

2001

Look Who's Extrapolating: A Reply to Hoffmann

Valerie West

Jeffery Fagan

Columbia Law School, jfagan@law.columbia.edu

James S. Liebman

Columbia Law School, jl Liebman@law.columbia.edu

Follow this and additional works at: https://scholarship.law.columbia.edu/faculty_scholarship



Part of the [Criminal Law Commons](#), [Criminal Procedure Commons](#), and the [Legal Writing and Research Commons](#)

Recommended Citation

Valerie West, Jeffery Fagan & James S. Liebman, *Look Who's Extrapolating: A Reply to Hoffmann*, 76 IND. L. J. 951 (2001).

Available at: https://scholarship.law.columbia.edu/faculty_scholarship/294

This Article is brought to you for free and open access by the Faculty Publications at Scholarship Archive. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of Scholarship Archive. For more information, please contact scholarshiparchive@law.columbia.edu.

Look Who's Extrapolating: A Reply to Hoffmann[†]

VALERIE WEST*
JEFFREY FAGAN**
JAMES S. LIEBMAN***

In late March, a reporter called with news of a pirated copy of Professor Joseph Hoffmann's soon-to-be-published "attack" on our study, *A Broken System: Error Rates in Capital Cases, 1973-1995*.¹ Did we care to comment?² Obtaining our own copy revealed that Professor Hoffmann's fusillade missed its mark (he misstates what we did) and boomeranged (his mischaracterizations of our analysis accurately describe his own). We do care to comment, and Hoffmann and the *Indiana Law Journal* have graciously let us do so.

Hoffmann's main claim is that we "extrapolated" the 68% rate of reversible error we reported for capital verdicts reviewed during the twenty-three-year study period.³ We did no such thing. But he did. His 40% counterstatistic is an "extrapolation" in the purest sense, which proceeds from deflationary assumptions that are factually wrong.

Every year nearly 300 death sentences roll off the production lines in states with the penalty.⁴ Most are inspected on state direct appeal, state postconviction review, and federal habeas.⁵ On average, it takes about twelve years for death sentences to pass all three inspections—if they do *pass*.⁶ Most don't.⁷ When so many capital cases are awaiting final inspection at any given time, how should error rates be calculated?

Consider an analogous production line. Suppose, hypothetically, Ford Explorers made at a particular plant are said to pose an unreasonable risk of injury or death. Investigators want to know how often Explorers fail the plant's three-stage safety inspection, requiring them to be reworked or scrapped. The plant made 5760

† Copyright 2001 by Valerie West, Jeffrey Fagan, and James S. Liebman.

* Ph.D. Candidate, Department of Sociology, New York University.

** Professor of Law, Columbia Law School, and Professor of Public Health, Joseph L. Mailman School of Public Health, Columbia University.

*** Simon H. Rifkind Professor of Law, Columbia Law School.

1. James S. Liebman, Jeffrey Fagan & Valerie West, *A Broken System: Error Rates in Capital Cases, 1973-1995*, at http://www.law.columbia.edu/instructionalservices/liebman/liebman_final.pdf (June 2000) [hereinafter *A Broken System*], reprinted in part in James S. Liebman et al., *Capital Attrition: Error Rates in Capital Cases, 1973-1995*, 78 TEX. L. REV. 1839 (2000) [hereinafter *Capital Attrition*].

2. See Ann Woolner, *How Much Error Is Tolerable in Death Cases*, Bloomberg News Service, at http://quote.bloomberg.com/fgcgi.cgi?ptitle=Ann%20Woolner&touch=1&s1=woolner&tp=ad_topright_bbco&T=markets_fgcgi_content99.ht&s2=blk&bt=blk&s=AOruLRBWLSG93IE11 (Mar. 23, 2001).

3. See Joseph L. Hoffmann, *Violence and the Truth*, 76 IND. L.J. 939, 945-46 (2001).

4. See James S. Liebman, *The Overproduction of Death*, 100 COLUM. L. REV. 2030, 2065 (2000) [hereinafter *Overproduction of Death*].

5. See *A Broken System*, supra note 1, at 19-22.

6. See BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, BULLETIN: CAPITAL PUNISHMENT 1999 (2000), <http://www.ojp.usdoj.gov/bjs/abstract/cp99.htm>.

7. See *Capital Attrition*, supra note 1, at 1851.

Explorers over twenty-three months. 1182 are awaiting a first inspection. The remaining 4578 got that inspection, which 41% (1885) flunked and were sent back. Of the remaining 2693 (4578-1885) getting a second inspection, 10% (255) failed. Of the remaining 2438, 599 received a third inspection (1839 are still awaiting review), and 40% (237) flunked. What is the plant's error rate?

That rate is a composite of the error rates revealed at each inspection stage during the relevant period. For every hundred vehicles actually inspected at the first stage, forty-one failed and fifty-nine passed. For every fifty-nine inspected at the second stage, an additional 10%, or six, failed and fifty-three, passed. For every fifty-three reviewed at the third and final stage, 40%, or twenty-one failed, and thirty-two passed. So, for every hundred vehicles actually going through the inspection process and either being rejected or passed on to market, sixty-eight were rejected and thirty-two were approved: a 68% error rate. This error rate (1) uses all available information, counting the full results of every inspection that actually occurred, and (2) does not extrapolate, making no assumption about inspections that have not yet occurred.

If you were an anti-SUV crusader, you might propose a different rate, namely; the proportion of vehicles failing all *final* inspections. Here, the number inspected is 1885 (failing the first inspection and being sent back) + 255 (failing the second inspection) + 599 (which received a third and final inspection) = 2739. The number failing is 1885 + 255 + 237 (the last is the number failing the third inspection) = 2377. The fail rate is 2377/2739, or 80%. The problem with this approach is that it doesn't use all available information about actual inspection results. By assuming a 100% failure rate at both the first inspection (1885/1885) and the second (255/255), it misleadingly ignores the actual results of those inspections—respectively, a 40% and 10% failure rate. (This approach does provide useful information: the ratio of vehicles entering the inspection process that made it to market. But that is not an error rate.)

If you were the plant manager, with your job on the line, you might try to fudge in the opposite direction, by computing the ratio of all vehicles the plant made (5760) that failed inspection (1885 + 255 + 237 = 2377), about 40% (2377/5760). The problem here is revealed by what the 40% rate says about how many vehicles *passed* inspection: 3383 (5760-2377), or about 60%. That number is the sum of a thousand-plus vehicles that *did not even reach the first inspection*, plus nearly 2000 vehicles that *failed to complete the other two inspections*, plus 362 vehicles that passed all three. The 40% rate thus is an extrapolation based on an assumption that thousands of vehicles passed (and *none* failed) a plethora of inspections that *never occurred*. Moreover, assuming a 100% *pass* rate for uninspected vehicles is obviously absurd, given how many vehicles made at the same plant in the same period *were* inspected and *flunked*—at rates of 41%, 10%, and 40% in the successive stages, and 68% overall. Why make a grossly inaccurate and deflationary assumption unless there is something to hide? (And what good is hiding when, even after faulty extrapolation, the fail rate is still 40%?)

This Ford plant is made up. But the numbers aren't. They are the actual, total number of death verdicts pronounced in the United States during our twenty-three-year study period and the number reversed and sent back at each stage of the review

process.⁸ Contrary to Hoffmann's claim, they are *not* "samples." They are the entire universe of cases during the period.

The three proposed reversal rates also are real. 88% is the error rate that critics of the death penalty compute from our study.⁹ As the analysis above indicates, that rate is suspect because it suppresses the crucial information that some cases awaiting *later* stages of review were *affirmed* at earlier stages. To that extent, the measure is misleading, and we have rejected it.

The 40% proposed rate is Hoffmann's.¹⁰ The problem with this figure is that it is an extrapolation based on two assumptions: (1) that thousands of court reviews that in fact have *not* occurred, *have* occurred and (2) that all resulted in *affirmances*. The first assumption is false. The second assumption—a 0% reversal rate for not-yet-completed court reviews—is (with all due respect) absurd: The same appellate courts reviewing thousands of death sentences imposed by the same trial courts in the same years had a 41% reversal rate at the first stage, a 10% reversal rate for the remaining cases reviewed at the second stage, and a 40% reversal rate for the remaining cases reviewed at the third stage, for an overall error rate of 68%.¹¹ Because Hoffmann's figure is an estimate based on plainly wrong assumptions about undecided cases, we rejected it. (It does show, however, that even with absurdly deflationary assumptions, the error rate is still depressingly high.)

The error rate is 68%. That rate counts *all* actual outcomes at *all* stages (no sampling; no data suppression), and *only* actual outcomes (no extrapolation), of court decisions reviewing every death sentence imposed and inspected during the twenty-three-year study period. It accounts for the condition that Hoffmann correctly notes must be the lynchpin of any method of calculating error—that "during the study period many of the cases . . . had been reviewed only on direct appeal, but not on state postconviction or federal habeas"¹²—but that he then ignores in his own calculation.¹³

8. See *A Broken System*, *supra* note 1, at 29-33.

9. Telephone interview with Richard Dieter, Director, Death Penalty Information Center (Sept. 6, 2000).

10. Hoffmann, *supra* note 3, at 946.

11. See *A Broken System*, *supra* note 1, at 29-33.

12. Hoffmann, *supra* note 3, at 945.

13. In his rejoinder, Hoffmann admits our 68% figure shows what *actually* happened to verdicts that were finally reviewed during the study period. See Joseph L. Hoffmann, *A Brief Response to Liebman, Fagan, and West*, 76 IND. L.J. 957, 957 (2001). This leads him to redefine his 40% figure, not as an alternative to our 68%, but as an admitted extrapolation from the answer to a counterfactual question about the many 1973-95 verdicts that got stuck in the system and were *not* reviewed in the period: "How many of the as-yet-unreviewed cases *would be reversed*, if they were completely reviewed?" *Id.* (emphasis in original). Noting that those results cannot actually be counted, Hoffmann makes an avowedly "absurd" guess—that *all* "as-yet-unreviewed" cases "would be" *affirmed*. See *id.* Then adding his made-up 0% reversal rate for undecided cases to the actual 68% rate for decided cases, he conjures up a 40% rate for the two sets of cases combined and claims that the combined rate is between 40% and 68%. See *id.* But if you were coaching the Indiana Pacers basketball team, which, say, had missed 68% of its free throws in the regular season, would you plan for the playoffs based on the avowedly absurd assumption that the team would *make* all its "as-yet-unattempted" foul shots? Or, would you plan for the team's miss rate to be about what it was in the past, give or take a little? And, if projecting a 40% miss rate for the entire season ("depressing" as that still admittedly would

Hoffmann's second criticism also misdescribes *A Broken System*. Our side study of the reasons for and retrial results of state postconviction reversals found that (1) 80% of the violations based on which relief was granted required proof of at least a reasonable probability that, but for the violation, the substantive outcome at the guilt or sentencing phase would have been different and (2) 82% of those reversals led, on retrial, to an outcome less than death, including 7% leading to acquittals.¹⁴ *A Broken System* nowhere claims that these bases for relief and retrial results are anything other than what they are: the reasons for, and retrial outcomes of, all actual reversals at the state postconviction stage during the study period. Hoffmann admits this is clear "[i]f you read the study carefully."¹⁵ It also is clear if you just read the study.¹⁶

Hoffmann is concerned that this finding not be extended to the other review stages where, he assumes—based on *no* data, a sample of zero—that reversals are for merely "technical" violations.¹⁷ We will soon publish findings that the bases for relief at the third, federal habeas, stage are very like those at the state postconviction stage, belying Hoffmann's "technicalities" assumption as to habeas.

We also question that assumption when applied to state direct appeal, where data are not available. Why assume elected state supreme court justices reverse capital verdicts for *less* serious, *more* frivolous reasons than unelected federal habeas judges? Why assume state supreme courts ignore the harmless error rule barring reversals based on violations with no effect on the substantive outcome?¹⁸ All available evidence suggests those courts *overuse* the rule to *deny* relief¹⁹—as in the Anthony Porter miscarriage that Hoffmann himself decries.²⁰ Is it sensible to prefer

be) is too reckless when it's just a game at stake, how helpful is it when life is in the balance?

14. See *Capital Attrition*, *supra* note 1, at 1850, 1852.

15. Hoffmann, *supra* note 3, at 946.

16. See, e.g., *A Broken System*, *supra* note 1, at ii (discussing "[t]he most common errors—prompting a majority of reversals at the *state post-conviction stage*," and reporting that "82% of the people whose capital judgments were overturned by *state post-conviction courts* due to serious error were found to deserve a sentence less than death when the errors were cured on retrial; 7% were found to be innocent of the capital crime" (emphasis added, other emphasis omitted)); *id.* at 5 (same); *Capital Attrition*, *supra* note 1, at 1850, 1852.

17. See Hoffmann, *supra* note 3, at 947.

18. See, e.g., *Chapman v. California*, 386 U.S. 18 (1967) (barring reversal of conviction in death-penalty case even though prosecutor's argument and trial judge's jury instructions unconstitutionally directed the jury to draw in favor of prosecution all inferences from defendants' failure to testify if the state could show beyond a reasonable doubt that the "constitutionally forbidden statements" did not influence defendants' convictions).

19. See *Overproduction of Death*, *supra* note 4, at 2128-29 & n.239 (listing sources).

20. Despite Anthony Porter's innocence, later proven by intrepid undergraduate students as a journalism project, see Hoffmann, *supra* note 3, at 948 & n.48, a full complement of state and federal courts had upheld his conviction and death sentence, holding that violations that may have occurred at his trial should be overlooked, as harmless and nonprejudicial, because "the evidence against the defendant was overwhelming." *Illinois v. Porter*, 489 N.E.2d 1329, 1337 (Ill. 1986) (denying direct appeal), *cert. denied*, 479 U.S. 898 (1986); see also *Illinois v. Porter*, 647 N.E.2d 872, 975 (Ill. 1995) (affirming denial of state postconviction relief, although Porter's trial lawyer was incompetent, because the "considerable" evidence against Porter prevented him from showing "prejudice"); *United States ex rel. Porter v. Warden Pontiac Correctional Ctr.*, No. 95-C-4111, 1996 WL 167340 (N.D. Ill. Apr. 4, 1996) (same),

postconviction review, which Hoffmann concedes finds truly serious error, to state direct appeal on the assumed ground that state supreme court justices reverse for the heck of it? Why not instead assume that the state supreme court decisions analyzed in our state postconviction study fairly suggest how clear and serious error must be before the *same* courts will reverse capital verdicts on direct appeal? Or why not assume, as is usually true, that initial inspections find the worst, most glaring errors, leaving more subtle flaws for later inspections? Where everything we actually know—including from our state postconviction study—belies Hoffmann's assumption that state supreme courts routinely reverse capital verdicts on "technicalities," isn't it fair to ask for some representative examples? Hoffmann gives none.²¹

Finally, Hoffmann complains that our study "identified only twenty-two cases nationwide" in which innocent men and women on death row were subsequently acquitted, and that "many more cases involved a defendant who was guilty of the capital crime, but was [merely?] undeserving of the death penalty."²² In fact, since the death penalty was reinstated, ninety-five death-row inmates have been acquitted—just

aff'd sub nom. Porter v. Gramley, 112 F.3d 1308 (7th Cir. 1997), *reh'g denied*, 122 F.3d 351 (7th Cir. 1997), *cert. denied*, 522 U.S. 1093 (1998), *reh'g denied sub nom.* Porter v. Gilmore, 523 U.S. 1042 (1998).

21. Hoffmann claims that some reversible errors are too piddling to bear on the reliability of the conviction or death sentence. We asked him to cite some real reversals that illustrate his point. He doesn't. Instead, his rejoinder poses a hypothetical "loosely based on the real case of *Mills v. Maryland*, in which the U.S. Supreme Court invalidated the jury instructions . . . that were routinely used with respect to mitigating circumstances in Maryland capital cases, based on the existence of a 'substantial risk' . . . that the jury . . . might have been . . . misled." Hoffmann, *supra* note 13, at 958 n.8 (citation omitted) (quoting *Mills v. Maryland*, 486 U.S. 367, 381 (1988)). "Loose[]" is right. Hoffmann assumes that "ninety out of a hundred cases would have come out the same way even without the challenged instruction" because "most of the . . . defendants . . . were . . . deserving of a death sentence." *Id.* We, again, have *actual* data on retrial outcomes at the state postconviction stage, where most *Mills* reversals probably occurred. In fact, 75% of the inmates whose death sentences were reversed based on *Mills* at that stage were found on retrial before a properly instructed jury (or in one case on clemency) to be *undeserving* of a death sentence. See *A Broken System*, *supra* note 1, at C-29 to C-30. Three-fourths of the time, the error meant the difference between life and death—hardly piddling.

Looser still is Hoffmann's assumption that courts found *reversible* error—the only error serious enough to be counted in our study—in all seventy or hundred cases the press initially thought *Mills* might affect. He forgets that courts often *fail to grant* relief on *Mills* claims for reasons that show why *reversible* error (the kind found in 68% of the death verdicts reviewed during the study period) is *serious* error—error whose cure on retrial usually changes the outcome. Some *Mills* claims were not preserved at trial or on appeal; others were mooted by reversals due to different violations; and still others were found nonprejudicial because the defendant could not show a "substantial risk" that the bad instruction misled the jury.

Hoffmann accepts our underlying policy conclusions but quibbles with our numbers. He believes precision matters. It does. That's why we prefer actual counts of real events to demonstrably inaccurate assumptions and hypotheticals.

22. Hoffmann, *supra* note 3, at 947.

under one for every seven executed.²³ More fundamentally, Hoffmann's indifference to faulty verdicts sentencing people to die for acts for which death is not a legal punishment, as long as they committed *some* crime, is contrary to the usual assumption that unreliable death sentences are matters of utmost concern.²⁴ The defects in Professor Hoffmann's analysis are illustrated by putting his words in the mouth of our hypothetical plant manager:

So what if our inspectors found 68% of our cars too flawed to let on the road? And what if 82% of the cars flunking our second inspection had to be scrapped instead of fixed? Only twenty-two (or was it ninety-five?) actually risked killing test drivers. Many others only endangered people who break the law. And some of the rest (sorry, I can't give examples) are due to overregulating and overlitigating, which is why we pay inspectors to reject cars on technicalities. As for the 82% scrap rate, that was only for a small set of cars where we know what happened. (Don't ask me why I think other rejected cars do better.) And they were scrapped because our supervisors thought they weren't worth fixing. Why trust them?²⁵

In the end, Hoffmann agrees that "the current system remains plagued by examples of overzealous police and prosecutors, inadequate defense lawyers, and strained resources," that "substantively unjust outcomes [occur] in . . . too many" cases, and that "prosecutors . . . [should] join with defense attorneys in searching for, and correcting, such error."²⁶ But despite this common ground, it is important that Hoffmann not (to paraphrase him) "greatly [under]stat[e] the statistical case about substantive injustice in death-penalty cases."²⁷

23. Death Penalty Info. Ctr., *Innocence and the Death Penalty*, at <http://www.deathpenaltyinfo.org/innoc.html> (last visited May 14, 2001).

24. See, e.g., *Herrera v. Collins*, 506 U.S. 390, 405 (1993) ("We have . . . held that the Eighth Amendment requires increased reliability of the process by which capital punishment may be imposed." (citing *McKoy v. North Carolina*, 494 U.S. 433 (1990), *Eddings v. Oklahoma*, 455 U.S. 104 (1982), and *Lockett v. Ohio*, 438 U.S. 586 (1978))); see also James S. Liebman, Jeffrey Fagan & Valerie West, *Death Matters: A Reply to Latzer and Cauthen*, 84 JUDICATURE 72 (2000); Jeffrey Fagan, James S. Liebman & Valerie West, *Death Is the Whole Ball Game*, 84 JUDICATURE 144 (2000).

25. See Hoffmann, *supra* note 3, at 946-47.

26. *Id.* at 948, 949.

27. *Id.* at 948.