

FILED
United States Court of Appeals
Tenth Circuit

August 7, 2006

Elisabeth A. Shumaker
Clerk of Court

PUBLISH

UNITED STATES COURT OF APPEALS
TENTH CIRCUIT

KARLA CARPENTER; LINDA
WILKERSON; SHERYL LANDON;
SANDY WILCYNKI; SONYA
PHILLIPS; CHARLENE CHAPMAN;
CHERYL LEE PERSINGER; NENA
HOLDER; RUBY RYHERD,
individually & on behalf of all others
similarly situated; MARY DEAN;
FAITH BRIDGEWATER; VERLENE
MAHOLMES, individually,

Plaintiffs - Appellants/Cross-
Appellees,

v.

Nos. 04-3334, 04-3350, 04-3351

THE BOEING COMPANY,

Defendant - Appellee/Cross-
Appellant.

KARLA CARPENTER; LINDA
WILKERSON; SHERYL LANDON;
SANDY WILCYNKI; SONYA
PHILLIPS; CHARLENE CHAPMAN;
CHERYL LEE PERSINGER; NENA
HOLDER; RUBY RYHERD,
individually, and on behalf of all other
persons similarly situated,

Petitioners,

v.

THE BOEING COMPANY,

Respondent.

**APPEALS FROM THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF KANSAS
(D.C. NO. 02-CV-1019-WEB)**

Jeffrey T. Sprung, Hagens Berman Sobol Shapiro LLP, argued for Plaintiffs - Appellants/Cross-Appellees, (Steve W. Berman, Andrew M. Volk, Ivy D. Arai, Hagens Berman Sobol Shapiro LLP, Seattle, Washington, and Mark B. Hutton and Derek S. Casey, Hutton & Hutton, Wichita, Kansas, with him on the brief).

James M. Armstrong, Foulston Siefkin LLP, argued for Defendant - Appellee/Cross-Appellant, (Mary Kathleen Babcock, Trisha A. Thelen, Carolyn L. Matthews, Foulston Siefkin LLP, Wichita, Kansas, and C. Geoffrey Weirich, Paul, Hastings, Janofsky & Walker LLP, Atlanta, Georgia, with him on the brief).

Mary Dean, Faith Bridgewater and Verlene Maholmes, pro se Plaintiffs - Appellants/Cross-Appellees, submitted a brief.

Before **HARTZ, ANDERSON**, and **O'BRIEN**, Circuit Judges.

HARTZ, Circuit Judge.

Plaintiffs appeal from the district court's disposition of the employment-discrimination claims of female employees at the Boeing Company's Wichita, Kansas, facility. They have sought to bring class-action claims alleging several

unlawful employment practices under both disparate-impact and disparate-treatment theories of discrimination. The two subclasses relevant to this appeal are a subclass of hourly female workers (the Hourly Subclass) and a subclass of salaried female workers (the Salaried Subclass). Before us now are both (1) the district court's summary judgment on the Hourly Subclass's disparate-impact claim relating to overtime assignments, certified by the district court as a final judgment under Fed. R. Civ. P. 54(b); and (2) several of the district court's class-certification decisions relating to both the Hourly and Salaried Subclasses, on which we provisionally granted interlocutory appeal under Fed. R. Civ. P. 23(f). Boeing has cross-appealed to challenge the district court's class certification of the Hourly Subclass's disparate-impact claim in the event that we reverse the district court's grant of summary judgment on that claim.

We affirm the district court's summary judgment because Plaintiffs' statistical evidence is not adequately based on data restricted to persons eligible for overtime assignments. This affirmance moots the cross-appeal. Also, we dismiss Plaintiffs' appeal of the district court's class-action decisions because they were not filed within 10 days of the district court's initial decision denying class certification. Finally, we reject the claims of three former class representatives who were stripped of that designation by the district court on the ground that they could not "fairly and adequately protect the interests of the class," Fed. R. Civ. P. 23(a)(4).

I. BACKGROUND

Title VII of the Civil Rights Act of 1964 prohibits, among other things, discrimination on the basis of sex. *See* 42 U.S.C. § 2000e-2(a). Two types of claims are recognized under Title VII: disparate treatment and disparate impact.

“Disparate treatment” . . . is the most easily understood type of discrimination. The employer simply treats some people less favorably than others because of their race, color, religion, sex, or national origin. Proof of discriminatory motive is critical, although it can in some situations be inferred from the mere fact of differences in treatment. . . .

Claims of disparate treatment may be distinguished from claims that stress “disparate impact.” The latter involve employment practices that are facially neutral in their treatment of different groups but that in fact fall more harshly on one group than another and cannot be justified by business necessity. Proof of discriminatory motive . . . is not required under a disparate-impact theory. Either theory may, of course, be applied to a particular set of facts.

Int’l Bhd. of Teamsters v. United States, 431 U.S. 324, 335 n.15 (1977) (citations omitted). In a disparate-impact claim the plaintiff is challenging an employment practice that is “‘fair in form, but discriminatory in operation.’” *Bullington v. United Air Lines, Inc.*, 186 F.3d 1301, 1312 (10th Cir. 1999) (quoting *Griggs v. Duke Power Co.*, 401 U.S. 424, 431 (1971), *overruled on other grounds by Nat’l R.R. Passenger Corp. v. Morgan*, 536 U.S. 101 (2002)). “[A] plaintiff may establish a *prima facie* case of disparate impact discrimination by showing that a specific identifiable employment practice or policy caused a significant disparate impact on a protected group.” *Id.* (internal quotation marks omitted). This

burden, which had been imposed by caselaw, *see, e.g., Ortega v. Safeway Stores, Inc.*, 943 F.2d 1230, 1242 (10th Cir. 1991), was codified by statute in 1991. *See* 42 U.S.C. § 2000e-2(k); Civil Rights Act of 1991, Pub. L. No. 102-166, § 105(a), 105 Stat. 1071, 1074-75. The 1991 statute departed from case law in several respects, but none are relevant here.

Discrimination suits are often filed as putative class actions. Whether a suit can proceed as a class action is governed by Fed. R. Civ. P. 23. Under that rule the district court must determine “at an early practicable time,” Fed. R. Civ. P. 23(c)(1)(A), whether a suit (or a particular claim within a suit) satisfies the prerequisites of numerosity, commonality, typicality, and adequacy of representation, *see id.* 23(a), and falls within one of the categories of actions maintainable as class actions, *see id.* 23(b). We review *de novo* whether the district court applied the correct legal standard in its decision to grant or deny class certification; when the district court has applied the proper standard, the decision will be reversed only for abuse of discretion. *See Shook v. El Paso County*, 386 F.3d 963, 967-68 (10th Cir. 2004). The district court can modify or amend its class-certification determination at any time before final judgment in response to changing circumstances in the case. *See* Fed. R. Civ. P. 23(c)(1)(C).

In 2000, Plaintiffs, among others, filed a putative nation-wide class-action suit in the United States District Court for the Western District of Washington, alleging gender discrimination in a variety of Boeing’s compensation practices.

The district court, however, certified only a class of female employees working at Boeing's Washington facilities. In 2002 non-Washington plaintiffs filed suits in several states, including this suit in the District of Kansas.

Boeing's Wichita facility includes operations of three major business units: Boeing Commercial Airplanes, which is the largest group at the facility and is responsible for commercial production; the Wichita Development and Modification Center, which is responsible for the site's military business; and the Shared Services Group, which provides infrastructure support. According to the complaint, the Wichita facility is Boeing's largest manufacturing business. In December 2001 Boeing had approximately 16,700 employees in Kansas.

This appeal concerns Plaintiffs' Title VII claims alleging gender discrimination in Boeing's compensation and overtime policies. Nine of the Plaintiffs (the Carpenter Plaintiffs) seek to represent themselves and a class of similarly situated current and former female employees at Boeing's Wichita facility. The other three Plaintiffs (the Dean Plaintiffs) are members of the class but represent only themselves on appeal.

The Hourly Subclass's overtime claims were brought under both disparate-impact and disparate-treatment theories. The claims are based on the allegation that the discretion given to supervisors in assigning overtime resulted in women receiving consistently fewer overtime assignments than their male counterparts. In their disparate-treatment claim, they allege further that Boeing's failure to act

upon knowledge of the denial of those assignments constituted intentional discrimination against its female employees. The Salaried Subclass made a disparate-impact claim that Boeing's company-wide practices for setting both starting salaries and raises systematically disadvantaged its female employees and a disparate-treatment claim that the company had failed to take action to correct the discriminatory impact since learning of it in 1995. Not at issue are other claims brought by the two subclasses and the claims of a putative subclass of female salaried engineers.

II. CLASS CERTIFICATION

There have been several class-certification proceedings before the district court in this case. The court's initial certification decision, on April 25, 2003, granted certification under Rule 23(b)(2) to both the Hourly and Salaried Subclasses on their disparate-impact claims. Certification was denied on all disparate-treatment claims. On February 24, 2004, following merits discovery, the court granted Boeing's motion to decertify the disparate-impact claim of the Salaried Subclass, leaving the overtime disparate-impact claim of the Hourly Subclass as the only claim certified for class-action treatment under Rule 23.

Plaintiffs filed a Renewed Motion for Class Certification (First Renewed Motion) on April 2, 2004, seeking recertification of the Salaried Subclass's disparate-impact claims. The court promptly denied the motion. Plaintiffs filed a Second Renewed Motion for Class Certification (Second Renewed Motion) on

August 27, 2004, seeking certification of the disparate-treatment claims of both the Hourly and Salaried Subclasses and again asking for recertification of the Salaried Subclass's disparate-impact claim. The district court denied the motion on September 8, 2004.

Plaintiffs then filed with this court an application to appeal under Rule 23(f) the denial of their Second Renewed Motion as it related to the claims of the Salaried Subclass. Although Plaintiffs sought certification of the disparate-treatment claim of the Hourly Subclass in their Second Renewed Motion, they abandoned that issue on appeal. Despite asserting in the application that the relief sought was "leave to appeal the district court's decision denying Plaintiffs' Second Renewed Motion for Class Certification," Pet. for Permission to Appeal at 11, the application refers only to the claims of the Salaried Subclass and its arguments relate only to the certification determinations made with respect to that subclass. Therefore, we will address only the claims of the Salaried Subclass.

Boeing argues that Plaintiffs' application was untimely under Rule 23(f), and that we therefore lack jurisdiction to consider it. We provisionally granted the application pending briefing and argument on our jurisdiction and the merits of the appeal. Upon further consideration we dismiss the application as untimely and do not reach the merits of the appeal.

A. Fed. R. Civ. P. 23(f)

Rule 23 was amended in 1998 to add subsection (f), which permits interlocutory appeals of district court orders granting or denying class certification. It states:

A court of appeals may in its discretion permit an appeal from an order of a district court granting or denying class action certification under this rule if application is made to it within ten days after entry of the order. An appeal does not stay proceedings in the district court unless the district judge or the court of appeals so orders.

Fed. R. Civ. P. 23(f).

Interlocutory appeals have long been disfavored in the law, and properly so. They disrupt and delay the proceedings below. *See* 19 James Wm. Moore, Moore's Federal Practice § 201.10[1] (3d ed. 2006) ("The purposes of the final judgment rule are to avoid piecemeal litigation, to promote judicial efficiency, and to defer to the decisions of the trial court. Unfettered interlocutory appeals would disrupt both the trial and appellate processes."); 15A Charles Alan Wright, Arthur R. Miller & Edward H. Cooper, Federal Practice and Procedure § 3907, at 269 (2d ed. 1991) ("When courts attempt to explain the policies that underlie the final judgment rule, . . . [they] speak of 'efficiency,' protecting the role of the trial judge, and the need to avoid such evils as interference with the trial court, deciding unnecessary issues, and deliberate delay or harassment."). But sometimes countervailing considerations predominate. The consideration that led to adoption of subsection (f) is that a class-certification determination can force a resolution of the case that is independent of the merits. When class-action status

is denied, the plaintiffs may need to abandon the case, or settle for a pittance, because the cost of continuing will far outweigh any potential recovery in the individual actions remaining. And when class-action status is granted, the defendant may be facing such enormous potential liability that a significant settlement becomes the only prudent course. As the Advisory Committee note puts it:

[S]everal concerns justify expansion of present opportunities to appeal. An order denying certification may confront the plaintiff with a situation in which the only sure path to appellate review is by proceeding to final judgment on the merits of an individual claim that, standing alone, is far smaller than the costs of litigation. An order granting certification, on the other hand, may force a defendant to settle rather than incur the costs of defending a class action and run the risk of potentially ruinous liability. These concerns can be met at low cost by establishing in the court of appeals a discretionary power to grant interlocutory review in cases that show appeal-worthy certification issues.

Fed. R. Civ. P. 23 advisory committee's note, 1998 Amendments, Subdivision (f).

But this opportunity for an interlocutory appeal is tightly confined. First, “[t]he court of appeals is given unfettered discretion whether to permit the appeal, akin to the discretion exercised by the Supreme Court in acting on a petition for certiorari.” *Id.* And second, there is a short time limit—10 days—within which the aggrieved party can ask the court of appeals to exercise its discretion. *See id.* (“The 10-day period for seeking permission to appeal is designed to reduce the risk that attempted appeals will disrupt continuing proceedings.”). Because this

timeliness requirement is mandatory¹, we must first determine whether Plaintiffs satisfied it.

B. Timeliness

The district court denied Plaintiffs' Second Renewed Motion by order entered on September 8, 2004. Plaintiffs assert that their application filed on September 22, 2004, was timely because it was filed within 10 days of the district court's disposition. *See* Fed. R. Civ. P. 6(a) (computation of time); *Beck v. Boeing Co.*, 320 F.3d 1021, 1022-23 (9th Cir. 2003) (Fed. R. Civ. P. 6(a) governs the timeliness of applications under Rule 23(f)). The validity of that assertion depends on whether the district court's denial was "an order . . . granting or denying class action certification." Fed. R. Civ. P. 23(f). Boeing contends that the district court's order was simply a refusal to reconsider its prior rulings

¹ We are among several circuits that have treated the timeliness requirement as jurisdictional. *See, e.g., Delta Airlines v. Butler*, 383 F.3d 1143, 1144 (10th Cir. 2004) (per curiam) ("Because the petition was not filed within the mandated time period, we dismissed for lack of jurisdiction."); *McNamara v. Felderhof*, 410 F.3d 277, 280 (5th Cir. 2005) ("Unless some exception applies, we lack appellate jurisdiction to entertain the [untimely] petition."). The Supreme Court's recent decision in *Eberhart v. United States*, 126 S.Ct. 403, 406 (2005), however, casts doubt on the notion that the timeliness of notices of appeal generally is jurisdictional, *see In re Special Grand Jury 89-2*, 450 F.3d 1159, 1166 n.2 (10th Cir. 2006), and could have similar implications for Rule 23(f), *see Coco v. Incorporated Village of Belle Terre*, 448 F.3d 490, 491 (2d Cir. 2006) (per curiam). Even if it is not jurisdictional, however, it is unquestionably "mandatory" if properly raised by the opposing party, as was the case here. Because we must dismiss the appeal in either event, we need not analyze *Eberhart's* impact on Rule 23(f).

denying certification to the Salaried Subclass and not itself an order appealable under Rule 23(f). We agree with Boeing.

In a sense, an order denying a motion to reconsider a decision on class certification is an “order . . . granting or denying class action certification.” But that cannot be the sense in which the term is used in Rule 23(f), because that construction of the term would undermine the 10-day time limit for filing an application for review. One who failed to file an application in time could simply file a motion to reconsider; and when that is denied, the 10-day period would restart. *See Gary v. Sheahan*, 188 F.3d 891, 893 (7th Cir. 1999) (“Accepting an appeal from such a decision [leaving the class definition in place] would abandon the time limit for all practical purposes. That step would be both unauthorized and imprudent.”); *cf. McNamara*, 410 F.3d at 281 (“[T]o hold that—no matter how styled—a motion under Rule 23(c) [to alter or amend a class-certification decision] is *always* distinct from a motion to reconsider would allow a party to subvert the ten-day time limit prescribed in Rule 23(f).”).

One might argue, as Plaintiffs do, that this reasoning does not apply when the motion for reconsideration raises new arguments, based on new developments in the case. But the need to avoid causing delay and disruption to the district court proceedings cautions against an appellate court’s engaging in detailed inspection and analysis of the record to determine how new an argument is and whether the underlying evidence was reasonably available when certification was

originally litigated. Moreover, there can be little doubt that review of an order denying a motion for reconsideration would have to be limited to the new elements in the motion—the original order regarding certification must be presumed correct, or there would be a clear end run around the 10-day limit. Yet given the multifactor analysis that courts must apply in deciding the propriety of class certification, such a limited review would often require contorted thinking that exceeds the capacities of even appellate courts. How can an appellate court say that one particular new factor would require a different result regardless of how the district court weighed the factors presented originally? In stating that the new factor required a different result, the appellate court must engage in weighing the factors weighed by the district court in its original ruling but cannot know precisely how much weight the district court granted to each. In particular, what if the district court clearly erred in giving dispositive weight to one factor? How is the appellate court to ignore such error (in keeping with the presumption that the original decision was correct) even when it addresses a motion for reconsideration that raises only a rather inconsequential new factor? To be sure, we do review motions to reconsider in certain circumstances, such as denials of motions under Fed. R. Civ. P. 60(b). But the predicate for that review—for example, fraud or newly discovered evidence—is largely collateral to the merits of the decision. We are not inclined to adopt a construction of Rule 23(f) that would regularly require mental gymnastics just for the purpose of giving litigants

a second bite at the interlocutory-appellate-review apple. We note that the very absence of a prompt appeal by the party aggrieved by the decision on certification suggests that the concerns justifying Rule 23(f) are, at the least, less significant in the particular case. If the decision whether or not to certify the class was truly outcome determinative, one would not expect the losing party to continue the litigation for months before launching a new challenge to the ruling. Any value in permitting a belated interlocutory appeal is overridden by the desirability of the district court's proceeding expeditiously.

We recognize that Rule 23(c)(1)(C) permits the district court to alter or amend a certification decision. And parties may suggest such changes as the factual record and legal theories develop. All we are saying is that there can be no Rule 23(f) appeal from the denial of such a suggestion. An order that leaves class-action status unchanged from what was determined by a prior order is not an order "granting or denying class action certification." Of course, when the district court accepts a suggestion and the certification decision is changed, the new order, to the extent it modifies the prior order, is indeed such an order and an interlocutory appeal under Rule 23(f) is permitted. *See Gary*, 188 F.3d at 893 ("[I]f in response to a belated motion for reconsideration the judge materially alters the decision, then the party aggrieved by the alteration may appeal within the normal time.").

In addition, we note the special case of motions to reconsider filed within 10 days of the district court's certification decision. The Supreme Court has long recognized that motions to reconsider toll the time for appeal when they are filed within the time for filing a notice of appeal. *See United States v. Dieter*, 429 U.S. 6, 8 & n.3 (1976) (“[T]he consistent practice in civil and criminal cases alike has been to treat timely petitions for rehearing as rendering the original judgment nonfinal for purposes of appeal for as long as the petition is pending.”). This recognition stems from the clear advantage of providing the district court an opportunity to correct its own error, as long as doing so does not undermine the time limit for pursuing an appeal. We assume, without having to decide in this case, that such motions to reconsider would also toll the time limit in Rule 23(f). *See, e.g., McNamara*, 410 F.3d at 281 (recognizing the tolling effect of a timely motion to reconsider in the Rule 23(f) context).

The district court first denied certification of the Salaried Subclass's disparate-treatment claim on April 25, 2003. Plaintiffs did not exercise their right to file a Rule 23(f) petition within 10 days of that order. The district court's determination as to that claim has not changed in all the subsequent proceedings, so there has not been any other order “granting or denying class action certification” that would trigger another period for seeking interlocutory appeal. As for class certification of the Salaried Subclass's disparate-impact claim, the district court initially certified the class in its April 25, 2003, order, which Boeing

did not appeal. On February 24, 2004, however, it granted Boeing's motion to decertify. At that juncture Plaintiffs could have sought our review under Rule 23(f), but they did not do so. Despite two attempts to have the district court recertify that claim, in both Plaintiffs' First Renewed Motion and their Second Renewed Motion, the district court's ruling on certification has not changed again. Therefore, neither of the orders denying those motions was an order granting or denying certification and neither triggered a new period for filing a Rule 23(f) application. Plaintiffs' Rule 23(f) application on September 22, 2004, must be dismissed as an untimely attempt to have us review the court's orders of April 25, 2003, and February 24, 2004. Any appeal of those certification decisions must await final judgment. *See Gary*, 188 F.3d at 892.

III. SUMMARY JUDGMENT

On February 24, 2004, the district court granted Boeing's motion for summary judgment on the disparate-impact claim of the Hourly Subclass. On August 11, 2004, the court certified that ruling as a final judgment, *see* Fed. R. Civ. P. 54(b), and Plaintiffs filed a timely appeal. Boeing filed a cross-appeal seeking decertification of this subclass should we reverse the summary-judgment ruling. Because we affirm the district court's grant of summary judgment, we need not address the cross-appeal.

A. Standard of Review

Our standard of review on summary judgment is de novo; we apply the same legal standard to be used by the district court. *Garrison v. Gambro, Inc.*, 428 F.3d 933, 935 (10th Cir. 2005). Summary judgment should be granted if “the pleadings, depositions, answers to interrogatories, and admissions on file, together with the affidavits, if any, show that there is no genuine issue as to any material fact and that the moving party is entitled to a judgment as a matter of law.” Fed. R. Civ. P. 56(c). Neither “mere assertions and conjecture,” *York v. AT&T Co.*, 95 F.3d 948, 955 (10th Cir. 1996), nor “the existence of a scintilla of evidence in support of the nonmovant’s position,” *Lawmaster v. Ward*, 125 F.3d 1341, 1347 (10th Cir. 1997), is sufficient to show a genuine issue of material fact; “an issue of material fact is genuine only if the nonmovant presents facts such that a reasonable jury could find in favor of the nonmovant,” *id.*

B. Plaintiffs’ Claim

Plaintiffs complain that women have been the victims of discrimination in the assignment of overtime at Boeing’s Wichita facility, being offered and receiving less than their proportionate share. They have raised both disparate-treatment and disparate-impact claims relating to overtime. The summary judgment disposed of only their disparate-impact claim.

“An unlawful employment practice based on disparate impact is established . . . only if . . . a complaining party demonstrates that a respondent uses a particular employment practice that causes a disparate impact on the basis of . . .

sex” 42 U.S.C. § 2000e-2(k)(1)(A)(i). The first step in raising a disparate-impact claim is to identify the specific employment practice allegedly causing the discriminatory impact. *See Wards Cove Packing Co. v. Atonio*, 490 U.S. 642, 657 (1989) (“[A] plaintiff must demonstrate that it is the application of a specific or particular employment practice that has created the disparate impact under attack.”); *Maldonado v. City of Altus*, 433 F.3d 1294, 1304 (10th Cir. 2006). The specific practice identified by Plaintiffs is that “Boeing supplies no guidance to managers on how to choose among eligible employees, and there are no centralized rules for how to choose among equally eligible male and female employees.” R. Doc. 340 at 23 (Plaintiffs’ Mem. in Opp’n to Boeing’s Mot. for Summ. J.); *see also* Aplt. Br. at 26. (In the summary-judgment proceedings below, and on appeal, Plaintiffs have also claimed discriminatory impact from Boeing’s failure to monitor managers or hold them accountable for the gender impacts of their overtime decisions. That allegation was not addressed by the district court. Plaintiffs’ appellate briefs, however, contain no further elaboration of this claim, so we will not specifically address it. *See Gross v. Burggraf Constr. Co.*, 53 F.3d 1531, 1546-47 (10th Cir. 1995) (“[I]t is insufficient merely to state in one’s brief that one is appealing an adverse ruling below without advancing reasoned argument as to the grounds for appeal.” (internal brackets and quotation marks omitted)). In any event, our discussion of their claim that

supervisors were given inadequate guidance in overtime assignments would likely also be dispositive of this claim.)

“Under the disparate impact theory, a plaintiff must first make out a prima facie case of discrimination by showing that a specific identifiable employment practice or policy caused a significant disparate impact on a protected group.” *Murphy v. Derwinski*, 990 F.2d 540, 544 (10th Cir. 1993) (internal quotation marks omitted). In other words, a plaintiff must “show that there is a legally significant disparity between (a) the [gender] composition, caused by the challenged employment practice, of the pool of those enjoying a job or job benefit; and (b) the [gender] composition of the qualified applicant pool . . . [, i.e.,] the pool from which potential qualified applicants might come.” *Crum v. Alabama (In re Employment Discrimination Litig. Against Ala.)*, 198 F.3d 1305, 1312 & n.11 (11th Cir. 1999). The court compares the gender composition of those who are subject to the challenged employment practice with the gender composition of those enjoying the benefit for which the practice selects. In assessing whether a plaintiff has established a prima facie case, it is, of course, irrelevant what happens to those who do not qualify for consideration. *See Wards Cove*, 490 U.S. at 650-51 (“The proper comparison is between the racial composition of the at-issue jobs and the racial composition of the *qualified* population in the relevant labor market.” (emphasis added; internal quotation marks, brackets, and ellipsis omitted)).

Plaintiffs' claim rests on the assertion that supervisors are exercising their discretion (intentionally or subconsciously) to award males a disproportionate share of available overtime assignments. To establish a prima facie case, it is not enough for Plaintiffs to show simply that more overtime assignments go to men than to women, or even that men get a higher percentage of those assignments than their percentage in the work force. They must compare *qualified* men to *qualified* women. That is, they must show that among men and women who are eligible for overtime assignments, a disproportionate share of overtime assignments go to men.

The qualifications for overtime assignment are established in the collective bargaining agreement (CBA) between Boeing and the International Association of Machinists and Aerospace Workers AFL-CIO. The CBA provides:

6.10(b) Overtime Scheduling Procedures for Extended Workday or Workweek.

- (1) The normal practice for the advance scheduling of overtime within the shop and shift will be to:
 - (a) First, ask the employee regularly assigned to either the machine, job, crew, or position providing the employee is in attendance when the overtime is being assigned
 - (b) Then, ask other qualified employees in the same job classification who are in attendance when the overtime is being assigned.
 - (c) If sufficient volunteers are not obtained, the Company may designate any employee to satisfy remaining requirements.
- (2) Management may exclude an employee from overtime, even if the employee is in attendance when the overtime is being assigned, if:
 - (a) The employee has been absent during the week

- (b) An employee is asked to work overtime (Saturday and/or Sunday) and is subsequently absent due to illness or bereavement leave on the workday preceding the overtime day.
- (c) Two (2) consecutive weekends have been worked by the employee.
- (d) One hundred forty-four (144) overtime hours have been worked in the budget quarter.
- (e) Eight (8) overtime hours have been worked on the Saturday or the Sunday.
- (f) An employee's schedule performance or work quality is currently documented as being deficient.

Rep. Aplt. Supp. App. Vol. 1 at 10-11. (This provision is from the September 2, 1999, CBA. The 1995 CBA may be applicable to a portion of the class period, which runs from April 2, 1999. Plaintiffs acknowledge, however, that the overtime provisions did not change materially during the class period.) Plaintiffs do not dispute that the CBA applies to the challenged overtime assignments. Their claim is that the discretion exercised by managers in “choos[ing] among employees who worked in the area where overtime was required and wanted the extra pay” has created a disparity between similarly situated men and women. Aplt. Br. at 26. Boeing for its part does not dispute that supervisors have some measure of discretion within the terms of the CBA. It argues, however, that Plaintiffs have failed to make a sufficient showing that this discretion has been exercised in a manner adverse to women.

On appeal Plaintiffs contend that their statistical evidence of disparate impact suffices to preclude summary judgment. (In district court Plaintiffs also presented a variety of anecdotal evidence to support this claim. But because on

appeal they do not rely on that evidence in challenging summary judgment, we will consider only the statistical evidence.) They rely on a study by their expert, Dr. Bernard Siskin. Dr. Siskin performed a regression analysis that compared the overtime worked by male and female employees whom he defined as “similarly situated.” R. Doc. 346 (Decl. of Bernard P. Siskin, Ph.D. in Opp’n to Boeing’s Mot. for Summ. J. (hereinafter “Siskin Study”)) at 22. The Siskin Study examined overtime assignments from April 2, 1999 (the beginning of the liability period for this claim) through June 20, 2002, using Boeing’s electronic daily payroll records. For weekday overtime the Siskin Study defined similarly situated employees as those who “[w]orked that day and are in the same job, grade, budget code and shift.” *Id.* at 23. Similarly situated employees with respect to weekend overtime were defined as those who “[w]orked Friday and are in the same job, grade, budget code and shift.” *Id.* For each cohort of similarly situated employees, the Siskin Study calculated three measures for men and women: (1) the likelihood of working any overtime; (2) the average number of overtime hours worked; and (3) the average number of overtime hours paid (overtime is paid at either 1.5 or 2 times a normal hour). It then computed a shortfall number for females that described how much greater each measure would be were females represented in proportion to their percentage representation in each cohort. “That is, if females were 25 percent of the cohort,

they should be 25 percent of those working overtime and receive 25 percent of the overtime hours and pay.” *Id.* at 22.

The Siskin Study concluded that “[h]ourly female employees who are similarly situated to males with respect to job, grade, shift, department, and budget code are consistently and highly statistically significantly less likely to work overtime, to work less overtime, and to receive less overtime pay. This pattern is consistent across time.” *Id.* at 3. It observed that “[c]learly, something in the overtime process consistently results in males obtaining more overtime and working more overtime than females.” *Id.* at 23.

There is no dispute that “something” causes men to work proportionately more overtime than women at Boeing. The district court said that the following summary was uncontroverted for purposes of summary judgment:

Between April 2, 1999, and December 31, 2001, disparities in overtime adverse to women ranged between a low of 17.06 standard deviations and a high of 38.03 standard deviations. For the last period for which Dr. Siskin has analyzed data, ending June 20, 2002, the disparities were 10.23 standard deviations for weekday overtime and 7.95 standard deviations for weekend overtime.

Rep. Apls. App. Vol.1 at 280.

Boeing concedes that these differences are statistically highly significant. The Supreme Court has recognized that a disparity of more than two or three standard deviations in a large sample makes “suspect” the contention that the differential occurs randomly. *See Hazelwood Sch. Dist. v. United States*, 433

U.S. 299, 308 n.14 (1977); *Castaneda v. Partida*, 430 U.S. 482, 496 n.17 (1977).

Several circuit courts have adopted a similar level of significance in Title VII cases. *See, e.g., Smith v. Xerox Corp.*, 196 F.3d 358, 366 (2d Cir. 1999); *Brown v. Philip Morris Inc.*, 250 F.3d 789, 809 (3d Cir. 2001); *Lewis v. Bloomsburg Mills, Inc.*, 773 F.2d 561, 568-69 (4th Cir. 1985) (five to eight standard deviations); *Adams v. Ameritech Servs., Inc.*, 231 F.3d 414, 424 (7th Cir. 2000).

But despite recognizing that the statistics show that men have worked proportionately more overtime than women, Boeing claims that the Siskin Study nonetheless fails to establish a prima facie case. It contends that the Siskin Study does not show that the “something” causing men to work more overtime than women is the manager discretion that Plaintiffs have identified as the challenged employment practice. Boeing’s argument appears to be that the “something” is a variable other than those that the Siskin Study included in the statistical model—namely, job, grade, budget code, and shift. According to Boeing, other variables affecting overtime assignments—such as the CBA criteria and potential differences in the rates at which men and women volunteer for overtime—are not controlled for in the Siskin Study and could be responsible for the observed disparities. The district court agreed with Boeing that a statistical study could not establish a claim without considering such variables and granted Boeing’s motion for summary judgment on that basis. Before addressing Boeing’s arguments, with

which we agree in part, we review the legal framework for the use of statistical evidence in Title VII cases.

C. Statistical Evidence - General Principles

Statistical evidence is an acceptable, and common, means of proving disparate impact. *See, e.g., Sandoval v. City of Boulder*, 388 F.3d 1312, 1326 (10th Cir. 2004); *Bullington*, 186 F.3d at 1312 (“As is typical in disparate impact cases, [plaintiff] relies on statistical evidence to establish her *prima facie* case.”); *Mountain Side Mobile Estates P’ship v. Sec’y of HUD*, 56 F.3d 1243, 1251 (10th Cir. 1995) (“In Title VII employment discrimination cases, plaintiffs may rely solely on a statistical showing of disparate effect to establish a *prima facie* case of disparate impact.”). The statistics must, however, relate to the proper population. For example, when the claim is disparate impact in hiring, the statistics should be based on data with respect to persons qualified for the job. *See Wards Cove*, 490 U.S. at 650-51 (“It is such a comparison—between the racial composition of the qualified persons in the labor market and the persons holding at-issue jobs—that generally forms the proper basis for the initial inquiry in a disparate-impact case.”); *see also Bullington*, 186 F.3d at 1314 (“[Plaintiff’s] applicant pool was appropriately limited to persons who sought out and were at least minimally qualified for the position”). The same requirement applies to other job benefits. *See Crum*, 198 F.3d at 1309, 1312 (relating to alleged discrimination in “layoffs, recalls from layoffs, terminations, discipline, hiring, rehiring,

evaluations, compensation, transfers, job duty assignments, recruitment, screening, selection procedures, denial of promotions, demotions, rollbacks, sick leave, subjective decision-making practices, and other terms and conditions of employment” (internal quotation marks omitted)). The essential requirement is that the data concern those persons subject to the challenged employment practice.

After specifying the employment practice allegedly responsible for excluding members of their protected class from a benefit, plaintiffs must identify the correct population for analysis. In the typical disparate impact case the proper population for analysis is the applicant pool or the eligible labor pool. The composition of this population is compared to the composition of the employer's workforce in a relevant manner, depending on the nature of the benefit sought.

Smith, 196 F.3d at 368. When the selection process is only partially subjective, a disparate-impact plaintiff should control for the constraints placed upon the decisionmaker's discretion. *See Anderson v. Westinghouse Savannah River Co.*, 406 F.3d 248, 266-67 (4th Cir. 2005); *cf. Watson v. Fort Worth Bank & Trust*, 487 U.S. 977, 994 (1988) (O'Connor, J., plurality opinion) (“Especially in cases where an employer combines subjective criteria with the use of more rigid standardized rules or tests, the plaintiff is in our view responsible for isolating and identifying the specific employment practices that are allegedly responsible for any observed statistical disparities.”).

To be sure, the population selected for statistical analysis need not perfectly match the pool of qualified persons. Such perfection may be impossible to obtain. When reliable data regarding that pool are unavailable, a different population may be used if it adequately reflects the population of qualified persons. *See* Ramona L. Paetzold & Steven L. Willborn, *The Statistics of Discrimination* § 5.04 (2002) (“In some instances, where applicant data are not available, reliable, or are believed to be biased, and where statistical information regarding the labor market is difficult to ascertain, the general population might adequately reflect the population of qualified job applicants.”); *see also Malave v. Potter*, 320 F.3d 321, 326-27 (2d Cir. 2003) (“[I]t was error [to reject] out of hand [Plaintiff’s] statistical analysis simply because it failed to conform to the *preferred* methodology described in *Wards Cove*, given the Supreme Court’s express endorsement in that decision of alternative methodologies if the preferred statistics are ‘difficult’ or ‘impossible’ to obtain.”); *cf. Trout v. Lehman*, 702 F.2d 1094, 1102 (D.C. Cir. 1983) (in disparate-treatment case brought before the Civil Rights Act of 1991, “plaintiffs cannot legitimately be faulted for gaps in their statistical analysis when the information necessary to close those gaps was possessed only by defendants and was not furnished either to plaintiffs or to the Court” (internal quotation marks omitted)), *vacated on other grounds by Lehman v. Trout*, 465 U.S. 1056 (1984), and *abrogated on other grounds by Berger v. Iron Workers Reinforced Rodmen, Local 201*, 170 F.3d 1111, 1124-25 (D.C. Cir.

1999). For example, in *Dothard v. Rawlinson*, 433 U.S. 321 (1977), the Supreme Court determined that plaintiffs who were challenging Alabama’s height and weight requirements for prison guards could use height and weight statistics based on national data for comparison. “[R]eliance on general population demographic data was not misplaced where there was no reason to suppose that physical height and weight characteristics of Alabama men and women differ markedly from those of the national population.” *Id.* at 330.

Nevertheless, absent a close fit between the population used to measure disparate impact and the population of those qualified for a benefit, the statistical results cannot be persuasive. “[S]tatistics based on an applicant pool containing individuals lacking minimal qualifications for the job would be of little probative value.” *Watson*, 487 U.S. at 997.

Thus, a statistical analysis cannot establish a plaintiff’s prima facie case unless it is based on data restricted to qualified employees, or (1) reliable data with respect to that group are unavailable and (2) the plaintiff establishes that the statistical analysis uses a reliable proxy for qualification. This approach holds plaintiffs to their statutory burden to “demonstrate[] that a respondent uses a particular employment practice that causes a disparate impact on the basis of . . .

sex,” 42 U.S.C. § 2000e-2(k)(1)(A)(i), without imposing an insurmountable burden when reliable data on a qualification are not available.²

D. Application to this Case

The employment practice challenged by Plaintiffs is the exercise of discretion by supervisors in assigning overtime. As stated in *Smith*, “[T]he proper

² The statute provides another avenue for plaintiffs when the particular aspect of the process that is claimed to be objectionable cannot be isolated:

[T]he complaining party shall demonstrate that each particular challenged employment practice causes a disparate impact, except that if the complaining party can demonstrate to the court that the elements of a respondent’s decisionmaking process are not capable of separation for analysis, the decisionmaking process may be analyzed as one employment practice.

42 U.S.C. § 2000e-2(k)(1)(B)(i). Plaintiffs argue in their reply brief to this court that they “presented facts below that defendant’s collective bargaining agreements imposed no meaningful objective standards on supervisors in assigning overtime” and that “[s]uch a showing is sufficient to trigger subsection 2000-e2(k)(1)(B)(i).” Rep. Aplt’s Reply Br. at 10. But they provide no citation to the record showing that they raised this issue in district court, and we cannot find in the record anything indicating to that court that they were attempting to make the required showing of analytical inseparability. We will not address the potential application of § 2000e-2(k)(1)(B)(i) to Plaintiffs’ claim, because our general rule is not to address arguments that were not first presented to the district court, *see Cummings v. Norton*, 393 F.3d 1186, 1190 (10th Cir. 2005) (the “general rule that issues not raised below are waived on appeal” is particularly important on appeal of summary judgment); *Bancamerica Commercial Corp. v. Mosher Steel of Kansas, Inc.*, 100 F.3d 792, 798-99 (10th Cir. 1996) (“Where a litigant changes to a new theory on appeal that falls under the same general category as an argument presented at trial or presents a theory that was discussed in a vague and ambiguous way the theory will not be considered on appeal.” (brackets and internal quotation marks omitted)), and we particularly frown on the making of new arguments in a party’s reply brief, *see Stump v. Gates*, 211 F.3d 527, 533 (10th Cir. 2000).

population for analysis is the . . . eligible labor pool.” 196 F.3d at 368. The Boeing hourly employees eligible for an overtime assignment are those who satisfy the CBA requirements for the assignment; that is, the challenged practice operates only with respect to employees eligible under the CBA. The CBA requires that overtime first be offered to “the employee regularly assigned to either the machine, job, crew, or position” for which overtime is to be scheduled, and then to others within the same “shop or shift.” Rep. Aplt. Supp. App. Vol. 1 at 10. The Siskin Study, however, did not incorporate the CBA’s eligibility requirements in its analysis. Instead, it controlled for “job,” “grade,” “budget code,” and “shift.” R. Doc. 346 at 23. The implicit assumption is that two hourly workers with the same job, grade, budget code, and shift have equal opportunities for overtime assignments under the CBA, subject to the supervisor’s discretion. There is certainly overlap between the Siskin Study variables and those used to determine overtime assignments. But the two sets of variables are not the same. Among the qualifications included in the CBA, the Siskin Study controlled for only “job” and “shift.” It did not account for whether women worked on the “machine” or were in the “crew,” “position,” or “shop” to which the overtime was assigned. This failure can skew the results. If, for example, overtime assignments were concentrated in a handful of shops and almost no women worked in those shops, a discrepancy found in the Siskin Study between the overtime worked by men and by women would not at all represent a disparate

impact created by the supervisor's discretionary choice among eligible employees. Rather, it could simply be a reflection of the gender distribution among those eligible for overtime. At the outset, therefore, it appears that the Siskin Study cannot establish a prima facie case based on a comparison "between (a) the [gender] composition, caused by the challenged employment practice, of the pool of those enjoying a . . . job benefit; and (b) the [gender] composition of the qualified applicant pool," *Crum*, 198 F.3d at 1312, because the study is not limited to data regarding those qualified people subject to the challenged practice. The study does not isolate the effect of supervisor discretion from the effect of the CBA requirements. *See Anderson*, 406 F.3d at 260.

Accordingly, we look to whether Plaintiffs have adequately established that (1) reliable data on the omitted CBA criteria were unavailable and (2) they used a reliable proxy. They have not. To begin with, Plaintiffs have not established that the data necessary to establish the impact on CBA-qualified workers were unavailable. Plaintiffs acknowledge that variables such as crew, position, and shop are relevant to qualification for overtime under the CBA but claim that they cannot be held responsible for including them in their statistical analysis because "Boeing did not maintain electronic data on any of the omitted variables." Rep. Aplt's. Br. at 35. But data may be available in nonelectronic form. Electronic data are undeniably more convenient, especially for use in statistical studies, but inconvenience does not excuse failure to collect the data. Plaintiffs have

presented no reason why the omitted information could not have been procured through other methods, such as depositions or interrogatories. It appears that they were simply satisfied with Boeing's indication that the data were unavailable in their electronic payroll records.

Furthermore, even were we convinced that the data are unavailable, Plaintiffs have failed to demonstrate that the variables in the Siskin Study's statistical analysis produce a reliable surrogate for qualifications for overtime; that is, that the results accurately reflect comparisons between individuals who were equally eligible for overtime assignments under the CBA. Plaintiffs make the bald claim that the "grade" and "budget code" variables used by the Siskin Study are equivalent to the omitted variables "crew," "position," and "shop." *See* Rep. Aplt's. Reply Br. at 5 ("[I]n the absence of specific electronic data maintained by Boeing identifying employees' 'shops' or 'crews,' [Dr. Siskin] closely tracked this information by using as a proxy budget codes and grade levels that reflected their area and level of work."). But they make no attempt to explain the basis of this claim. We cannot agree that those relationships are as self-evident as Plaintiffs apparently believe them to be. The record does not even indicate what a "budget code" is. Plaintiffs' "mere assertion" will not suffice. *See York*, 95 F.3d at 955. Accordingly, we agree with the district court that the Siskin Study was insufficient to establish a prima facie disparate-impact case.

Plaintiffs rely on *Bazemore v. Friday*, 478 U.S. 385, 400 (1986), and *Bullington*, 186 F.3d at 1314, to argue that failure to define perfectly the population of qualified employees does not prevent the Siskin Study from establishing their claim. Both decisions are distinguishable, however, because the missing variables considered by the two courts did not relate to minimal, objective qualifications. A jury could decide that the missing variable in those cases was not likely to affect the exercise of discretion to a significant extent—a rather different matter from ignoring a factor that disqualifies a candidate before discretion comes into play.

In *Bazemore* the United States and others brought a pattern-or-practice suit against the North Carolina Agricultural Extension Service, alleging racial discrimination in salaries; the plaintiffs offered statistical evidence that controlled for race, education, tenure, and job title. *See* 478 U.S. at 398. The court of appeals had upheld the district court’s rejection of the statistical evidence, ruling that “the regression analysis presented here must be considered unacceptable as evidence of discrimination,” because it “omitted . . . variables which ought to be reasonably viewed as determinants of salary,” *id.* at 399-400 (quoting *Bazemore v. Friday*, 751 F.2d 662, 672 (4th Cir. 1984)), particularly geographic variations in salary, *Bazemore*, 478 U.S. at 399. The Supreme Court disagreed. Introducing its analysis it noted that “if the defendants have not succeeded in having a case dismissed on the ground that plaintiffs have failed to establish a prima facie case,

and have responded to the plaintiffs' proof by offering evidence of their own, the factfinder then must decide whether the plaintiffs have demonstrated a pattern or practice of discrimination by a preponderance of the evidence." *Id.* at 398. It then said, "[I]t is clear that a regression analysis that includes less than 'all measurable variables' may serve to prove a plaintiff's case Whether, in fact, such a regression analysis does carry the plaintiff's ultimate burden will depend in a given case on the factual context of each case in light of all the evidence presented by both the plaintiff and the defendant." *Id.* at 400 (internal citation omitted). The Court remanded for consideration of the particular characteristics of the regression analysis to determine whether it was sufficiently probative. As we understand the opinion, the regression analysis was not used to establish the prima facie case, but to prove discrimination once the presence of a prima facie case was established, or at least uncontested. The prima facie case required a showing that qualified blacks were receiving lower salaries than qualified whites. The regression analysis then examined whether other factors—such as education, tenure, and job title—could account for this difference. Failure to take into account all potential factors did not necessarily render the regression analysis unprobative.

Moreover, there was no question in *Bazemore*, as there is here, regarding whether the data concerned persons who were not qualified or eligible for the benefit at issue—namely, a higher salary. The statistical study compared salaries

of persons employed by the extension service. The issue was whether discretion in setting salaries was exercised in a discriminatory manner. The statistical study took into account some factors that might influence the exercise of discretion and omitted others; but there is no indication that it omitted any factor that was a nondiscretionary determinant of salary (such as a maximum salary established for a specific job title).

Bullington considered a disparate-impact claim of gender discrimination in the hiring of airline flight officers. *See* 186 F.3d at 1312. The plaintiff offered statistics indicating that the interview pass rate for women was only 60% of the pass rate for men. *See id.* The district court rejected the statistical study and granted summary judgment to the airline, because the pass rates were not adjusted for hypothesized differences in aeronautical experience between men and women. *Id.* at 1312-13. We reversed, concluding that the statistics were “sufficiently reliable” because the study was properly limited to individuals who were minimally qualified for the positions (otherwise they would not have been given an interview in the first place). *Id.* at 1314. Again, the statistical study examined only the exercise of discretion—this time in the interview process. A jury can weigh whether omission of a factor that could affect the exercise of discretion renders an analysis unpersuasive. But no one could disagree that an objective eligibility requirement is a necessary component of the analysis.

Thus, the variables omitted from the regression analyses in both *Bazemore* and *Bullington* related to characteristics that did not affect whether the population studied was “minimally qualified” for the benefit sought. The geographic variations in salary at issue in *Bazemore* had nothing to do with whether a particular individual was minimally qualified to receive a higher salary. *See* 478 U.S. at 398. In *Bullington* the level of aeronautical experience was certainly a permissible consideration in the interview process, but it was only one of many factors considered in a subjective determination, not a mandatory criterion for being hired. In contrast, the Siskin Study did not confine itself to the persons eligible for an overtime assignment. *See Ortega*, 943 F.2d at 1245 (statistics not probative because they did not take into account qualifications for the jobs available).

Our conclusion is not undermined by the “massive overtime disparities” that Plaintiffs allege are revealed by the Siskin Study’s analysis. Rep. Aplt. Br. at 38. They contend that these disparities are so large that a substantial disparity would certainly be present even if the statistical analysis were adjusted to account for the CBA requirements. Plaintiffs argue:

Boeing’s expert report affirmatively demonstrated that the gender of the employee regularly assigned to the machine or position for which overtime work was needed was largely irrelevant to the overtime disparities. Boeing’s expert determined that, for 78% of overtime opportunities, more than one employee was eligible to work the overtime. Thus, even if there had been reason to believe that more men than women were the sole employees regularly assigned to the

machine or position for which overtime work was required, less than 25% of the overtime opportunities studied by Siskin would have been affected by his omission of this variable. Viewed in the light most favorable to plaintiffs, this fact casts grave doubt on whether the omission of this variable explains the massive overtime disparities found by Siskin, precluding the district court from rejecting Siskin's analysis on this ground.

Id. at 37-38 (record citation omitted). As we now proceed to explain, however, this argument misapprehends the statistical evidence by confusing the *magnitude* of the disparities with their *level of statistical significance*, as measured in standard deviations.

There is no dispute that the Siskin Study's regression analysis reflected a difference in the amount of overtime worked by men and women that was many standard deviations removed from equality. The Siskin Study computed departures from equal treatment of men and women whose statistical significance ranged from 7.95 standard deviations (weekend overtime during 2002) to 38.03 (weekday overtime during 1999). That statistical significance, however, does not necessarily mean that the departure from equality was large. For example, the Siskin Study calculated that women worked an average of 19% fewer hours of weekday overtime in 1999, 17% fewer in 2000 and 2001, and 11% fewer in 2002. For weekend overtime it calculated that women worked an average of 18% fewer overtime hours in 1999, 19% fewer in 2000, 18% fewer in 2001, and 10% fewer in 2002. Although notable, these are not what most would call "massive disparities"—it is nothing like men receiving proportionately even twice as much

overtime as women. Indeed, guidelines from the Equal Employment Opportunity Commission draw a line (albeit not a rigid one) at a 20% disparity:

A selection rate for any race, sex, or ethnic group which is less than four-fifths (4/5)(or eighty percent) of the rate for the group with the highest rate will generally be regarded by the Federal enforcement agencies as evidence of adverse impact, while a greater than four-fifths rate will generally not be regarded by Federal enforcement agencies as evidence of adverse impact.

29 C.F.R. § 1607.4(D); *accord* 28 C.F.R. § 50.14(4)(D) (Department of Justice Guidelines); *see Smith*, 196 F.3d at 365 (treating “four-fifths” guideline as persuasive); *Thomas v. Metroflight, Inc.*, 814 F.2d 1506, 1511 n.4 (identifying the EEOC guideline as “[o]ne possible index of substantial disparity”); *cf. Maldonado*, 433 F.3d at 1305 (“EEOC guidelines, while not controlling upon the courts by reason of their authority, do constitute a body of experience and informed judgment to which courts and litigants may properly resort for guidance.” (internal quotation marks omitted)).

What the large number of standard deviations means is that the departure from equality, whatever its magnitude, is highly unlikely to be random. Of course, when there are massive disparities, the difference may be many standard deviations. But when, as here, there is a great deal of data, even a relatively small difference may be highly statistically significant (that is, unlikely to be random). Consider an experiment involving 1,000,000 flips of a coin. The canonical result, of course, would be 500,000 heads and 500,000 tails. Say, the

results were 510,000 heads and 490,000 tails. Although the magnitude of the difference is small, only about 4% more heads than tails, the odds of such a difference occurring in the absence of a weighted coin are exceedingly small—the departure from equality is 20 standard deviations. The difference strongly indicates *some* influence on the results other than the operation of pure chance.

Likewise, under the Siskin Study’s analysis, it is very, very unlikely that the difference in the assignment of overtime to men and women with the same job, grade, budget code, and shift is a random event. As the Siskin Study observed, “Clearly, something in the overtime process consistently results in males obtaining more overtime and working more overtime than females.” R. Doc. 346 at 23. The large number of standard deviations tells us nothing about what that “something” is, however, other than that it is not based on differences in job, grade, budget code, or shift.

As a result, it could be very important, contrary to Plaintiffs’ brief, that nearly a quarter of the overtime opportunities were in work done by only one person if, as impliedly assumed in the above-quoted passage from Plaintiffs’ brief, men disproportionately held those positions or the offers of overtime were concentrated in such positions held by men. (In such situations there is very little, if any, supervisor discretion in the assignment of overtime, because the CBA provides that the person who normally performs the work should be offered it first.) Similarly, it could be quite important if men are disproportionately

employed in crews in which overtime is available to everyone in the unit. Of course, such gender disparities in these positions could indicate discrimination in hiring for those jobs, but that is not the claim made by Plaintiffs. *See Price v. City of Chicago*, 251 F.3d 656, 661 (7th Cir. 2001) (plaintiff's statistical showing that eligibility test may produce disparate impact could not establish prima facie case when the test's use is not the employment practice complained of).

An illustration may make this proposition clearer. Boeing's expert, Dr. Ward, conducted a study on overtime assignments that controlled for the CBA criteria by surveying individual managers about the actual offers made to eligible employees. For each overtime assignment, Dr. Ward's study determined who was eligible under the CBA and then measured whether men were disproportionately selected for the overtime. These data were collected for only a portion of the Wichita facility and only for a two-month period in 2003, so the study is hardly dispositive of whether discrimination occurred. But the results are instructive.

When only one employee normally performed the work and was eligible for the overtime assignment, women received 14% (76 of 535) of the overtime offers, precisely what would be expected (according to the report) given their representation in the jobs from which those overtime assignments were made. When multiple workers were eligible under the CBA for the assignment, women received 23% (430 of 1855) of the offers, very slightly *more* than would be expected. Overall, women received 21% (506 of 2390) of the overtime offers.

From these data it appears likely that women were significantly underrepresented in those jobs for which only one worker was eligible for particular overtime assignments, and even though those jobs accounted for only 22% (535 of 2390) of the overtime assignments studied, this underrepresentation decreased women's percentage of overtime offers from 23% (when more than one employee was eligible) to 21% (the overall rate). That is approximately a 9% reduction in the offer rate to women (21 is 91% of 23). In other words, contrary to what one would expect if the above-quoted argument of Plaintiffs were valid, women received 9% fewer offers than one would expect if one looked only at positions for which more than one worker was eligible.

Dr. Ward's study illustrates how disparities of the magnitude of those found in the Siskin Study could result solely from underrepresentation of women in jobs for which only one employee is eligible for overtime assignments. Likewise, even in situations in which several employees are eligible to work overtime, women could be underrepresented in the crews most likely to receive calls for overtime work. Yet, Plaintiffs have not shown how the Siskin Study parameters would account for such underrepresentation.

We do not mean to make too much of Dr. Ward's study. We certainly are not saying that it disproves Plaintiffs' allegations of disparate impact. But that study clearly shows the flaws in the reasoning of Plaintiffs' brief—that the large number of standard deviations calculated in the Siskin Study makes it

unnecessary to determine whether the parameters used in that study are good proxies for the CBA eligibility requirements. To repeat, the very large number of standard deviations does not mean that the gross difference in the amount of overtime worked by men and women is itself large; it just means that the difference is very unlikely to be random. But since the CBA requirements not included in the Siskin Study model are not random, and may well impact men and women differently (as Dr. Ward's study suggests), the results of the Siskin Study are consistent with the CBA requirements being the cause of the disparity in overtime assignments—at least in the absence of evidence that the Siskin Study's parameters are reliable proxies for the CBA requirements. There being no such evidence, the Siskin Study does not satisfy Plaintiffs' burden to establish a prima facie case.

IV. CLAIMS OF DEAN PLAINTIFFS

Also on appeal are claims by the Dean Plaintiffs. They ask us to review the district court's denial of reconsideration of their removal as class representatives and denial of their motion for recusal of Judge Brown. The district court included these rulings in its certification of a final judgment under Fed. R. Civ. P. 54(b). We review the rulings on both motions for an abuse of discretion. *See Price v. Philpot*, 420 F.3d 1158, 1167 n.9 (10th Cir. 2005) (motion to reconsider is reviewed for abuse of discretion regardless of whether it is construed as raised under Fed. R. Civ. P. 59 or 60); *Fymbo v. State Farm Fire and Cas. Co.*, 213 F.3d

1320, 1321 (10th Cir. 2000) (finding that individual is not adequate class representative is subject to abuse-of-discretion review); *Higginbotham v. Oklahoma ex rel. Oklahoma Transp. Comm'n*, 328 F.3d 638, 645 (10th Cir. 2003) (denial of motion to recuse is reviewed for abuse of discretion). Neither decision by the district court was incorrect, let alone an abuse of discretion.

The district court based its initial ruling removing the Dean Plaintiffs as class representatives on their ongoing demand to be paid a “consultant’s fee” of 15% of any attorney fees obtained by class counsel. In its denial of their motion to reconsider, the court stated that the Dean Plaintiffs’ repeated public references to privileged conversations with class counsel only strengthened its initial conclusion that they put their own interests above those of the class. Given such conduct, we agree with the district court that the Dean Plaintiffs would not “fairly and adequately protect the interests of the class” as required by Rule 23(a)(4). *See Rutter & Wilbanks Corp. v. Shell Oil Co.*, 314 F.3d 1180, 1187-88 (10th Cir. 2002) (“Resolution of two questions determines legal adequacy: (1) do the named plaintiffs and their counsel have any conflicts of interest with other class members and (2) will the named plaintiffs and their counsel prosecute the action vigorously on behalf of the class?” (internal quotation marks omitted)).

Similarly, the district court did not err in denying the Dean Plaintiffs’ motion for recusal. As the court thoroughly explained in its order, their “unsubstantiated suggestions, speculations, [and] opinions,” are insufficient to

establish even the appearance of any bias, prejudice, or misconduct that would warrant judicial recusal. Rep. Aplt's. Supp. App. Vol.1 at 73 (Dist. Ct. Order of 1/7/2004); *see Bryce v. Episcopal Church in the Diocese of Colo.*, 289 F.3d 648, 659-60 (10th Cir. 2002) (discussing the standards for recusal). “[A] judge . . . has as strong a duty to sit when there is no legitimate reason to recuse as he does to recuse when the law and facts require.” *Bryce*, 289 F.3d at 659 (internal quotation marks omitted). The district judge correctly recognized his duty to continue to sit in this case.

V. CONCLUSION

For the reasons stated above, Plaintiffs’ petition for permission to appeal is DISMISSED. The district court’s summary judgment and its denial of Plaintiffs’ motion to reconsider that decision (04-3334) are AFFIRMED. The district court’s disposition of the motions by the Dean Plaintiffs (04-3350) is AFFIRMED. Boeing’s cross-appeal (04-3351) is DISMISSED as moot.