I. THE CITY CONCEDES ITS INABILITY TO JUSTIFY PROMOTION OF THOSE WITH THE 108 HIGHEST SCORES

The City's brief presents an unexpected but realistic concession that the speculative views of its expert fail to justify the City's decision to promote to lieutenant those with the 108 highest test scores. This concession requires that the judgment of the district court be reversed and the case remanded with instructions to fashion appropriate promotional relief.

The context for the City's concession is as follows:

- The City of Chicago promoted to lieutenant the 108 sergeants who had received the highest scores on the promotional test. (Short Appendix 60 n.1.)
- The use of test scores to promote those with the 108 highest scores had an undisputed disparate impact on minorities. (Def.Br. 29.)
- The district court rejected our challenge to rank order promotions and concluded that the use of the test scores to promote those with the highest 108 scores was "approved by the EEOC guidelines and industry standards in connection with a content valid test." (Short Appendix 52.)

In our opening brief, we argued at length that the district court had applied an erroneous legal standard in approving the City's use of final test scores to promote those with the 108 highest scores. (Pl.Br. 35-36.) We asserted that, under the circumstances of this case, Title VII requires evidence that "those with a higher test score do better on the job than those with a lower test score." (Pl.Br. 35, citing *Ensley Branch of NAACP v. Seibels,* 616 F.2d 812, 822 (5th Cir. 1980), opinion following remand, 31 F.3d 1548.) We argued that the evidence offered by the City to justify the promotion of those with the 108 highest scores — the testimony of an industrial psychologist who admitted

that he could not measure actual police performance — was unscientific speculation that was inadmissible under *Daubert v. Merrell Dow Pharmaceuticals, Inc.,* 509 U.S. 579 (1993).[1]

In this Court, the City agrees that its use of test scores to promote those with the 108 highest scores requires evidence that a higher test score is related to better job performance: The City states that "we have no quarrel with the cases that plaintiffs cite holding that test results cannot be used in strict rank order when the use of rank ordering cannot itself be validated."[2] (Def.Br. 47

---

1. We respond to the City's halfhearted defense of its expert at 9-11 below.

2. The City cites *Police Officers for Equal Rights v. City of Columbus,* 916 F.2d 1092 (6th Cir. 1990) as an example of a case where the employer successfully validated rank order selection. (Def.Br. 59 n.11.) The City does not, however, rely on the methodology accepted in *Police Officers for Equal Rights, supra,* to justify rank order selection.

   In *Police Officers for Equal Rights, supra,* the Sixth Circuit accepted "reliability calculations" to justify ranking. Id. at 1102-03. In cases following *Police Officers for Equal Rights,* though, the Sixth Circuit appears to have disavowed the use of "reliability calculations" to justify rank order selection.

   In *Brunet v. City of Columbus,* 1 F.3d 390 (6th Cir. 1993), *appeal following remand,* 58 F.3d 251 (6th Cir. 1995), the Court cited *Police Officers for Equal Rights, supra,* for the proposition that a finding of a relationship between test scores and job performance is reviewed "only for clear error," id. at 410, and cited *Williams v. Vukovich,* 720 F.2d 909, 924 (6th Cir. 1983) for the proposition that "Ranking is a valid, job-related selection technique only where the test scores vary directly with job performance." Id. More recently, in *Gonzales v. Galvin,* 151 F.3d 526 (6th Cir. 1998), the Sixth Circuit wrote of the need for evidence of a "correlation between higher test scores and better job performance," id. at 533, and cited *Police Officers for Equal Rights, supra,* only for the proposition that a test must be "sufficiently representative of the job." Id.

n.8.)

The City also agrees that the evidence fails to justify its decision to promote those with the highest 108 scores: Rather than seeking to defend the speculative opinions of its expert that a higher test score implies better job performance, (Pl.Br. 27), the City offers an Orwellian rewriting of history and states that the district court "*rejected* [the City's] evidence that defended strict rank-order use of the examination results." (emphasis supplied) (Def.Br. 47.)

We agree that the district court *should* have "rejected [the City's] evidence that defended strict rank-order use of the examination results." But the district court did not reject this evidence. Instead — at least in the opinion released on June 30, 1998, reproduced in the short appendix to plaintiffs' opening brief — the district judge wholeheartedly accepted the testimony of the City's expert that final scores on the test could be used for rank order selection: "Promotions were made in rank order to fill vacancies in the rank or lieutenant, *a procedure that Dr. Barrett convincingly defended and that is approved by the EEOC Guidelines and industry standards in connection with a content valid test.*" (emphasis supplied) (Short Appendix 52.)

The district court's finding that the expert's testimony was "convincing" and that rank order promotions are "approved by the EEOC Guidelines and industry standards" cannot be squared with the City's outrageous assertion that the district court "ultimately rejected the testimony [of the City's expert supporting rank order promotions]." (Def.Br. 58) To borrow language used by this Court in a different context in *Douglas v. Agricultural Stabilization and Conservation Service,* 33 F.3d 784, 785 (7th Cir. 1994), when a district court errs, the

right response is not to rewrite history, changing the record in Orwellian fashion to present that it had reached some other conclusion. Changing the record is precisely what the City seeks to do in this appeal.

The rule applied by this Court is that "silence about facts does constitute a waiver of the specific factual contentions made by the opposing party in a brief filed earlier." *Blackwell v. Cole Taylor Bank,* 152 F.3d 666, 673 (7th Cir. 1998). An analogous rule is appropriate when, instead of defending application of a legal standard applied by the district court, a party seeks to revise history and deny that the district court made the challenged ruling. The Court should treat the City's dissembling as a confession of error. We show below that the same is true for the City's refusal to respond to our challenge to the scoring system of the test.

II. THE CITY'S REFUSAL TO DEFEND THE SCORING SYSTEM SHOULD BE VIEWED AS A CONFESSION OF ERROR

In our opening brief, we challenged the scoring system for two of the three components of the test.[3] We argued that the record did not contain any

3. Our challenge to the written job knowledge test, which was graded by the number of correct answers. takes issue with the district court's finding that the written test "measures the job knowledge required to perform the major work behaviors of the lieutenant job." (Pl.Br. 20-21.)

   The City defends the written job knowledge test by asserting that "[w]e know of no legal or professional standard that prohibits an employer from expecting candidates for promotion to have, or attain by studying, a certain level of knowledge required for the job." (Def.Br. 38.)

   The City did not, however, use the written test to determine if a sergeant had a "certain level of knowledge." Instead of seeking to determine a passing score that would indicate had enough job knowledge to work as a lieutenant, the City used scores on the written test for rank ordering, without any evidence that sergeants who scored higher on the test will do better as lieutenants

comprehensible explanation for the system used to grade the "oral communication component." (Pl.Br. 33.) We also pointed out that the scoring system on the "objective in-basket" (the number of correct answers on a multiple choice test) had never been tested by the scientific method. (Pl.Br. 29.) The City ignores these challenges and appears to abandon any defense of the scoring systems for either the "oral communication component" or the "objective in-basket."

The City asserts that the "oral communication component" was designed to approximate tasks that are actually performed by lieutenants on the job, (Def.Br. 44), and argues at length that a content valid approach is appropriate to measure "analytical ability and organization and oral communication skills." (Def.Br. 45-47.) Nowhere in the City's defense of the "oral communication component," however, is there any response to our challenge to the scoring system.

The City's defense of the "oral communication component" is to cite *Gillespie v. State of Wisconsin,* 771 F.2d 1035 (7th Cir. 1985) to support its argument that a content validation approach is appropriate to measure "analytical ability and organization and oral communication skills."[4] The City, however,

---

4. The City mistakenly cites *Gillespie v. State of Wisconsin,* 771 F.2d 1035 (7th Cir. 1985) as authority for the proposition that "analytical ability and organization and oral communication skills" may be measured by a content valid test. *Gillespie,* however, does not support the City's position.

   In *Gillespie,* the Court adopted the three part test articulated in *Guardians v. Civil Service Commission,*, 630 F.2d 79, 93 (2d Cir. 1980) to determine whether "a characteristic is too abstract to be measured by a content validated test." 771 F.2d at 1042.

   Thus, the court must evaluate the test for: (1) the degree to which the

ignores the extensive discussion in *Gillespie* of the "two requirements to justify the choice of a cut-off score: The test scores must be reliable,[5] *Guardians [v. Civil Service Commission,]*, 630 F.2d at 101-02, and the employer must have some justifiable reasoning for adopting the cut-off score. Id. at 105; APA Standards at 66-67." 771 F.2d at 1044-45. Neither requirement is met in this case.

The Court in *Gillespie* stated that "[a]n employer may establish a justifiable reason for a cut-off score by, for example, "using a professional estimate of the requisite ability levels, or, at the very least by analyzing the test

---

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

Gillespie, 771 F.2d at 1043.

The test in *Gillespie* was a three question essay test that had been designed "to screen out those persons who did not possess the fundamental skills necessary [for the job]." 771 F.2d at 1043. The Court concluded that the essay test, used as a screening device, was not measuring characteristics that were too abstract to be measured by a content validated test. The Court concluded that "[t]he abilities to communicate in standard written English, to prepare a written job description, and to place the analysis of a recruiting problem in acceptable form are concrete, observable and concrete characteristics. Id. at 1043.

In this case, the "oral communication component" was not used as a screening device and was intended to measure "analytical ability and organization and oral communication skills." The product of the test was not a work sample but, as described by one of the City's subject matter experts, "was an attempt by the individual to follow the instructions that were contained in the exercise." (Klein Testimony, Tr. 869.)

5. The City offers the opinion of its expert, unsupported by statistics, that the examination "should be considered reliable." (Def.Br. 56.) See *infra* at 10.

results to locate a logical 'break-point' in the distribution of scores." [citing] *Guardians,* 630 F.2d at 105." 771 F.2d at 1045. The latter approach — of locating a logical "break point" — cannot be used when, as here, scores are used for rank order selection: There is no "break point" when the employer simply promotes those with the highest scores. Nor can the alternative of a "professional estimate" be employed to justify rank order selection.

In *Gillespie,* a cut-off score was chosen "in order that the DER could interview as many minority candidates as possible while at the same time assuring that the candidates possessed the minimum skills necessary to perform as a Personnel Specialist/Personnel Manager I. In addition to selecting as many qualified minority applicants as possible, the cut-off score was set so that the interviewers would not be overwhelmed by the sheer number of candidates. Therefore, the record demonstrates that the DER's cut-off score was based upon an estimate of the ability levels needed, maximized the number of minority candidates and preserved the integrity of the interview process." 771 F.2d at 1045.

In this case, unlike *Gillespie,* there was no cutoff score — the City simply combined scores on three subtests and promoted those with the 108 highest scores. Moreover, the City remains unable to explain how the oral component was scored — the best that the City can offer is to rely on the testimony of its expert that the oral test was scored "according to an objective checklist." (Def.Br. 56.) But when asked at trial to describe how the "checklist" was applied, the expert admitted that he was unable to provide any explanation because "I haven't been trained as a rater." (Tr. 545.) *Gillespie* makes plain that more evidence is required to justify a scoring system that has a disparate impact.

The City also argues that the "in basket" was "content valid."[6] (Def.Br. 39.) The City recounts in great detail the manner in which the test was created (Def.Br. 39-43) and concludes that the in-basket "was based on, derived from, and linked from the job analysis findings." (Def.Br. 43.) But nowhere in the City's discussion of the in-basket is there any attempt to explain how the "objective scoring system" for the in-basket is correlated to "the requisite ability levels" referred to in *Gillespie.* Instead, the City refers to the opinions of its expert that he believed that the in-basket had been scored appropriately. (Def.Br. 56.)

In our opening brief, we pointed out that this particular type of in-basket had never been experimentally validated (Pl.Br. 30) and that none of the three research papers cited by the City's expert supported his claims. (Pl.Br. 30-32.) The City characterizes the "objective in-basket" as an "innovation," (Def.Br. 56), but does not disagree with the complete absence of any experimental validation for this "innovation." Nor does the City dispute our reading of the research reports cited by its expert. The City's only defense of the scoring system for the in-basket is to assert that the lack of scientific validity for its expert's opinion "goes to the weight, and not to the admissibility" of the expert's testimony. (Def.Br. 56.)

---

6. The City does not disagree with the district court's finding that the in-basket was "not intended to be exact replica of the job." (Short Appendix 49, cited in Pl.Br. 23.)

In the City's view, the opinions of its "expert" — the industrial psychologist it hired to prepare the test — are all that is needed to prove test validity under *Griggs v. Duke Power Co.,* 401 U.S. 424 (1971) and its progeny. This exclusive reliance on expert opinion was rejected by the Supreme Court in *Griggs,* where the Court approvingly quoted from the EEOC's 1966 guidelines that "The fact that a test was prepared by an individual or organization claiming expertise in test preparation does not, without more, justify its use within the meaning of Title VII." Id. at 433. In this case, it was the City's burden to come forward with evidence beyond the opinions of the psychologist who prepared the test that the sergeants with the 108 highest scores are the most qualified to be promoted to lieutenant. The City's silence on this issue is a concession that it failed to meet its burden.

III. THE SPECULATIVE OPINIONS OF THE CITY'S EXPERT WERE INADMISSIBLE UNDER DAUBERT

The City responds with conclusory assertions to our argument that the speculative opinions of its expert should not have been admitted into evidence. Each of these assertions fails critical analysis.

According to the City, its expert relied on "the general scientific literature." (Def.Br. 52.) The City, however, fails to respond to our opening brief, where we pointed out that under cross-examination, the expert admitted that the "scientific literature" consists of a single unpublished study. (Pl.Br. 27.)

The City also relies on the expert's testimony that the job analysis showed that "the more knowledge, and greater administrative skills and oral communication skills, the more effective a candidate is in the job of Chicago police lieutenant." (Def.Br. 52-53.) The City fails to explain how the job analysis showed that a candidate could be "more effective" when, as its expert admitted, he was

unable to measure actual job performance. (Pl.Br. 27.)

The City also states that its expert relied on "the EEOC guidelines and SIOP principles." (Def.Br. 53.) But once again, the City is unable to reconcile this bold claim with the requirement of the EEOC guidelines on which we relied in our opening brief (Pl.Br. 28-29) that the employer demonstrate that differences in scores are "associated with differences in levels of job performance." (Pl.Br. 29.)

As its penultimate reason for accepting the opinion of its expert, the City states that "the overall reliability of the test battery" was .93. (Def.Br. 53.) But reliability is different from validity — a clock that is stopped will reliably show the same time, but does not present a valid measurement of time. (Tr. 216.) Moreover, the expert could not describe with certainty the reliability of the test — all that he could say was "I think our overall reliability of test battery is .93." (Tr. 408.)

The City's final reason for accepting the opinion testimony of its expert is his claim that "several other scholars support the position that content validation strategy can be used for rank order promotions." (Def.Br. 53.) This self-serving testimony does not provide the scientific validity required by *Daubert.* None of these "other scholars" have seen fit to publish their opinions in peer reviewed journals; that particular views may be shared by four psychologists falls short of providing scientific validity.

The City appears to recognize the dangers of reliance on the speculative opinions of its expert, and asserts that the "district court ultimately rejected the testimony about which plaintiffs complain." (Def.Br. 57-58.) But as we pointed

out above (at 3-4), this is an Orwellian rewriting of the district court's opinion, which described the expert's testimony as "convincing" (Short Appendix 52). The district court erred in accepting the opinions of the City's expert that he had developed a test that could be used to promote those with the highest 108 scores.

IV. THE DISTRICT JUDGE ABUSED HIS DISCRETION IN REFUSING TO ORDER ADDITIONAL MERIT PROMOTIONS

The City offers two misleading assertions in its attempt to defend the refusal of the district court to order additional merit promotions.

First, the City states that "we cannot now know how many more merit promotions the Superintendent might have made had he known that federal law did not require him to obey the *McArdle* injunction." (Def.Br. 66.) But before entry of the *McArdle* injunction, the Superintendent had proposed to use merit selection to make 20% of all promotions to lieutenant and had selected 13 sergeants for merit promotions to complement the 54 rank-order promotions. It is obvious that, but for the injunction, the Superintendent would have made an additional 13 merit promotions when he authorized another 54 rank-order promotions.

Second, the City asserts that it would be impossible to identify any additional merit promotees because "no one else has been identified or selected on the basis of merit at that time *or any time thereafter.*" (Def.Br. 62.) This is a false and outrageous claim. The City made 32 merit promotions to lieutenant on October 26, 1998. (Among those who received a merit promotion were plaintiffs Sandra Day and Bernard Porter.) If ordered by the district court, it is obvious that the Superintendent would have made an additional 13 merit promotions.

None of the arguments offered by the City justify the refusal of the district court to require additional merit promotions.

V. CONCLUSION

For the reasons above stated and those previously advanced, the order finding that the test could be used to make rank order promotions and the order refusing to require the City to make additional merit promotions should be reversed and the case should be remanded with instructions to fashion full relief.

Respectfully submitted,

____

KENNETH N. FLAXMAN
CRAIG B. FUTTERMAN
122 South Michigan Avenue
Suite 1850
Chicago, Illinois 60603
(312) 427-3200
attorneys for plaintiffs-appellants

INDEX

I. THE CITY CONCEDES ITS INABILITY TO JUSTIFY PROMOTION OF THOSE WITH THE 108 HIGHEST SCORES 1

II. THE CITY'S REFUSAL TO DEFEND THE SCORING SYSTEM SHOULD BE VIEWED AS A CONFESSION OF ERROR . . . 4

III. THE SPECULATIVE OPINIONS OF THE CITY'S EXPERT WERE INADMISSIBLE UNDER DAUBERT 9

IV. THE DISTRICT JUDGE ABUSED HIS DISCRETION IN REFUSING TO ORDER ADDITIONAL MERIT PROMOTIONS . . . 11

V. CONCLUSION 12