content studies, and there is no dispute that a content study *is* an adequate method of validation. Similar to a

criterion study, the regulations for a content study require a selection procedure that: (1) ensures a “representative sample of the content of the job”; and (2) the skill to be tested is a “necessary prerequisite to successful job performance.” 29 C.F.R. § 1607.14(C)(1). Thus, in determining whether the content of the test is valid, the Seventh Circuit directed district courts to consider:

(1) the degree to which the nature of the examination procedure approximates the job conditions; (2) whether the test measures abstract or concrete qualities; and (3) . . . combin[ing] of these factors, . . . whether the test attempts to measure an abstract trait with a test that fails to closely approximate the working situation.

Bryant v. City of Chi., 200 F.3d 1092, 1099 (7th Cir. 2000) (quoting *Gillespie v. State of Wis.*, 771 F.2d 1035, 1043 (7th Cir. 1985)).

As detailed above, defendant’s experts Rick Jacobs, the developer of the original PAT in 1997, and Carl Swander, who updated and validated the PAT in 2013, both provided extensive testimony about: (1) the job studies conducted to select physical tasks as a proxy for on-the-job requirements, which were conducted with the assistance of senior members of the Department; and (2) the process for setting cut and disqualification scores, utilizing a statistical method and checking it against the subjective assessment of senior members of the Department. (Facts ¶¶ 7-20.) Members of the Department also testified to their involvement in developing and updating the PAT leading up to the challenged 2014 PAT. Given this record, the court concludes that the 2014 PAT was valid under the Uniform Guidelines set forth in § 1607.14.

In fairness, while plaintiff’s expert Arthur Weltman offered some challenges to the PAT: criticizing how the cut and disqualification scores were set; questioning whether the

PAT developers were sufficiently experienced in exercise physiology, which is Weltman's area of expertise; raising concerns about the qualifications of the Department's command officers to serve as subject matter experts; and faulting defendant for failing to engage in a test-retest review. However, none of these concerns directly challenge to the validity of the 2014 PAT under § 1607.14. Indeed, on this critical point, Weltman conceded that he had little familiarity with the Uniform Guidelines and offered no opinion as to whether the PAT was valid in light of them. (Trial Tr. (dkt. #67) 61.)

In addition to failing to show that any of these criticisms -- whether viewed alone or in combination -- undermine defendant's evidence of the validation process for its PAT § 1607.14, including the 2014 PAT at issue, defendant also countered with its own evidence that LJA's and EAPRI's reliance on statistical analysis to set cut and disqualification scores was a widely accepted practice, and witnesses from both consulting groups reliance on the Departments' senior officials in conducting a job study and utilizing other forms of reliability testing, including focusing on execution and administration of the test. EAPRI's President Swander also credibly testified to the difficulty in conducting a test-retest study with a sufficient number of participants given Madison's size (among other constraints), and critiqued the Austin study of the CPAT as having an insufficient survey base to serve its purpose.

The court further concludes that the Department met its obligation to investigate "alternative selection procedures with evidence of less adverse impact . . . to determine the appropriateness of using or validating it in accord with [the Uniform] guidelines." 29 C.F.R. § 1607.3(B). Specifically, Swander testified that he reviewed the CPAT test

components with the Department, but it rejected the pike pole component of the CPAT. The court credits this testimony, and further finds that the Department offered significant evidence in support of the specific designs of its own ladder and pike pole test. In particular, the evidence demonstrated that the ladder test reflected the equipment used by the Department at the time of the test, and the pike pole configuration took into account safety considerations. The court sees no reason to find that these discrete decisions somehow invalidated the PAT.

Plaintiff also faulted the Department for its failure to consider the fairness of the test on female applicants. Specifically, plaintiff pointed out that other than Firefighter Frances Tatar's research into the CPAT and attempts to collect data about hiring of female applicants by other departments using the CPAT, the Department did not engage in a "fairness" assessment as required under the portion of § 1607.14 addressing criterion validity studies. However, defendant argued that a fairness assessment was not required for content validity studies, given the regulation's construction, a reading which the court indicated seemed correct during the closing discussion. (Def.'s Br. (dkt. #70) 10-13; Trial Tr. (dkt. #68) 178-79.) Moreover, in post-trial briefing, defendant again made this argument, which plaintiff did not dispute, effectively conceding that the fairness analysis under 29 C.F.R. § 1607.14(B)(8) was not applicable to content validity studies. (Def.'s

Reply (dkt. #76) 2.)⁴

Based on the record at trial as a whole, therefore, the court finds that defendant has met its burden of proving by a preponderance of the evidence that the 2014 PAT was job-related for the position and consistent with business necessity.

III. Alternative Test, the CPAT, Was Not Adequate.

As a result of this finding, the burden shifts back to plaintiff to prove that the employer refused to adopt an alternative hiring practice resulting in less disparate impact and serving the employer's legitimate needs. 42 U.S.C. §§ 2000e-2(k)(1)(A)(ii), (C); *Ernst*, 837 F.3d at 794; *see also* 29 C.F.R. § 1607.3B. (“[T]he user should use the procedure which has been demonstrated to have the lesser adverse impact.”). Here, Erdman contends that the CPAT is such an alternative test.⁵ In determining whether “the CPAT would be equally as effective as the challenged practice in serving the employer’s legitimate business goals,” the court may consider “[f]actors such as the cost or other burdens of the proposed alternative selection devices.” *Watson*, 487 U.S. at 998.

As an initial matter, defendant challenges whether the evidence demonstrates that the CPAT has a lesser adverse impact. Specifically, defendant questions whether the

⁴ Plaintiff did attempt to argue that *another* regulation dealing with documentation requirements nonetheless requires employers to consider adverse impact on female applicants. (Pl.’s Br. (dkt. #74) 10-14 (citing 20 C.F.R. § 1607.15(C)(5), (6)).) However, the court agrees with defendant that plaintiff’s focus on the documentation regulation and any shortcomings in satisfying this requirement does not undermine a finding that the studies the substantive requirements of a valid study set forth in § 1607.14. (Def.’s Reply. (dkt. #76) 3.)

⁵ In the alternative, plaintiff contends that the Department could modify the PAT to look like the CPAT, but this is a difference with no material distinction.

CPAT's passage rate cited by plaintiff of 48% is a fair comparison to the female passage rate under the 2014 PAT. Specifically, defendant points out that the CPAT study showed of the "naïve" or non-firefighter participants (as compared to the incumbent participants) in the Austin study only two of nine passed, resulting in a passage rate of 22%, which is certainly much closer to the 14% passage rate experienced using the 2014 PAT. (Trial Tr. (dkt. #67) 85-86; Ex. 6 at p.5.) Moreover, the CPAT revalidation study noted that the "comparison of male and female sample size is so disparate, it does not allow statistically meaningful comparison." (Ex. 6 at p.6.) In considering the pike pole event in particular -- the element that eliminated plaintiff Erdman from the 2014 PAT-- 4 out of 20 women or 20% participating in the CPAT validation study did not complete that component, as compared to 1 out of 7 participants or 14% participating in the 2014 PAT.

Still, even conceding the likely lack of a statistically significant testing sample between a single department's use of CPAT and the Madison Fire Department's use of the 2014 PAT, plaintiff's expert Weltman also relied on a broader study of fire departments using the CPAT, in which the pass rates for women was 68.0% as compared to 49.0% in departments using other physical ability tests. (Weltman Rept. (dkt. #18-2) 24.) Based on the results of this broader survey, the court finds it more probably true than not true that the CPAT has less of an adverse impact on female applicants. The court further finds that the CPAT is valid generally, which the defendant does not really challenge. Indeed, defendant's expert Charles Swander testified that, in his role as a co-owner of a testing business, he has administered the CPAT over 10,000 times. (Trial Tr. (dkt. #68) 108.)

This leads to the final issue addressed at trial: whether plaintiff proved by a preponderance of the evidence that the CPAT meets Madison Fire Department's legitimate needs. In answering this question, both parties focus on certain, distinct elements of the PAT compared to the CPAT. At the outset, the court recognizes that this analysis proves an ill fit, since plaintiff is challenging the PAT as a whole, which is necessary to meet her *prima facie* burden of showing a disparate impact. Moreover, if plaintiff were pursuing a claim based on a discrete element of the PAT, the available statistics are not nearly so clear cut to support her claim of disparate impact, as described above. Nevertheless, plaintiff seeks to latch onto discrete elements of the PAT, as compared to the CPAT, to argue that the latter meets the Madison Fire Department's legitimate needs.

Even adopting plaintiff's approach, the court must consider the Department's specific arguments in support of the two elements that plaintiff asserts are unnecessary as compared to CPAT's alternative elements -- the ladder and pike pole. In particular, the Department points out that those elements of the test were specifically designed to replicate the tasks Madison firefighters would be expected to execute in light of the equipment available to the Department at that time and of concerns about safety. Evidence of the time spent by the Department's personnel and their expert consultants in developing these two elements of the 2014 PAT was overwhelming. In crediting this testimony, however, the court is skeptical of the Department's argument that its role as a forerunner in developing this type of physical abilities test -- one that is premised on job tasks, rather than just general fitness requirements -- as well as its relatively strong record of hiring women more generally when compared to other fire departments around the

country, should somehow excuse it from considering an alternative test. On the other hand, plaintiff simply points to the CPAT, assuming that it would fit Madison's needs without attempting to validate the test locally.

Regardless, defendant proffered credible evidence of numerous burdens associated with adopting the CPAT as an alternative test, including: (1) the need to perform a transferability study; (2) the PAT having been a good predictor of outcome historically, as defined by a high passage rate out of the academy; (3) the Department's comparatively high percentage of female firefighters, leading to a possible inference that the CPAT may have a favorable disparate impact on women but results in the washing out of ultimately unsuccessful applicants after the additional expenditure of time and money at the academy phase; and (4) certain elements of the PAT were designed specifically for Madison, in light of characteristics of the city, the Department's equipment or other considerations, including safety. Given plaintiff bears the burden to prove the CPAT would serve the Madison Fire Department's legitimate needs, when coupled with the Seventh Circuit's admonition that "courts are generally less competent than employers to restructure business practices, and unless mandated to do so by Congress they should not attempt it," *Ernst*, 837 F.3d at 794 (quoting *Furnco Constr. Corp. v. Waters*, 438 U.S. 567, 578 (1978)), the court concludes that plaintiff has not demonstrated by a preponderance of the evidence that the CPAT meets the Department's legitimate needs as an alternative to the 2014 PAT.

ORDER

IT IS ORDERED that the court finds in defendant's favor on plaintiff's Title VII claim. The clerk's office is directed to enter judgment accordingly.

Entered this 20th day of July, 2022.

BY THE COURT:

/s/

WILLIAM M. CONLEY
District Judge

IN THE UNITED STATES DISTRICT COURT
FOR THE WESTERN DISTRICT OF WISCONSIN

CATHERINE ERDMAN,

Plaintiff,

Case No. 16-cv-786-wmc

v.

CITY OF MADISON,

Defendant.

JUDGMENT IN A CIVIL CASE

IT IS ORDERED AND ADJUDGED that judgment is entered in favor of defendant City of Madison against plaintiff Catherine Erdman dismissing this case.

s/ R. Swanson, Deputy Clerk
Joel Turner, Clerk of Court

7/20/2022
Date

