

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