IN THE SUPREME COURT OF IOWA
Supreme Court No. 20–0617

STATE OF IOWA,
Plaintiff-Appellee,

vs.

KENNETH L. LILLY,
Defendant-Appellant.

APPEAL FROM THE IOWA DISTRICT COURT
FOR NORTH LEE COUNTY
THE HONORABLE MARY ANN BROWN, JUDGE

APPELLEE'S BRIEF

THOMAS J. MILLER
Attorney General of Iowa

LOUIS S. SLOVEN
Assistant Attorney General
Hoover State Office Building, 2nd Floor
Des Moines, Iowa 50319
(515) 281-5976
Louie.Sloven@ag.iowa.gov

CLINTON BODDICKER
Lee County Attorney

ANDREW PROSSER
Assistant Attorney General

ATTORNEYS FOR PLAINTIFF-APPELLEE

FINAL

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Federal Cases

State Cases

State Rules/Statutes

Other Authorities

STATEMENT OF THE ISSUE PRESENTED FOR REVIEW

I. Did the trial court err in determining that Lilly failed to establish that representation of African-Americans on his jury pool (or over a five-year period) was not unfair or unreasonable, relative to the proportion of the jury-eligible population of North Lee County that is African-American?

Authorities

State v. Veal, 930 N.W.2d 319 (Iowa 2019)
State v. Williams, 929 N.W.2d 621 (Iowa 2019)
Varnum v. Brien, 763 N.W.2d 862 (Iowa 2009)
Washington v. People, 186 P.3d 594 (Colo. 2008)
Iowa R. App. P. 6.801
Iowa R. App. P. 6.1101(2)(f)
Iowa R. App. P. 6.1101(3)(a)
Shamena Anwar et al., *The Impact of Jury Race in Criminal Trials*, Q.J. Econ 1017 (2012)
Peter A. Detre, *A Proposal for Measuring Underrepresentation in the Composition of the Jury Wheel*, 103 Yale L.J. 1913 (1994)
Nina W. Chernoff, *Black to the Future: The State Action Doctrine and the White Jury*, 58 Washburn L. J. 103 (2019)
Paula Hannaford-Agor, *Systematic Negligence in Jury Operations: Why the Definition of Systematic Exclusion in Fair Cross Section Claims Must Be* Expanded, 59 Drake L. Rev. 761 (2011)
Norbert Hirschauer et al., *Pitfalls of Significance Testing and p-Value Variability: An Econometrics Perspective*, 12 Stat. Surveys 136 (2018)
Jerry Kang et al., *Implicit Bias in the Courtroom*, 59 UCLA L. Rev. 1124 (2012)
Mark A. Musick et al., *Much Obliged: Volunteering, Normative Activities, and Willingness to Serve on Juries*, 40 L. & Soc. Inquiry 433 (2015)
Denez Szucs, *A Tutorial on Hunting Statistical Significance by Chasing N,* 7 Frontiers in Psychol. 1 (Sept. 2016)
Last Week Tonight: Juries (Aug. 16, 2020) at 3:10–3:30, https://youtu.be/1f2iawp0y5Y?t=191
WolframAlpha, https://www.wolframalpha.com/input/?i=1+%3D+%2814-P*1939%29%2F%E2%88%9A%281939*P*%281-P%29%29
WolframAlpha, https://www.wolframalpha.com/input/?i=-1+%3D+%281-P*75%29%2F%E2%88%9A%2875*P*%281-P%29%29
WolframAlpha, https://www.wolframalpha.com/input/?i=-1+%3D+%280-P*75%29%2F%E2%88%9A%2875*P*%281-P%29%29

ROUTING STATEMENT

Lilly seeks retention because he is appealing from a ruling on his fair-cross-section challenge. *See* Def's Br. at 8. That is not a basis for retention. This case can be resolved on prong #3 of *Duren*/*Lilly* by applying principles that were set out in Lilly's previous appeal, because Lilly offered no actual proof of systematic exclusion. *See State v. Lilly*, 930 N.W.2d 293, 304–09 (Iowa 2019). Transfer to the Iowa Court of Appeals would be appropriate. *See* Iowa R. App. P. 6.1101(3)(a).

Analyzing prong #2 is more difficult. There are substantial questions about data aggregation that can be sidestepped in this case, due to the absence of any proof of systematic exclusion on prong #3. Even so, Iowa district courts still need guidance on how to apply *Lilly*. If retained, the Iowa Supreme Court should answer lingering questions about how to select the data to analyze for fair-cross-section claims. *See* Iowa R. App. P. 6.1101(2)(f).

STATEMENT OF THE CASE

Nature of the Case

Kenneth L. Lilly was tried and convicted on one count of aiding and abetting robbery in the first degree, a Class B felony, in violation of Iowa Code sections 703.1, 711.1, and 711.2 (2017). On direct appeal, the Iowa Supreme Court affirmed the conviction but remanded for

further litigation on Lilly's claim that his jury was not drawn from a fair cross section of the jury-eligible population of North Lee County, under a new framework for assessing those challenges. *See Lilly*, 930 N.W.2d at 308. On remand, the parties presented additional evidence and argument. The district court found that (1) Lilly failed to establish an unfair or unreasonable level of underrepresentation; and (2) he also failed to prove systematic exclusion. This is his appeal from that ruling.

Facts

The facts relating to the underlying robbery are set out in the decision that resolved Lilly's direct appeal. *See Lilly*, 930 N.W.2d at 296–98, 308–09. Those facts are not relevant to this appeal.

Relevant Course of Proceedings

Two weeks before trial, Lilly filed a motion that challenged the representativeness of the jury pool under the Sixth Amendment and under Article I, Section 10 of the Iowa Constitution. *See* Challenge to Jury Pool (9/14/17); App. 6. The State resisted. *See* Resistance (9/14/17); App. 12. A hearing was held, and the parties presented testimony from the North Lee County jury manager (Dawn Willson) and the IT director for the State Judicial Branch (Mark Headlee). *See* Transcript (9/22/17). The district court noted that, from the records

compiled about North Lee County jury pools over the past five years, there appeared to be a pattern of underrepresentation (because levels of African-American representation were consistently falling short of then-available estimates of the percentage of the jury-eligible people in North Lee County who were African-American). *See* Transcript 32:25–33:6. But it rejected Lilly's argument that it "stands [to] reason" that proof of persistent underrepresentation meant that "some aspect of the jury selection procedure is causing that underrepresentation." *See* Transcript 30:14–31:2; *accord Lilly*, 930 N.W.2d at 305–06 & n.8 (rejecting similar argument from dicta in *State v. Plain*, 898 N.W.2d 801, 824 (Iowa 2017) and explaining that *Plain* "repeatedly noted that the defendant had the burden to establish systematic exclusion, not merely underrepresentation"). Instead, the district court ruled:

> Even if the jury panels are not representative of the African-American population in the community, in order for the defendant to challenge the panel he must still prove that the underrepresentation is due to a systematic exclusion of the group in the jury selection process. . . .
>
> [. . .]
>
> . . . Mr. Headlee testified that the judicial branch utilize the services of Conduent Company to merge the files they receive from the Iowa Secretary of State's Office and the DOT. . . . Consequently, the list created through the State Court Administrator's office complies with Iowa Code. . . .

> No evidence has been presented to this court that there is any other list available that could be used in a systematic random selection process that would increase the representation of African-Americans on the jury list. What other readily available and discernable list of names is available? What more could those creating the list do to increase the number of African-Americans on the list? Without that information even being discussed, there is no evidence that the underrepresentation of African-Americans is due to a systematic exclusion of the group in the jury selection process.

Order Denying Challenge (9/25/17) at 5–8; App. 27–30.

Lilly was subsequently tried before a jury that was drawn from that jury pool. Although no African-American people were among the 50 potential jurors who appeared on the morning of trial, there were 25 other respondents who received and returned jury questionnaires. One of them indicated their race was "black/white." Another person listed their race as "mixed." *See* RemandTr. (1/23/20) 26:2–34:6; *see also* Remand Ex. D, at 12–13; C-App. 73–74.

Lilly was convicted. He appealed. The Iowa Supreme Court retained the appeal and adopted a new analytical framework for these claims under Article I, Section 10 of the Iowa Constitution. *See Lilly*, 930 N.W.2d at 301–08; *cf. State v. Veal*, 930 N.W.2d 319, 328–30 (Iowa 2019) (adopting new analytical framework for similar claims under the Sixth Amendment). It rejected Lilly's other challenges and

conditionally affirmed Lilly's conviction. Then, it remanded the case "to give Lilly a further opportunity to develop his arguments that his Sixth Amendment and article I, section 10 rights to an impartial jury were violated" and to hold a new trial "[i]f the district court concludes a violation occurred." *See Lilly*, 930 N.W.2d at 308.

On remand, Lilly presented additional testimony from Willson, the North Lee County jury manager. He argued that jury panel data from the past five years that showed a pattern of underrepresentation, together with data on income disparities, was sufficient to establish that the source lists for jury selection were systematically excluding African-Americans living in North Lee County. *See* Def's Remand Br. (1/22/20) at 9–14; App. 93–98; RemandTr. 64:11–72:3. In response, the State noted that this was proof of correlation, but not causation. *See* RemandTr. 74:3–77:10. It also explained that adding a source list at the last stage, just before drawing names of potential jurors, would distort the jury selection process by adding duplicate names:

> . . . The software that we use [at the state level] combines the list for . . . voter registration and for driver's license and non-operator I.D. and then removes duplicates before it gets to the jury manager.
>
> If the jury manager takes that and then adds in SNAP recipients, adds in some list of mobile phone subscribers, suddenly you're reintroducing duplicates. People who both register to vote and receive SNAP benefits are selected for

> jury service twice as often. People who also have a mobile phone subscription, three times as often.
>
> In addition, Your Honor, to what you identified that there can be possible hidden biases in additional lists, you're also compounding a duplication problem with every additional list that you have. And you're sacrificing the randomness that's created by getting as many names as possible, running a de-duplication process, and then randomly selecting them.

RemandTr. 78:12–79:7. The State also pointed out that, while there was not a reliable source for demographic data that would specifically describe North Lee County, it was still possible to remove residents of the largest city in South Lee County from the calculations—and it was important to do that, because Keokuk had a disproportionately high concentration of African-American residents that had been inflating expected levels of African-American representation on jury pools in North Lee County to unrealistic levels. *See* RemandTr. 85:24–88:7; Remand Ex. K; App. 193. Lilly argued that this was still inaccurate, because it would still include any people who lived in other parts of South Lee County. *See* RemandTr. 61:6–19. The State agreed, and it explained that it could not find data on the non-Keokuk population of South Lee County. But the State showed that, after removing Keokuk, even implausibly generous assumptions could not rescue Lilly's claim:

> Let's just say that [subtracting the population of South Lee County who does not live in Keokuk] doubles the percentage [of African-Americans] . . . If I take out another 9,000 or so people [who are not African-American] from the total amount of North Lee County, let's just say now that we're at 0.832 percent. Turns out the expected value when you multiply that by 1,939 [which is the number of potential jurors who identified their race over five years of aggregated jury pool data], you get 16.13. That's how many African-American people you would expect to be randomly selected for jury duty among that 1,939 over the course of five years.
>
> Now, what's the actual figure? As [Lilly's counsel] points out in his brief, it was 14 people. The *Lilly* test is if you're within one standard deviation of the mean of the expected value, it is not unfair or unreasonable.
>
> And the standard deviation for this parameter for a sample of this size works out to be to three decimal places, [4.000]. So we're within the one standard deviation range set out in *Lilly*.

RemandTr. 88:8–89:2. The State also explained why Iowa courts should not use aggregated data for this analysis, unless the jury pool and the distinctive group's jury-eligible population in that county are too small to create a fair expectation of non-zero representation on the jury pool that was used for the defendant's trial. *See* RemandTr. 80:4–83:14; Remand Tr. 85:4–23; State's Br. on Prong #2 (1/21/20) at 8–20; App. 70–82.

The district court accepted the files that were attached to the briefs on remand as exhibits. *See* Ruling (4/7/20) at 2; App. 198. It noted that residents of South Lee County were not eligible to serve on

juries in North Lee County. *See id.* at 4–5; App. 200 (citing *Froman v. Keokuk Health Systems*, Inc., 755 N.W.2d 528, 531 (Iowa 2008)). As a result, "if the court uses only jury lists from North Lee County but uses demographic population for the entire county, it will be mixing apples and oranges." *See id.* at 5; App. 201. It found that deducting the known population of Keokuk would produce a range of estimates for North Lee County that were closer to its actual demographics:

> The State acknowledged that its calculation for North Lee County versus South Lee County is not precise. The imprecision results because any African Americans that might live in South Lee County but not in the city of Keokuk are still included in the North Lee County population numbers. . . . The State's method also makes the total number of all race adults, 18,735; higher than would be eligible be for North Lee jury duty, as it includes people of all races that live in rural South Lee as well as the handful of small towns in South Lee. Recognizing that the potential African-American percentage could be slightly low, the State agreed to double the percentage to .832%. The numbers provided by the state show 7806 adults in Keokuk and 343 of them are African-Americans. If Keokuk alone was evaluated, the percentage of jury-eligible African-Americans in Keokuk is 4.3%. With the county-wide percentage being 1.59%, it is clear including Keokuk residents would inordinately increase the percentage for North Lee. With the data available there is no way the North Lee percentage of jury-eligible persons in the distinctive group of African-Americans could be greater than .832%.
>
> In *Lilly* the Court directed this court on remand to "… rely on 'the statistical data that best approximates the percentage of the jury-eligible' persons in the distinctive group." *Lilly* at 305, quoting *United States v. Torres-*

> *Hernandez*, 447 F.3rd 699, 704 (9th Cir. 2006). That data has established that .461% to no more than .832% of the eligible North Lee jurors are African-American.

See id. at 5–7; App. 201–03. It used that 0.832% figure to analyze the aggregated data from five years of jury pools in North Lee County:

> If we use the aggregated data proposed by the Defendant, the data shows that over a five-year period leading up to and including the Lilly trial, 1,939 people were chosen at random who identified their race. Of those, 14 jurors identified as African-American. Using the generous percentage of .832% of the total being African-Americans who are jury eligible of the population in North Lee County; that would mean 16.13 jurors would be expected to be African-American. The records reflect that 14 jurors identified as African-American. A standard deviation of one, with these numbers is 4. The 14 jurors are within one standard deviation of the expected value.

Id. at 7; App. 203. Based on that, Lilly's challenge failed on prong #2.

The district court also remarked that, on systematic exclusion, Lilly's argument that it should exercise supervisory authority to order the North Lee County jury manager (or any other) to use additional source lists would jeopardize "uniformity across the state of Iowa on how jury pools are created," and it warned that "[b]ias and prejudice could result from a single judge making his or her own decision on what lists should be used." *See id.* at 7–8; App. 203–04. It doubted that it had authority to demand lists of names and addresses from other government agencies or from private businesses, and it also doubted

that it had "power or authority to dictate which lists are used." *See id.* at 8–9; App. 204–05. In any event, it found that Lilly failed to establish that using these two source lists was systematic exclusion, because its analysis on prong #2 had already shown that "in this particular county, the master list used to create the jury pools has not created pools that underrepresented African-Americans." *See id.* at 9; App. 205.

Additional facts will be discussed when relevant.

ARGUMENT

I. The district court did not err in rejecting Lilly's claim.

Preservation of Error

Error was preserved for any challenges and arguments that Lilly raised and that the district court rejected in its ruling. *See Lamasters v. State*, 821 N.W.2d 856, 864 (Iowa 2012). This means, on prong #3, error is only preserved to renew his argument that proof of correlation between racial disparities in income and racial disparities in levels of representation on jury pools is sufficient to prove systematic exclusion was caused by the use of voter registration records and DOT records as source lists. *See* Def's Remand Br. (1/22/20) at 9–14; App. 93–98; RemandTr. 64:11–72:3; Ruling (4/7/20) at 7–9; App. 203–05.

For the State, error is preserved to defend the basis for the district court's ruling, and to renew any argument that would provide an alternative basis for denying Lilly's claim, as long as that argument was urged below. *See King v. State*, 818 N.W.2d 1, 11–12 (Iowa 2012); *DeVoss v. State*, 648 N.W.2d 56, 61–63 (Iowa 2002).

Standard of Review

Review of a ruling on a fair-cross-section challenge is de novo. *See Lilly*, 930 N.W.2d at 298 (quoting *Plain*, 898 N.W.2d at 810).

Merits

In *Lilly*, the Iowa Supreme Court confirmed that Iowa follows *Duren* and requires proof of three elements to "establish a prima facie violation of the fair-cross-section requirement":

> (1) that the group alleged to be excluded is a "distinctive" group in the community; (2) that the representation of this group in venires from which juries are selected is not fair and reasonable in relation to the number of such persons in the community; and (3) that this underrepresentation is due to systematic exclusion of the group in the jury-selection process.

See Lilly, 930 N.W.2d at 299 (quoting *Plain*, 898 N.W.2d at 821–22 (quoting *Duren v. Missouri*, 439 U.S. 357, 364 (1979))); *cf. id.* at 301 (noting Iowa courts "will apply the *Duren*/*Plain* three-part test under the Iowa Constitution, reserving the right to apply it differently").

There is no dispute on prong #1. African-Americans are a distinctive group. There is no group membership requirement; like the petitioner in *Duren* (who was a man who alleged underrepresentation and systematic exclusion of female jurors), *any* defendant may allege underrepresentation and systematic exclusion of African-Americans. *See Duren*, 439 U.S. at 360–64; *Peters v. Kiff*, 407 U.S. 493, 496–504 (1972) ("[T]he existence of a constitutional violation does not depend on the circumstances of the person making the claim."). *But see Plain*,

898 N.W.2d at 822 (stating that first prong of *Duren* requires that "a defendant must establish membership in a distinctive group"); *State v. Shaw*, No. 18–0421, 2019 WL 5790884, at *3 n.2 (Iowa Ct. App. Nov. 6, 2019) (quoting that segment of *Plain* and noting "the State correctly identifies that this assertion is contradicted by the United States Supreme Court's Sixth Amendment precedent").

However, Lilly failed to establish prong #2 and prong #3. Because Lilly's challenge fails if he cannot establish all three prongs, this case can be resolved on prong #3. It is not necessary to analyze prong #2. However, this Court may use the opportunity to resolve lingering questions about the permissible uses of aggregated data.

A. Lilly failed to establish systematic exclusion.

Lilly's argument is that there is a correlation between racial disparities in income and racial disparities in levels of representation on jury pools in North Lee County. *See* Def's Br. at 19–26. Lilly argues that source lists must be causing the apparent underrepresentation, and that it is "mismanagement" amounting to systematic exclusion "to not supplement the jury lists when it was well known that the existing lists underrepresented African Americans." *See* Def's Br. at 26–31. In its amicus brief, the NAACP reiterates the same argument:

that African-Americans "almost certainly are underrepresented" on those source lists, because African-Americans are disproportionately likely to be "living at poverty level or below, of low per capita income, and of low household income in Lee County." *See* Amicus Br. at 37–38. But the amicus brief also admits that there is a total absence of proof for that hypothesis. *See* Amicus Br. at 38–39.[1] This burden of proof was expressly assigned to Lilly, in the ruling on his direct appeal. *See Lilly*, 930 N.W.2d at 307–08 ("[T]he defendant must prove that the [challenged] practice has caused systematic underrepresentation."). Lilly has failed to carry that burden—he has not actually proven that African-Americans are underrepresented on these source lists.

---

[1] The amicus brief introduces new factual material from outside the record, in the form of an e-mail from a former Iowa DOT director. *See* Amicus Br. at 38–39. It is proper for appellate briefs to introduce "constitutional facts"—but the NAACP's brief is attempting to provide something equivalent to testimony about new *adjudicative facts* to establish causation, as "a finding of facts relating to the parties and their particular circumstances." *See Varnum v. Brien*, 763 N.W.2d 862, 880–81 (Iowa 2009) (citing *Greenwood Manor v. Iowa Dep't of Pub. Health*, 641 N.W.2d 823, 836 (Iowa 2002)). This is improper. *See* Iowa R. App. P. 6.801 (defining scope of record on appeal); *accord In re Marriage of Keith*, 513 N.W.2d 769, 771 (Iowa Ct. App. 1994) ("[C]ounsel has referred to matters apparently not a part of the record of this appeal. We admonish counsel to refrain from such violations of the rules of appellate procedure. We are limited to the record before us and any matters outside the record on appeal are disregarded.").

Both Lilly and the NAACP seize upon *Lilly*'s use of scholarship from Paula Hannaford-Agor—but both ignore the specific language that *Lilly* quoted from her work, to describe the burden of proof:

> Litigants alleging a violation of the fair cross section requirement would still have to demonstrate that the underrepresentation was the result of the court's failure to practice effective jury system management. This would almost always require expert testimony concerning the precise point of the juror summoning and qualification process in which members of distinctive groups were excluded from the jury pool and a plausible explanation of how the operation of the jury system resulted in their exclusion. Mere speculation about the possible causes of underrepresentation will not substitute for a credible showing of evidence supporting those allegations.

Lilly, 930 N.W.2d at 307 (quoting Paula Hannaford-Agor, *Systematic Negligence in Jury Operations: Why the Definition of Systematic Exclusion in Fair Cross Section Claims Must Be Expanded*, 59 DRAKE L. REV. 761, 790–91 (2011)). Here, there was *no* testimony about the demographic make-up of the people included on the source lists, nor was there any testimony about demographic characteristics of the people who were selected for jury service whose questionnaires were undeliverable, who did not respond, or who did not mark their race. *See* RemandTr. 23:13–24:7; RemandTr. 32:20–34:6; RemandTr. 37:11–38:4. Lilly offers nothing beyond a claim equating correlation to causation, which the Iowa Supreme Court has already rejected.

Lilly seems to recognize this, because he "asks the court to reconsider its holding in *Lilly* requiring the defendant to show the underrepresentation was caused by some aspect of the system." *See* Def's Br. at 32 (citing *Lilly*, 930 N.W.2d at 306). That would be a profound change in the law, and it would make it impossible to bring any criminal defendant to trial in any Iowa county, until we succeed in eradicating every kind of disparity that directly or indirectly affects willingness to participate in jury service—if that is even possible. *See Lilly*, 930 N.W.2d at 307 ("We are reluctant to impose an open-ended obligation on lower courts to follow unspecified 'known best practices,' whatever those best practices may turn out to be."). Researchers have found that, even after controlling for income, age, prior jury service, and a wide variety of other factors, African-American identity made a respondent's willingness to serve on a jury decline by about 60%.

> . . . The race effect is consistent with a broader pattern of results showing that African Americans experience less support for and a greater sense of alienation from the legal system (e.g., Hagan, Payne, and Shedd 2005), including, apparently, greater reservations about serving as a decision maker. Consistent with past research looking at people who have served (Shuman and Hamilton 1992; Rose 2005; Denver 2011), the lower level of willingness to serve among African Americans in these data survives controls for past experience on a jury, and it was unique to African Americans (i.e., Hispanics did not differ from whites in the final model).

See Mark A. Musick et al., *Much Obliged: Volunteering, Normative Activities, and Willingness to Serve on Juries*, 40 L. & SOC. INQUIRY 433, 442–52 & tbl.3, 457 (2015). Other researchers found "evidence that African-Americans and Latinos fail to respond to jury summons at a disproportional rate," and attributed that finding to a correlation between low income and lower response rates. *See* Nina W. Chernoff, *Black to the Future: The State Action Doctrine and the White Jury*, 58 WASHBURN L. J. 103, 123–24 (2019) ("[W]hen income is controlled for, the response rate for African-Americans and Latinos is the same as whites."); *accord* Hannaford-Agor, *Systematic Negligence*, 59 DRAKE L. REV. at 774 & n.71 (summarizing results of 1998 study that determined that "a low-income white person was just as likely as a low-income black person to fail to appear for jury service," although it still "found a significant attitudinal difference regarding jury service between blacks and whites" that still persisted even "when education, income, and jurisdiction were controlled"). Each individual person makes their own decision about how to respond to a jury summons, and there will always be economic, social, and psychological factors that affect each recipient's decision. Patterns in response rates that emerge from those choices are neither "systematic" nor "exclusion."

In addition to being foreclosed by Iowa precedent and being a bad idea on the merits, Lilly's call to discard the requirement of proof of causation for systematic exclusion is foreclosed by law of the case. *See State v. Ragland*, 812 N.W.2d 654, 658 (Iowa 2012) (quoting *Bahl v. City of Asbury*, 725 N.W.2d 317, 321 (Iowa 2006)) ("[A]n appellate decision becomes the law of the case and is controlling on both the trial court and on any further appeals in the same case."); *accord State v. Grosvenor*, 402 N.W.2d 402, 405 (Iowa 1987). In its decision that resolved Lilly's direct appeal, the Iowa Supreme Court held that a claim under Article I, Section 10 of the Iowa Constitution *does* require proof of causation on prong #3, to establish that some feature of the jury selection process or jury management system is the actual cause of observed underrepresentation. *See Lilly*, 930 N.W.2d at 305–08. Even if that were wrong on the merits, it would be binding in this case. *See State v. Cromer*, 765 N.W.2d 1, 7 n.4 (Iowa 2009) (explaining that decision of Iowa Court of Appeals in prior appeal was incorrect when it held that error had not been preserved for a particular claim during the trial proceedings, but that "[n]evertheless, this ruling became the law of the case on remand, whether the ruling was right or wrong."). Lilly needed to carry that burden on remand, and he did not do that.

Lilly simply assumes that the reason why African-Americans appear to be underrepresented on jury pools in North Lee County is that they are not selected, because they are not on these source lists—but Lilly has done nothing to prove this, nor has he shown that this is a better explanation for any apparent pattern of underrepresentation than any competing explanation (and there are plenty). This case was remanded to grant Lilly that opportunity to develop a factual record to prove systematic exclusion, and he failed to do that. It was correct to deny this challenge on prong #3, because Lilly failed to establish that any underrepresentation is the result of systematic exclusion.

Both Lilly and the NAACP end their briefs by arguing that the State has waived its opportunity to argue that a government interest justifies the use of these source lists, notwithstanding any effect that qualifies as systematic exclusion. *See* Def's Br. at 32–33; Amicus Br. at 44–45. But the State did make the argument that adding new lists after the de-duplication process would reintroduce duplicates, and would exacerbate sampling problems if the extra lists are likely to overrepresent certain groups—especially if they are lists that would only include head of household, which can introduce a gender bias. *See* RemandTr. 78:12–79:7; *accord* Ruling (4/7/20) at 7–8; App. 203

(rejecting Lilly's argument for adding new source lists at local level because "[t]here must be some uniformity across the state of Iowa on how jury pools are created," and warning that "[b]ias and prejudice could result from a single judge making his or her own decision on what lists should be used"). Moreover, it is absurd to argue that the State has waived its opportunity to respond to proof of exclusion that was never presented. If Lilly had presented actual evidence that could establish systematic exclusion, the State could have responded—but that did not happen. The State had nothing but an argument about an apparent correlation to respond to; it is incorrect to claim that it has waived any responses to proof of causation that was never offered.

B. Lilly cannot show that the level of representation of African-Americans on his jury pool was unfair or unreasonable in relation to the prevalence of African-Americans in the jury-eligible population of North Lee County.

Prong #2 involves math. The math is made harder because this trial happened in North Lee County. The first step is to determine the prevalence of African-Americans among the jury-eligible population from which the jury pool was drawn. Obtaining accurate numbers for North Lee County is difficult. Population data exists for Lee County, but that does not reflect *North* Lee County; only residents who live in

the northern part of Lee County, including the city of Fort Madison, are eligible for jury service at the courthouse in North Lee County. Residents of the remaining area (including Keokuk) are only eligible for jury service at the courthouse in South Lee County. *See Froman*, 755 N.W.2d at 531; *see also State v. Morgan*, 559 N.W.2d 603, 609 (Iowa 1997). On remand, the State's suggestion was to use data for Lee County (with a population of about 27,000) and then subtract Keokuk (which has a population of about 10,000). While imperfect, this gets closer to approximating the population of North Lee County, which only includes Fort Madison (with a population around 10,000) and the surrounding rural area. And it matters, because Keokuk has a disproportionately high concentration of African-American residents, so including Keokuk would generate unrealistically high expectations for levels of African-American representation on jury pools that are drawn from North Lee County. *See* Ruling (4/7/20) at 6; App. 202. That is why Lilly objects to removing Keokuk from this calculation:

> Lilly submits that such calculation would be too arbitrary and skewed. Assuming a greater number of African American live in the urban areas such a Keokuk, the State's proposal fails to account for the higher population of whites in the rural areas. The State offered to double the percentage of African Americans but that is still an arbitrary number without any basis.

Def's Br. at 18. But the jury manager testified that Keokuk residents are never drawn for jury service in North Lee County. *See* RemandTr. 42:12–43:20. They should not be included in the calculation.

It is probably true that there is a "higher population of whites in the rural areas" of South Lee County that are still part of the adjusted population figure, after starting with all of Lee County and removing Keokuk. *See* Def's Br. at 18. The State acknowledged that critique, and it suggested an assumption to over-correct for that: assume that half of all remaining Lee County residents lived in rural South Lee County, and assume that *none* of those people were African-American. That assumption would take the percentage of "Lee County minus Keokuk" that is African-American, and double it. *See* RemandTr. 86:22–88:18; RemandTr. 97:23–99:11. That assumption is implausibly generous, because it assumes that rural areas of South Lee County do not have *any* African-American residents, and because Fort Madison accounts for more than half of the remaining population of Lee County after removing Keokuk (so whatever amount of people *should* be excluded as residents of rural South Lee County, it is less than half of the total). Still, this can help to show that no further analysis is necessary: if the claim still fails, then it cannot succeed under any plausible parameter.

Here are the State's calculations. They have changed slightly from the calculations described in its filings and argument on remand to remove non-citizen residents (none of whom are African-American).

TABLE 1	African-American residents	All residents
Total population in all of Lee County	872	34,785
-(people under 18)	-166	-7,524
-(non-citizen adults)	-0	-133
-(inmates in prison in Fort Madison)	-285	-700
-(adult citizens who live in Keokuk)	-343	-7,695
Adult citizens who live in Lee County, but not in Keokuk	78	18,733

See RemandTr. 86:15–87:17; *see also* Remand Ex. A; App. 99; Remand Ex. B; App. 101; Remand Ex. K; App. 193; Remand Brief (1/21/20) at 5–6; App. 67–68. A distinctive group of 78 people is only 0.416% of a population of 18,733. If doubled, that would be 0.832%.

The State objected to using the aggregated jury pool data from the last five years of jury pools in North Lee County. But even using that aggregated data, with these numbers, the State would prevail:

TABLE 2	Using 0.416%	Using 0.832%
Expected presence in sample of 1,939	8.066 people	16.132 people
Standard deviation	2.834	4.000
Difference between actual presence (14) and expected presence	+5.934 people (overrepresentative)	-2.132 people
…that result divided by standard deviation	+2.094 (overrepresentative)	-0.533

See RemandTr. 88:24–89:2; Ruling (4/7/20) at 6–7; App. 202. This is the mathematical analysis that the district court performed, and it is correct. Even after making that implausibly generous assumption in calculating the parameter, the actual amount of African-Americans who were included on jury pools over five years in North Lee County was within one standard deviation of the expected value.[2]

This is the end of the analysis of Lilly's arguments on prong #2. Error was not preserved for any of the new claims beyond this point, including the NAACP's proposals that would inflate the parameter.

---

[2] To establish a violation of prong #2 with that aggregated data, Lilly would need African-Americans to comprise at least 0.941% of the jury-eligible population. To calculate that, plug the observed level of representation (14) and the sample size (1939) into the equation for Z-score, set the Z-score to -1, and solve for P. *See* WOLFRAMALPHA, https://www.wolframalpha.com/input/?i=-1+%3D+%2814-P*1939%29%2F%E2%88%9A%281939*P*%281-P%29%29.

Hypothetically, if Lilly showed that African-Americans made up a greater portion of the jury-eligible population of North Lee County, then using five years of aggregated data *could* enable Lilly to prevail on prong #2. The State preserved error on its argument against using aggregated data when the actual jury pool is already adequate to assess whether fair and reasonable expectations of minority representation were met, along with its proposal for how to aggregate data when the jury pool and the group's presence in the jury-eligible population are too small to give rise to expectations of representation in that pool. *See* Remand Brief (1/21/20) at 11–19; App. 73–81; RemandTr. 79:25–85:18; *DeVoss*, 648 N.W.2d at 61–63. Because Lilly alleged that any apparent underrepresentation of African-Americans was the result of their disproportionate absence from the source lists that were used to select eligible residents for jury service, the analysis should focus on the group of people who were selected for jury service and responded to the jury questionnaire (as opposed to analyzing the smaller group of people who actually appeared on the morning of trial). *See State v. Williams*, 929 N.W.2d 621, 630 (Iowa 2019). 125 names were drawn for jury service. Of those, 75 people returned jury questionnaires and were able to serve. *See* RemandTr. 26:2–28:17; *cf.* Remand Ex. C & D;

C-App. 4-88. One of those people listed their race as "black/white." *See* Remand Ex. D at 12; App. 73.[3] Even assuming that none of the potential jurors who did not mark their race were African-American, a jury pool with one African-American person out of 75 respondents cannot be unfair or unreasonable, unless African-Americans comprise more than 3.437% of the relevant jury-eligible population.[4]

The NAACP's brief proposes adjustments to population data to include multi-racial residents who are also African-American. *See* Amicus Br. at 16–30. As the NAACP points out, the State's brief in *State v. Veal* did include a similar attempt to adjust population data to include mixed-race residents who are African-American, by calculating the percentage of all single-category racial minorities in that county who were African-American, and assuming that a similar portion of the county's multi-racial residents were also African-American. *See*

---

[3] Another person listed their race as "mix," but it is not clear whether this person was African-American. *See* Remand Ex. D at 13; C-App. 74. Even counting this juror as some other race, this jury pool is not underrepresentative, so it is unnecessary to speculate.

[4] Again, this can be calculated by plugging the observed level of representation (1) and the sample size (75) into the Z-score equation, setting Z-score to -1, and then solving for P. *See* WOLFRAMALPHA, https://www.wolframalpha.com/input/?i=-1+%3D+%281-P*75%29%2F%E2%88%9A%2875*P*%281-P%29%29.

Amicus Br. at 15 (citing *Veal*, No. 17–1453, State's Br. at 31). That adjustment still makes sense as a way to include African-Americans who are multi-racial; they may be part of the same "distinctive group" for purposes of a fair-cross-section challenge, even though they are categorized differently in census data. *See, e.g.*, *United States v. Green*, 435 F.3d 1265, 1271 (10th Cir. 2006) (describing multi-factor test for determining if a group is "distinctive" under *Duren*, which includes considering whether "the group has a definite composition such as race or sex" and whether "a common thread or basic similarity in attitude, idea, or experience runs through the group").

Of course, this depends on the "distinctive group" that is identified in the allegation of underrepresentation and exclusion—none of the data that Lilly used to argue his claim (like per-capita income or poverty status) included multi-racial African-Americans. *See* Def's Remand Ex. E–N; App. 120–62. And Lilly never argued that multi-racial African-Americans were being omitted from calculations. Error was not preserved for any claim of error that alleges that the district court should have adjusted population figures to add some percentage of multi-racial people to its count of African-Americans. *See State v. Krogmann*, 804 N.W.2d 518, 523 (Iowa 2011).

Even if considered, one problem with the NAACP's pro-rata calculation is that it includes prisoners in calculating the prevalence of African-Americans in Lee County, relative to the prevalence of other racial minorities. *See* Amicus Br. at 26–27. It would be more accurate to remove prisoners *before* assessing the relative prevalence of minority racial groups among the population of adult citizens who live in Lee County and are not incarcerated.

TABLE 3	Adult citizens in Lee County	Inmates at Fort Madison	Adult citizens, minus inmates
African-American	706	-285	421
American Indian and Alaska Native	71	-16	55
Asian	96	-4	92
Native Hawaiian or Pacific Islander	0	-0	0
Some other race	131	-0	131
AA/total	706/1004	285/305	421/699
...as percentage	70.32%	93.44%	60.23%

The next problem is that simply multiplying the total amount of multi-racial adult residents of Lee County by 60.23% would still count residents of Keokuk. The NAACP finds a county-wide rate for relative prevalence of African-Americans among multi-racial residents, *then*

subtracts out the multi-racial population of Keokuk and multiplies the multi-racial population of all other parts of Lee County by that county-wide percentage. But, as discussed, the population of Keokuk is disproportionately African-American, compared to all other parts of Lee County. To assess multi-racial populations *outside* of Keokuk, the population of Keokuk should be excluded.

TABLE 4	Adult citizens in Lee County, minus inmates	Adult citizen residents of Keokuk [5]	Adult citizens, minus inmates, minus Keokuk
African-American	421	-343 (B05003B)	78
American Indian and Alaska Native	55	-36 (B05003C)	19
Asian	92	-81 (B05003D)	9
Native Hawaiian or Pacific Islander	0	-0 (B05003E)	0
Some other race	131	-68 (B05003F)	65
AA/total	421/699	343/528	78/171
...as percentage	60.23%	64.96%	45.61%

[5] The NAACP's brief provides a link to the State Data Center, which is an excellent resource for county-level demographic data. *See* Amicus Br. at 19–20, 25 n.11. Still, it does not include city-level data. These figures are generated from the U.S. Census Bureau's tables that provide 2013–2017 ACS data for Keokuk, as indicated.

Note that this still includes rural parts of South Lee County. However, this calculates relative prevalence of African-Americans *among racial minorities*—it should be unaffected by inclusion of some extra rural areas of South Lee County, unless there is reason to believe that *other racial minorities* are disproportionately prevalent in those rural areas (and the State is not aware of any).[6]

The rest of this calculation uses the same ACS data that the NAACP has provided and referenced. *See* Amicus Br. at 23–25 & n.11.

Multi-racial adult citizens residing in Lee County	481
…minus multi-racial adult citizens in Keokuk	-136
Multi-racial adult citizens in Lee County, not Keokuk	345
Relative prevalence of African-Americans among that very specific group	45.61%
Estimate of multi-racial African-Americans to add:	157.37 (round up to 158)
New total	236 / 18,733
New percentage	**1.260%**

[6] The NAACP is correct to point out that Hispanic identity is counted as an ethnicity, not a race. A person who marks "Hispanic" and "African-American" is included in counts of African-Americans and in counts of Hispanic people, but is not counted as multi-racial.

The NAACP also suggests adjusting the data to remove the estimated population of rural areas of South Lee County; it suggests starting with the total population of Lee County and removing both Keokuk and Fort Madison to get the estimated rural population—at which point "[i]t would be reasonable to allocate the Lee County rural population equally between the two judicial divisions." *See* Amicus Br. at 31–35. That seems reasonable in the abstract—but, then again, seemingly reasonable assumptions have been disproved by actual data at every step of these demographic calculations. If Lilly had proposed this during the proceedings below, the parties could have assembled a record that would help gauge the reasonableness of that assumption. Again, error is not preserved for challenges based on that calculation.

In any event, the State's proposal is still the same: whatever the correct population figure would be, after removing the rural portion of South Lee County, it would surely fall somewhere between 1.260% and a hypothetical doubled figure (2.520%) that would represent the same number of African-Americans and half as many other residents. This is different from the NAACP's claim that it is "law of the case" that this Court must decide what it thinks the actual percentage is, and then double it. *See* Amicus Br. at 18, 29–30. Rather, this is a

recognition that it *does not matter* what the precise percentage is, provided that it falls within a certain range that would lead to the same outcome. As the district court noted, this sets an upper bound that is a "generous percentage." *See* Ruling (4/7/20) at 6–7; App. 202–03. The NAACP's calculations are substantially less generous. The NAACP's estimate for the percentage of eligible jurors in all parts of Lee County other than Keokuk who were African-American is 1.7%. *See* Amicus Br. at 29–30. After applying its proposed method for estimating and removing the population of rural South Lee County, that percentage only rises to 2.09%—it increases by less than 20% of the original parameter. *See* Amicus Br. at 35.

In the end, all of these proposed calculations that adjust the data to add multi-racial residents to the African-American population of North Lee County produce estimates between 0.941% and 3.437%. For percentages within that range, aggregation makes the difference.[7]

---

[7] Based on calculations described in footnote 2, Lilly could not carry prong #2 even *with* that aggregated data, if African-Americans were less than 0.941% of North Lee County's jury-eligible population. Conversely, from calculations described in footnote 4, Lilly's actual jury pool (containing one African-American among 75 respondents) would be underrepresentative enough to carry his burden of proof on prong #2 *without* using any aggregated data if African-Americans comprised at least 3.437% of the eligible jurors in North Lee County.

By applying standard-deviation analysis to the *aggregated* data from five years of jury pools, Lilly could show a Z-Score that falls below -1 and he could carry his burden on prong #2. However, if that analysis were limited to Lilly's actual jury pool (with one African-American among 75 potential jurors) then this claim would fail on prong #2—the Z-Score would be somewhere above -1, which would signify that expectations of a fair and reasonable level of representation *on that jury pool* were met. Thus, if Lilly's claim prevails on prong #3 and if this Court decides to apply an adjustment to population data that was not considered below, then this Court would need to resolve the two unanswered questions from the prior appeal. *See Lilly*, 930 N.W.2d at 305 n.7. First, *when* should aggregated data be used? Second, when aggregating data, *how much* aggregated data should be analyzed?

The State's answer to the first question can be explained like this:

Proposed decision matrix for *Lilly*	Pool is adequate; total absence results in Z-score below -1.	Pool is small; total absence results in Z-score above -1.
Actual representation is greater than zero.	Non-zero expectation, requires **analysis** of this jury pool only	Expectation is zero and is exceeded — prong #2 **failed**
Actual representation is exactly zero.	Non-zero expectation, and it is not met — prong #2 **proven**	Expectation is zero, and must aggregate data for **analysis**

During the prior appeal, the Iowa Supreme Court expressed concerns about situations where the jury pool is too small to enable a litigant to establish prong #2 under any set of facts—where total absence of the distinctive group from the jury pool is less than one standard deviation below the expected average level of representation for that group on a jury pool of that size. *See Lilly*, 930 N.W.2d at 305. In that situation, some amount of aggregation should be permitted, to enable challenges to the total exclusion of smaller distinctive groups from jury service. This makes sense. If the Iowa Supreme Court were comfortable with an approach that would automatically reject such claims on prong #2, it would not have overruled *Jones* in the first place. *See Plain*, 898 N.W.2d at 825 (overruling *State v. Jones* after deciding to reject its 10% absolute disparity threshold because "an African-American could not establish a racially unrepresentative jury [in any Iowa county] . . . even if the exclusion of African-Americans was total and systematic"); *accord Lilly*, 930 N.W.2d at 302–03 (rejecting State's proposal for a 3% threshold for absolute disparity, because it "has the same defect" and "gives a free pass to systematic underrepresentation so long as the absolute underrepresentation . . . falls below a certain threshold").

That situation is in the blue cell (lower-right) in the State's proposal. Aggregation of past data should only occur when expectations for the distinctive group's representation on the defendant's actual jury pool are so low that they are indistinguishable from zero (so that when the group is absent, that is still not enough to fall short of expectations).

But what if the jury pool and the distinctive group's presence in the jury-eligible population are already large enough to give rise to an expectation of fair and reasonable representation on that jury pool, at some non-zero level? If total absence of the group from that jury pool would satisfy prong #2 under *Lilly*, then there is no need to dilute the analysis by adding any additional data: the level of representation on this jury pool was either fair and reasonable, or it was not. With each additional jury pool that is aggregated into the dataset, the analysis is more detached from the central question of whether *this defendant's* jury was drawn from a jury pool with unfair or unreasonable levels of underrepresentation, compared to the jury-eligible population. True, the level of African-American representation on the jury pool that was drawn in September 2012 may be relevant in proving (or disproving) a claim that there was a pattern of observable underrepresentation over time, so it could matter for systematic exclusion on prong #3.

But nothing that happened in September 2012 could affect the analysis of whether Lilly's jury pool met quantifiable expectations for levels of African-American representation. *See* Pretrial Ex. A (9/27/17) at 7; App. 38. And the relevance of older jury pool data is especially weak because, as the NAACP points out, demographics can shift over time—past jury pools may seem underrepresentative (or overrepresentative) when measured against more recent demographic data, even if they were perfect cross-sections of the jury-eligible population that existed at that point in time. *See* Amicus Br. at 25 n.9 (stating that ACS data had estimated "a 65% increase in the number of multi-racial persons 18 years of age and older in Lee County" between 2010 and 2017).

Moreover, data from other jury pools is conceptually irrelevant because the constitutional harm is the impact of underrepresentation and systematic exclusion on *this defendant*, asserting a violation of *his* constitutional rights to a jury drawn from a fair cross-section of the jury-eligible population for *his* trial. Aggregated data will often drown out a claimant's actual jury pool. *See, e.g.*, *Veal*, No. 17–1453, Pet. for Rehearing (6/6/19) at 7–12; *accord* RemandTr. 79:25–83:14 ("To go five years back, most of those 1,939 people . . . have nothing to do with Mr. Lilly's trial, nothing to do with Mr. Lilly's eventual jury.").

Lilly insists that an apparent history of slight underrepresentation on prior jury pools in North Lee County should have more weight than his actual jury pool, from which his actual jury was drawn. Note that, with this five-year dataset, only a slight degree of underrepresentation would be required for Lilly to carry his burden on prong #2.

TABLE 6	P = 1.260%	P = 2.09%	P = 2.520%
Standard deviation for N=1,939	4.912	6.300	6.902
Expected level of representation among N=1,939	24.431	40.525	48.863
Lowest actual representation that is fair under *Lilly*	20 (Z=-0.902)	35 (Z=-0.877)	42 (Z=-0.994)
Highest actual representation that is unfair under *Lilly*	19 (Z=-1.106)	34 (Z=-1.036)	41 (Z=-1.139)
…as a percentage of aggregated sample of 1,939 people	0.980%	1.753%	2.114%
Degree of absolute disparity that would carry prong #2	0.280%	0.337%	0.406%

No fair-cross-section challenge to such a slight underrepresentation has ever succeeded in any American court. *See, e.g.*, *Washington v. People*, 186 P.3d 594, 605 (Colo. 2008) (collecting cases).

Lilly did adopt a requirement that a claimant must show that representation levels on their own jury pool were below-average. *See Lilly*, 930 N.W.2d at 305. That would bar Lilly's challenge if there were *two* African-Americans among 75 respondents, comprising 2.667% of the jury pool (which is above average for any parameter in this range). If Lilly can use five years of aggregated data for this kind of challenge, that would become the only way to defeat such a claim on prong #2 in North Lee County—the best jury pool that is *not* overrepresentative is the jury pool that Lilly actually had (with one African-American person out of 75 respondents), and it has a miniscule impact when combined with aggregated data from thirty other jury pools. *See* Pretrial Ex. A (9/27/17); App. 32. Conversely, in any county that has an apparent pattern of overrepresentation or adequate representation of that group on jury pools, aggregated data can make it impossible for a claimant to carry prong #2 of *Lilly*—even if a sizable group in the community is wholly absent from their jury pool.[8] *Cf. Lilly*, 930 N.W.2d at 302–03 (rejecting approaches that would give "a free pass" to total exclusion).

[8] Of course, those claims should fail on prong #3—unless the defendant can establish that members of the distinctive group were systematically excluded from *this* jury pool, despite records showing adequate levels of representation on prior jury pools in that county.

The majority in *Lilly* was "not persuaded" by arguments against aggregated data—it stated that "[i]t is unfair to restrict the defendant to the current jury pool that may have as few as seventy-five persons, and then at the same time require the defendant to furnish results that have a certain degree of statistical significance." *See Lilly*, 930 N.W.2d at 305.[9] But *Lilly* adopted a much lower threshold for its standard deviation analysis than the State had proposed (requiring a Z-Score that falls below -1, rather than -1.64). *See id.* at 304. Because of that lower threshold, a jury pool of 75 people will be large enough to enable analysis without aggregation (and total absence of members of a distinctive group will suffice as proof on prong #2), as long as the group comprises at least 1.316% of the local jury-eligible population.[10]

---

[9] By sheer coincidence, this jury pool had 75 people—but that fact was not in the record, during the prior appeal. Rather, *Lilly*'s remark about a jury pool with "as few as seventy-five persons" was a reference to data from *other* jury pools in North Lee County, showing that "[t]he typical number of responses appears to have ranged from 75 to 115" over the preceding five years. *See Lilly*, 930 N.W.2d at 299 n.3.

[10] This calculation uses the same equation for calculating Z-Score that was used in footnote 2 and footnote 4—but it sets the actual level of representation to reflect total absence (0). Then, it plugs in numbers for sample size (75) and the Z-Score that would be needed to carry the claimant's burden on prong #2 (which is -1). Then, just solve for P. *See* WOLFRAMALPHA, https://www.wolframalpha.com/input/?i=-1+%3D+%280-P*75%29%2F%E2%88%9A%2875*P*%281-P%29%29.

If the jury pool is large enough that total absence of the distinctive group produces a Z-Score that falls below -1, that means there is a non-zero expectation of fair and reasonable representation that is either met, or it is not met. Of course, with sizable groups or larger jury pools, those expectations will be higher.[11] But the point is that, as long as those expectations *exist* at a level where total absence of the distinctive group fails to meet them, a court can assess whether the level of representation in *the defendant's* jury pool satisfies them. When that happens, neither party should be entitled to a mulligan. If

---

[11] For example, in the equation used as the input in the link in the previous footnote, changing the "0" that would reflect total absence to some other number ("X") will calculate a minimum population figure that would be required for a defendant to carry prong #2 with a pool of 75 potential jurors that contains X members of the distinctive group.

Number of group members on a jury pool of 75 people	Minimum population size that would carry prong #2
1 person (1.333% of jury pool)	3.437% of eligible jurors
2 people (2.667% of jury pool)	5.240% of eligible jurors
3 people (4% of jury pool)	6.933% of eligible jurors
4 people (5.333% of jury pool)	8.565% of eligible jurors

those expectations are not met, the State should not be able to argue for aggregation of more data for prong #2, to add in more data from other jury pools with better levels of representation. But when those expectations exist and when they *are* met, the defendant should not be able to use underrepresentation on prior jury pools (which do not have any impact on the composition of the actual jury pool from which his jurors were drawn or will be drawn) to inflate his claimed injury. Neither party should be able to dodge unfavorable results by asking for aggregation of past data as a way to re-roll the dataset, if the math shows that the defendant's actual jury pool and the distinctive group were both large enough to give rise to some non-zero expectation of fair and reasonable representation within that jury pool.

Here, Lilly's actual jury pool would meet any such expectation:

TABLE 8	P = 1.260%	P = 2.09%	P = 2.520%
Standard deviation for N=75	0.966	1.239	1.357
Expected level of representation	0.945	1.568	1.89
Actual level (1) – expected level	+0.055 (overrepresentative)	-0.568	-0.89
…divided by SD (**Z-Score**)	+0.0569 (overrepresentative)	-0.458	-0.656

Note that, for the 1.260% population figure, it initially appears that aggregation of data would be required: because standard deviation is larger than the expected level of representation, a total absence of the distinctive group from that jury pool would produce a Z-Score that would be lower than -1. But the actual jury pool is *over*representative, so any further analysis becomes unnecessary. *See Lilly*, 930 N.W.2d at 305. This would occur for any hypothetical population figure that is lower than the percentage of this jury pool that was African-American, which is 1.333%. This is the situation that is described in the red cell (upper-right) of the proposed decision matrix: while the expected level of representation on this jury pool for such a small minority would be too close to zero to support a claim that total absence would be unfair, any non-zero level of observed representation is enough to *exceed* that fractional expectation and foreclose any convoluted proof of unfairness through aggregation of unrelated data from prior jury pools.

For both of the other population figures (2.09% and 2.520%), the situation is different: the standard deviation is *lower* than the average expected value. That means there is a fair and reasonable expectation of non-zero representation (under the analysis in *Lilly*) for that distinctive group on a jury pool of this size, and total absence

of that group from this jury pool would enable the defendant to carry his burden on prong #2. For those two population figures, this case would be sorted into the yellow cell (upper-left) of the decision matrix: the jury pool and the distinctive group are large enough to give rise to fair and reasonable expectations of a non-zero level of representation, which is either met or not met. When presented with this situation, Iowa courts should analyze the actual jury pool to determine whether those expectations were actually met, and the analysis on prong #2 should stop there—no matter who prevails. In these cases, data from prior jury pools should be saved for prong #3, where it may help support (or disprove) a causation theory for systematic exclusion.

Here, for both figures (2.09% and 2.520%), the actual jury pool was somewhat underrepresentative—but not to a degree that would disappoint the expectation (created by *Lilly*) that the observed level of representation should not fall below the average level by more than one standard deviation. Lilly should not be able to respond by adding in data from prior jury pools, until he gets the result that he wants. In every case where the defendant's actual jury pool can be measured against quantifiable expectations of fair and reasonable representation, that result should determine the outcome on prong #2.

Aggregation can help solve the problem of small numbers. But when claimants offer an aggregated sample that includes years of data from prior jury pools, that effectively hacks the statistical analysis and creates a problem of *large* numbers. Each time another jury pool is added to the aggregate sample, that increases sample size (N). The expected level of representation in that aggregated sample (NP) also increases, in direct proportion. But the standard deviation does not increase nearly as much—it increases in direct proportion to √(N), because standard deviation is calculated as √(NP(1-P)). That means adding more jury pools will *narrow* the range of acceptable levels of representation that fall within one standard deviation of the average. With enough aggregated data, *de minimis* underrepresentation can carry a claimant's burden on prong #2 (which raises the stakes of any fight over precise demographic estimates). These are the concerns that led some courts to reject standard-deviation analysis altogether. *See, e.g.*, *People v. Bryant*, 822 N.W.2d 124, 142 (Mich. 2012). The State is not proposing that, because standard-deviation analysis helps assess the significance of underrepresentation in relation to the size of both the jury pool and the distinctive group—it is undoubtedly superior to using either absolute disparity and comparative disparity. *See Lilly*,

930 N.W.2d at 303 (quoting *Jefferson v. Morgan*, 962 F.2d 1185, 1189 (6th Cir. 1992)). But, as a test for statistical significance, it only has a link to fairness and reasonableness of representation levels in manageable sample sizes, comprising one sizable jury pool (or a few smaller jury pools). It loses that link to fairness and reasonableness of representation levels when limitless aggregation makes the dataset arbitrarily large—with large amounts of data, even a miniscule level of underrepresentation will be statistically significant. *See United States v. Hernandez-Estrada*, 749 F.3d 1154, 1163 (9th Cir. 2014) (quoting Peter A. Detre, *A Proposal for Measuring Underrepresentation in the Composition of the Jury Wheel*, 103 YALE L.J. 1913, 1928 (1994)) ("By imagining larger and larger jury wheels, the probability of any degree of underrepresentation arising by chance can be made arbitrarily small."); *accord* Norbert Hirschauer et al., *Pitfalls of Significance Testing and p-Value Variability: An Econometrics Perspective*, 12 STAT. SURVEYS 136, 149–50 (2018) ("[A]ny effect, even if very small and irrelevant, eventually becomes statistically significant in large samples."); Denez Szucs, *A Tutorial on Hunting Statistical Significance by Chasing N,* 7 FRONTIERS IN PSYCHOL. 1, 7 (Sept. 2016) (noting that small effects will "inevitably reach [statistically] significant levels" as a dataset grows).

Lilly stated a principle of analytical integrity: litigants cannot "tip the scales in an aggregate analysis by including some earlier jury pools but not other, more recent jury pools." *See Lilly*, 930 N.W.2d at 305. The same concern for analytical integrity weighs in favor of a rule that prevents litigants from arbitrarily inflating the sample, to manufacture statistically significant underrepresentation in cases where the actual jury pool was fair and reasonable. *See United States v. Chanthadara*, 230 F.3d 1237, 1257 (10th Cir. 2000) (rejecting a statistical analysis that "merely represent[s] a manipulation of the same numbers that we have held were not sufficient to establish a prima facie violation of the Sixth Amendment"). This makes sense because it minimizes the impact of additional data that does not have any relationship to the jury pool that Lilly's jurors were drawn from. Aggregating data over a longer period of months or years may drown out any recent success at attaining more racial diversity on jury pools. If Lilly is claiming that his constitutional rights were violated, his claim should be primarily based on proof of levels of representation from the actual pool of jurors that produced his petit jury—and while aggregation can help fill in the gaps, it should not take center stage. This Court should reject any challenge based primarily on proof of underrepresentation in *other* jury pools.

The second question that *Lilly* left unanswered is: in situations where courts need to use aggregated data to assess a claim that alleges underrepresentation of a small distinctive group on a small jury pool, how much data from prior jury pools should be aggregated? *See Lilly*, 930 N.W.2d at 305 n.7. Without guidelines for aggregation, advocates will argue that data supporting their advocacy should be included in the aggregated sample, and data that hurts their advocacy should be excluded—and district courts will have no way to determine what data they should analyze. *See* RemandTr. 59:7–21 (noting that each prior jury pool either contains some African-Americans or none of them, which guarantees that selecting any stopping point for past data will give rise to a dispute between the parties over whether to include one additional jury pool in the analysis, and concluding "it's admittedly a shot in the dark about how far you aggregate and what you do").

The State prefers minimal aggregation, because every additional jury pool that is swept into the sample has the effect of lifting the focus of the standard-deviation analysis further away from the defendant's actual jury pool (which is the only jury pool, among all jury pools that were ever generated in that county, that could potentially cause some actual deprivation of the defendant's constitutional right to a fair trial).

Aggregation is only proper if the jury pool and the distinctive group are too small to enable an analysis that focuses *exclusively* on the fairness and reasonableness of the actual level of representation in the defendant's actual jury pool, because total absence of the group from that jury pool would never produce a Z-score that falls below -1. It stands to reason that aggregation of data should be calibrated to solve that problem and to enable analysis: it should continue until the aggregated sample becomes large enough to give rise to an expectation of a fair and reasonable level of representation that *would not be met* if the distinctive group were totally absent from the aggregate sample—and once that point is reached, no further data should be aggregated.

Here, there is no scenario where aggregation would be required. This is because Lilly's jury pool was large enough to analyze prong #2 for any distinctive group that comprises at least 1.316% of the relevant jury-eligible population (as shown in footnote 10: for any group that is larger than 1.316%, their total absence from a jury pool of 75 people would produce a Z-Score below -1). Aggregation would only be needed for claims that alleged underrepresentation of a group that is *smaller* than 1.316% of the population—but any such claim would instantly fail, because African-Americans comprised 1.333% of the actual jury pool.

Using any population parameter that is lower than 1.333% will result in a *positive* Z-Score, signifying overrepresentation that forecloses the challenge entirely. *See Lilly*, 930 N.W.2d at 305. It is impossible for a distinctive group to be more than 1.333% of the eligible population (which is what Lilly needs, to allege any underrepresentation at all) while also comprising less than 1.316% of the jury-eligible population (which would trigger analysis of aggregated data). Thus, logically, it is impossible for aggregation to be helpful in analyzing Lilly's challenge, no matter what the correct population figure might be.

"[S]tandard deviation varies considerably with sample size," and large enough samples can reach statistical significance through this analysis even when "no unfair underrepresentation exists by anyone's standards." *See Waller v. Butkovich*, 593 F.Supp. 942, 955 (M.D.N.C. 1984). If this Court needed to aggregate data, the State's proposal would be to aggregate data until a non-zero expectation of distinctive group representation emerged—that is, it should add in past jury pools until the total absence of the distinctive group from the aggregated sample would carry the burden of proof on prong #2. This keeps the sample size as small as possible, while focusing the analysis on the defendant's actual jury pool (as much as possible).

Hypothetically, if there had been no African-Americans on this jury pool, then it would be necessary to aggregate data if the group comprised less than 1.316% of the population. How many jury pools should be added? That would depend on the amount of eligible jurors in North Lee County who were African-American. The rule would be that aggregation should stop when total absence of African-Americans from that aggregated sample would establish a Z-Score below -1, and carry the defendant's burden on prong #2 of *Lilly*. It is possible to use the definition of Z-Score to derive an equation for the threshold that aggregation should reach (and then stop), as a function of the group's presence among the county's jury-eligible population: $N = (1-P)/P$. *See* Remand Ex. G; App. 103. But district courts may find it easier to "guess and check." Using data from Pretrial Exhibit A as an example, note that very few additional pools are needed, even for small groups:

TABLE 9	…with pool for July 2017	…with pool for May 2017	…with pool for March 2017
Total amount of potential jurors	139 (+64)	198 (+59)	260 (+62)
Total absence is -1 standard deviation if distinctive group is at least ____% of eligible population	0.714%	0.503%	0.383%

It would be unnecessary to go further back than March 2017, to analyze this challenge under any of the population figures that were proposed. But the key realization is that it is unnecessary to aggregate prior jury pools *at all*, because Lilly's jury pool (as randomly selected) contained an African-American respondent. In any situation where expected levels of African-American representation on this jury pool would be too close to zero or too fractional for a court to analyze a challenge to total absence of African-Americans, the presence of a single African-American person on the jury pool would *exceed* that fractional expectation of representation. Essentially, the presence of a single African-American person in this randomly selected jury pool means that Lilly will always be in the top row of the decision matrix for this challenge, where there is no need to aggregate data.

None of this is relevant to the preserved challenge, which did not ask the district court to adjust the population data to include any multi-racial residents of Lee County. But if this Court decides to use Lilly's unadjusted figure for Lee County that includes Keokuk (or any figures proposed by the NAACP), then aggregation becomes an issue. Because Lilly's actual jury pool was either large enough to analyze or overrepresentative, it would be inappropriate to use aggregated data.

CONCLUSION

The NAACP argues that "[w]hereas an all-white jury and a racially mixed jury convict a white defendant at about the same percentage rate, an all-white jury convicted Black defendants 81 percent of the time while a racially mixed jury did so only 66 percent of the time." *See* Amicus Br. at 36–37 (citing *Plain*, 898 N.W.2d at 825–26 (citing Shamena Anwar et al., *The Impact of Jury Race in Criminal Trials*, 127 Q.J. ECON 1017 (2012))). That is *still* not what the Anwar study really says. *See Lilly*, No. 17–1901, State's Brief at 63–67. Lilly's brief cites to a recent segment on juries from HBO's long-form comedy/news program, "Last Week Tonight." *See* Def's Br. at 31–32. During that segment, John Oliver does a better job of explaining that Anwar study's actual findings than either *Plain* or the NAACP:

> Researchers who examined felony trials in Florida found juries formed from all-white pools convict black defendants a full 16 percentage points more often than they do white defendants. But that gap in conviction rates is entirely eliminated when the pool includes at least one black member.

Last Week Tonight: Juries (Aug. 16, 2020) at 3:10–3:30, *available at* https://youtu.be/1f2iawp0y5Y?t=191 (quoting Anwar et al., *The Impact of Jury Race in Criminal Trials*, 127 Q.J. ECON at 1017).

Why does it matter that it is only *total absence from a jury pool* that affects conviction rates? For one thing, it informs the State's view that a fair-cross-section challenge arising from a group's total absence from a jury pool is a unique situation that raises special concerns—the kind of concerns that led *Lilly* to reject a flat ban on aggregated data, and led *Plain* to reject *Jones* and its 10% absolute-disparity threshold. Iowa courts should analyze prong #2 in a way that makes it possible for defendants in relatively homogenous Iowa counties to vindicate their rights to a jury pool drawn from a fair cross-section of *those* communities, even if the only way to find any reasonable expectation of non-zero minority representation in those counties is to aggregate recent jury pools into a larger sample. But when the actual jury pool contains one member of that distinctive group, that unique concern disappears—whatever fractional expectation of representation that existed in that jury pool was actually met and exceeded, and there is no need to torture the numbers until they say otherwise. Moreover, if the actual jury pool and the distinctive group are large enough that a non-zero expectation of representation exists—without aggregation—then none of the concerns identified in *Lilly* or *Plain* are implicated, and the court should assess the *actual* jury pool to analyze prong #2.

More importantly, a surface-level reading of that Florida study would support a belief that African-American jurors and white jurors are just different, in a way that would render a trial unfair whenever a petit jury is homogenous. But what the study *really* says is that, under certain conditions, an all-white petit jury behaves exactly like a more diverse petit jury. *See* Anwar et al., 127 Q.J. ECON at 1046 ("Strikingly, the coefficients that characterize the black–white conviction rate gap when there is at least one black member seated on the jury are almost exactly the same size as the estimated impact of having at least one black potential juror in the pool."). It means that it remains possible to do something, somehow, that makes an all-white group of petit jurors deliberate in a way that produces verdicts that mirror the outcomes associated with more diverse juries. *Accord* Jerry Kang et al., *Implicit Bias in the Courtroom*, 59 UCLA L. REV. 1124, 1143 (2012) ("When the case is racially charged, jurors—who want to be fair—respond by being more careful and thoughtful about race and their own assumptions and thus do not show bias in their deliberations and outcomes."). So, although our courts may not be able to control which eligible residents are drawn by random selection (or which of them choose to appear), that does not mean all criminal trials are hopelessly, inevitably unfair.

Lilly's jury panel agreed that racism still existed and that their dormant implicit biases should not affect their verdict. *See* TrialTr.V1 158:18–166:2. They received cautionary instructions on implicit bias, throughout the trial. *See* TrialTr.V1 182:24–183:4; Jury Instr. 11A; App. 62. Because of variance in random selection, Lilly's petit jury happened to contain only eligible jurors who belonged to groups that comprised the other 98% to 99% of North Lee County—but this was still a fair trial. There is no need to contort the analysis to grant relief.

Lilly failed to prove his claim, and this Court should affirm.

REQUEST FOR NONORAL SUBMISSION

If retained, this case should be set for oral argument.

Respectfully submitted,

THOMAS J. MILLER
Attorney General of Iowa

LOUIS S. SLOVEN
Assistant Attorney General
Hoover State Office Bldg., 2nd Fl.
Des Moines, Iowa 50319
(515) 281-5976
louie.sloven@ag.iowa.gov

CERTIFICATE OF COMPLIANCE

This brief complies with the typeface requirements and type-volume limitation of Iowa Rs. App. P. 6.903(1)(d) and 6.903(1)(g)(1) or (2) because:

- This brief has been prepared in a proportionally spaced typeface using Georgia in size 14 and contains **12,150** words, excluding the parts of the brief exempted by Iowa R. App. P. 6.903(1)(g)(1).

Dated: November 19, 2020

LOUIS S. SLOVEN
Assistant Attorney General
Hoover State Office Bldg., 2nd Fl.
Des Moines, Iowa 50319
(515) 281-5976
louie.sloven@ag.iowa.gov