

AGEE, Circuit Judge, concurring:

Although I concur in Judge Gregory's opinion, I write separately to address my concern with the EEOC's disappointing litigation conduct. The Commission's work of serving "the public interest" is jeopardized by the kind of missteps that occurred here. Gen. Tel. Co. of the Nw. v. EEOC, 446 U.S. 318, 326 (1980). And it troubles me that the Commission continues to proffer expert testimony from a witness whose work has been roundly rejected in our sister circuits for similar deficiencies to those we observe here. It is my hope that the agency will reconsider pursuing a course that does not serve it or the public interest well.

I.

As in other cases, the EEOC proffered expert testimony to establish the alleged disparate impact of Freeman's background check policies. Yet the expert testimony here was fatally flawed in multiple respects.

A.

The district court used harsh words to describe the work of the EEOC's "expert," Kevin R. Murphy. The court found that Murphy's reports contained a "plethora" of "analytical fallacies," reflected "cherry-picked" data, produced "a meaningless, skewed statistic," and included a "mind-boggling

number of errors." EEOC v. Freeman, 961 F. Supp. 2d 783, 793-96 (D. Md. 2013). Even when Murphy submitted late-in-the-day amendments, he still relied upon "a skewed database plagued by material fallacies." Id. at 796. The slapdash nature of Murphy's work convinced the district court that the EEOC had only a "theory in search of facts to support it." Id. at 803.

The majority opinion rightly agrees with the district court's view, as Murphy's work simply did not meet the standards for expert testimony that Federal Rule of Evidence 702 provides. But this was not a close question, and three problems merit special recognition.

First, courts often caution experts against drawing broad conclusions from incomplete data. In Lilly v. Harris-Teeter Supermarket, 720 F.2d 326 (4th Cir. 1983), for instance, this Court criticized an expert for using data from only a limited set of relevant locations and years to draw conclusions about a much broader class. See id. at 337 ("The first problem with this data, however, is that its scope -- covering the stores and warehouse for only 1976 and only the stores for 1975 -- is insufficient to prove discrimination from 1974 through 1978."); see also EEOC v. Am. Nat'l Bank, 652 F.2d 1176, 1195 (4th Cir. 1981) (deeming expert evidence unreliable where it drew conclusions about seven-year period from only one of those seven years). The principle espoused in Lilly derives from a common-

sense idea: expert work should not be considered “[w]hen the assumptions made by [the] expert are not based on fact.” Tyger Constr. Co. v. Pensacola Constr. Co., 29 F.3d 137, 144 (4th Cir. 1994).

Yet as the majority notes, Murphy made the very mistake identified in Lilly: he omitted important information from relevant periods and locations. The EEOC challenged credit check policies beginning in late March 2007 and ending in early August 2011; its criminal-background-check claims spanned November 30, 2007 to the present. For reasons unknown, Murphy’s data included barely any information on applicants after mid-October 2008 -- ignoring at least two-and-a-half years of relevant and available data for each claim. By arbitrarily putting aside those years, Murphy ignored 300 credit checks and 1,500 criminal background checks. Indeed, Murphy even ignored applicant data on persons that the EEOC identified as purported victims. Worse still, Murphy ignored relevant criminal background check data from 21 of Freeman’s 39 different locations.

Neither Murphy nor the agency explained these omissions. Although the EEOC speculates that Freeman produced incomplete data, the record says differently. Among other things, Freeman produced applicant logs, datasheets, and background check forms that Murphy could have used to compile relevant information.

Thus, as the majority indicates, the district court's finding that Freeman presented more than sufficient data is far from clearly erroneous. For his part, Murphy insisted that there was no need to look at more of the available information regardless of relevance. Yet he never explained why his model incorporated enough observations to ensure a valid statistical result.¹

Second, courts have consistently excluded expert testimony that "cherry-picks" relevant data. See, e.g., Bricklayers & Trowel Trades Int'l Pension Fund v. Credit Suisse Secs. (USA) L.L.C., 752 F.3d 82, 92 (1st Cir. 2014); Greater New Orleans Fair Hous. Action Ctr. v. U.S. Dep't of Hous. & Urban Dev., 639 F.3d 1078, 1086 (D.C. Cir. 2011); Barber v. United Airlines, Inc., 17

¹ Experts may use appropriate sampling methods to draw conclusions. But determining an appropriate sample size can be a "tricky" question in statistics, Am. Honda Motor Co., Inc. v. Allen, 600 F.3d 813, 818 (7th Cir. 2010) (per curiam), and Murphy never engaged with it. Some evidence suggests that Murphy used a convenience sample -- that is, he used only the information that was readily at hand. See J.A. 797-98 (indicating that Murphy analyzed only data that was entirely complete without the need for supplementation); see also Freeman, 961 F. Supp. 2d at 794 ("Murphy instead relied almost entirely on the two Excel spreadsheets in creating his database"). Although convenience samples are "easy to take," they "may suffer from serious bias." David H. Kaye & David Freeman, Reference Guide on Statistics in Reference Manual on Scientific Evidence 83, 162 (Fed. Judicial Ctr. 2d ed. 2000). Murphy was no stranger to having courts reject his work for improper sampling. See EEOC v. Kaplan Higher Learning Educ. Corp., No. 1:10 CV 2882, 2013 WL 322116, at *11 (N.D. Ohio Jan. 28, 2013) (criticizing Murphy for failing to explain why his selective use of data did not "skew the sample"), aff'd 748 F.3d 749 (6th Cir. 2014).

F. App'x 433, 437 (7th Cir. 2001); Fail-Safe, LLC v. A.O. Smith Corp., 744 F. Supp. 2d 870, 891 (E.D. Wis. 2010); In re Bextra & Celebrex Mktg. Sales Practices & Prod. Liab. Litig., 524 F. Supp. 2d 1166, 1176-77 (N.D. Cal. 2007). "Cherry-picking" data is essentially the converse of omitting it: just as omitting data might distort the result by overlooking unfavorable data, cherry-picking data produces a misleadingly favorable result by looking only to "good" outcomes.

Murphy undeniably "cherry-picked." The very few pieces of post-October-2008 data that Murphy included consisted of 19 applicants. Of those 19, one was a double-counted applicant, one was a "fail" miscoded as a "pass," and the remaining were all "fails" under one or the other (or both) checks. This 100% failure rate among the 19 post-October-2008 applicants wildly varies from the 3.5% failure rate for criminal checks and 9.9% failure rate for credit checks reflected in the rest of the data. See J.A. 326 (noting that "the likelihood of failing either [check] is low"). Thus, not only was Murphy capriciously selective in his use of post-October-2008 data, but the high number of "fails" among his few selections suggests that he fully intended to skew the results. The district court certainly thought so, terming Murphy's work "an egregious example of scientific dishonesty." Freeman, 961 F. Supp. 2d at 792.

Finally, Murphy's analysis contained many obvious errors and mistakes, and these "factual deficiencies" further evidence his "faulty methods and lack of investigation." Brown v. Burlington N. Santa Fe Ry. Co., 765 F.3d 765, 773 (7th Cir. 2014); see also Dart v. Kitchens Bros. Mfg. Co., 253 F. App'x 395, 399 (5th Cir. 2007) (noting that "basic mathematical errors and flaws in methodology" were appropriate reasons to exclude an expert); cf. Overton v. City of Austin, 871 F.2d 529, 539 (5th Cir. 1989) (per curiam) ("[A] trial court should not ignore the imperfections of the data used[.]"). For example, Murphy's initial statistical analysis was filled with basic arithmetic mistakes. Even once those fundamental errors were corrected, problems lingered. Murphy excluded applicants with known race and gender information, inaccurately claiming incomplete information. He miscoded criminal and credit check outcomes, as well as race and gender information. And he double-counted other applicants. As the majority recounts, within a sample of 41 known "victims" in Murphy's database, 29 of those 41 (or more than 70%) had errors or omissions.

In sum, Murphy's work was riddled with fundamental errors, mistakes, and misrepresentations. I certainly agree with the majority's determination that the district court appropriately excluded Murphy's evidence.

B.

These problems would be troubling enough standing alone, but they are even more disquieting in the context of what appears to be a pattern of suspect work from Murphy.

EEOC v. Kaplan Higher Education Corp., 748 F.3d 749 (6th Cir. 2014), provides only the most recent example. There, the EEOC sought to use Murphy's testimony to challenge an employer's use of credit checks, just as it did here.² A panel of the Sixth Circuit, however, unanimously affirmed the district court's decision to exclude Murphy's determinations. Like his work in this case, Murphy's analysis in Kaplan was filled with errors; among other things, he again "overrepresented 'fails' generally" and again drew conclusions from a skewed, unrepresentative sample. Id. at 752, 754. When the defendant in Kaplan noted several such problems, Murphy responded by filing a series of late reports attempting to repair his earlier ones -- much as he did in this case.³ The Sixth Circuit held that, despite Murphy's

² In Kaplan, "the EEOC sued the defendants for using the same type of background check that the EEOC itself uses." 748 F.3d at 750. The EEOC's claim here is largely the same. Still, the irony of that course is not the subject of this appeal, which focuses only upon the actions that the agency undertook in presenting its case.

³ In the present case, Murphy submitted additional reports right up to the day of the summary judgment hearing. As the majority notes, the district court correctly saw these last-minute changes for what they were: "poorly disguised attempts to (Continued)

eleventh-hour effort to patch his mistakes, his methodology “flunked” every test used to assess expert reliability. Id. at 752. After cataloguing a variety of flaws in Murphy’s analysis, the Sixth Circuit concluded that Murphy’s testimony amounted to “a homemade methodology, crafted by a witness with no particular expertise to craft it, administered by persons with no particular expertise to administer it, tested by no one, and accepted only by the witness himself.” Id. at 754. That account describes the EEOC’s expert evidence in this case to a tee.

Murphy’s flawed approach is not just a recent problem. Over a decade ago, in Cooper v. Southern Co., Murphy drew different but no less severe criticism. See 390 F.3d 695 (11th Cir. 2004), overruled in part by Ash v. Tyson Foods, Inc., 564 U.S. 454, 456-57 (2006) (per curiam). The Eleventh Circuit concluded a report from Murphy served only to “recapitulate[] the basic allegations of the plaintiffs in the guise of an expert report.” Id. at 716 n.10. Indeed, his report lacked any

counter [Freeman]’s arguments with new expert analyses.” Freeman, 961 F. Supp. 2d at 797. The EEOC nevertheless insists that the tardy reports were merely supplements. But “[t]o construe Rule 26(e) supplementation to apply whenever a party wants to bolster or submit additional expert opinions would wreak havoc in docket control and amount to unlimited expert opinion preparation.” Campbell v. United States, 470 F. App’x 153, 157 (4th Cir. 2012) (quotation marks and alterations omitted).

"statistical evidence to substantiate [its] broad claims." Id. Thus, the report and the "sweeping conclusions" within it were "of extremely limited use." Id.

Other recent cases provide additional examples of Murphy's lax attitude towards scientific rigor. In Boelk v. AT & T Teleholdings, Inc., No. 12-cv-40-bbc, 2013 WL 3777251 (W.D. Wis. July 19, 2013), for example, Murphy attempted to offer an expert opinion premised on "common sense," "obvious[ness]," and "foreseeab[ility]." Id. at *8. Unsurprisingly, the district court held that such testimony was "not the appropriate subject of expert testimony" and did not create a genuine dispute of material fact at summary judgment. Id. Echoing a familiar theme, the court dubbed Murphy's testimony "too general and speculative to be useful." Id. The Second Circuit too has rejected Murphy's conclusions, holding that Murphy had incorrectly accused another expert of making unfounded assumptions in her report. See M.O.C.H.A. Soc'y, Inc. v. City of Buffalo, 689 F.3d 263, 278-79 (2d Cir. 2012).

II.

Despite Murphy's record of slipshod work, faulty analysis, and statistical sleight of hand, the EEOC continues on appeal to defend his testimony. Conceding that Murphy's report was not an "A+ report," the EEOC nevertheless says that it meets some

indeterminate threshold of reliability. In doing so, however, the Commission advances positions that are not grounded in law. Most troubling is its view that problems in an expert's data are an inappropriate reason to exclude that expert.

Evidence is admissible only if "it rests on a reliable foundation." Daubert v. Merrell Dow Pharms., 509 U.S. 579, 597 (1993). Thus, the trial court must probe the reliability and relevance of expert testimony any time "such testimony's factual basis, data, principles, methods, or their application are sufficiently called into question." Kumho Tire Co. v. Carmichael, 526 U.S. 137, 149 (1999). Federal Rule of Evidence 702 likewise directs courts to verify that expert testimony is "based on sufficient facts or data." See Fed. R. Evid. 702(b). "A court may conclude that there is simply too great an analytical gap between the data and the opinion offered," and accordingly choose to exclude the opinion. Gen. Elec. Co. v. Joiner, 522 U.S. 136, 146 (1997).

The EEOC, however, ignores this threshold analysis by contending that the issue of the reliability of an expert's data is always a question of fact for the jury, except perhaps in some theoretical, rare case. See, e.g., Reply Br. 15 ("[P]urported flaws in Murphy's analyses concerned data . . . and therefore concerned weight/credibility issues for trial, not admissibility."). The agency's contention ignores Daubert's

instruction that the district court must act as a gatekeeper. Moreover, no court has accepted the agency's argument. Rather, courts widely agree that "trial judges may evaluate the data offered to support an expert's bottom-line opinions to determine if that data provides adequate support to mark the expert's testimony as reliable." Milward v. Acuity Specialty Prods. Grp., Inc., 639 F.3d 11, 15 (1st Cir. 2011) (quotation marks and alteration omitted); accord Blunt v. Lower Merion Sch. Dist., 767 F.3d 247, 276 (3d Cir. 2014); In re TMI Litig., 193 F.3d 613, 697 (3d Cir. 1999); United States v. City of Miami, 115 F.3d 870, 873 (11th Cir. 1997). The EEOC's contention was not simply meritless, but unsupported and without a legal foundation.

* * * *

The EEOC wields significant power, some of which stems from the agency's broad discretion to investigate, conciliate, and enforce, and some of which derives from public actions that exert influence outside the courtroom. The Commission's actions can be also expected to have broader consequences than those of an ordinary litigant given the "vast disparity of resources between the government and private litigants." EEOC v. Great Steaks, Inc., 667 F.3d 510, 519 (4th Cir. 2012).

In deciding when to act, the Commission must balance sometimes-competing responsibilities. On the one hand, the agency must serve the employee's interest by preventing an employer from "engaging in any unlawful employment practice" under Title VII. 42 U.S.C. § 2000e-5(a). On the other hand, "the EEOC owes duties to employers as well: a duty reasonably to investigate charges, a duty to conciliate in good faith, and a duty to cease enforcement attempts after learning that an action lacks merit." EEOC v. Argo Distrib., LLC, 555 F.3d 462, 473 (5th Cir. 2009). That the EEOC failed in the exercise of this second duty in the case now before us would be restating the obvious.

The EEOC must be constantly vigilant that it does not abuse the power conferred upon it by Congress, as its "significant resources, authority, and discretion" will affect all "those outside parties they investigate or sue." EEOC v. Propak Logistics, Inc., 746 F.3d 145, 156 (4th Cir. 2014) (Wilkinson, J., concurring). Government "has a more unfettered hand over those it either serves or investigates, and it is thus incumbent upon public officials, high and petty, to maintain some appreciation for the extent of the burden that their actions may impose." Id. The Commission's conduct in this case suggests that its exercise of vigilance has been lacking. It would serve the agency well in the future to reconsider how it might better

discharge the responsibilities delegated to it or face the consequences for failing to do so.