2005

Causing Constitutional Harm: How Tort Law Can Help Determine Harmless Error in Criminal Trials

Jason M. Solomon

Printed in U.S.A.
Vol. 99, No. 3

CAUSING CONSTITUTIONAL HARM: HOW TORT LAW CAN HELP DETERMINE HARMLESS ERROR IN CRIMINAL TRIALS

*Jason M. Solomon**

I. INTRODUCTION

Several years before Todd Brecht improbably found himself before the Supreme Court, he was the protagonist in a human tragedy. Just out of a

* Chief of Staff to the President, Harvard University. For helpful comments at earlier stages of this project, I would like to thank Tom Eaton, Cindy Estlund, Bob Iuliano, Christine Jolls, Dori Kaplan, Jonathan Sacks, Bill Stuntz, and Rebecca White. Thanks also to the editors and staff of the Northwestern University Law Review for superb editing and other assistance.

Georgia prison, Brecht was staying with his sister, Molly Hartman, in the home she shared with her husband, the local district attorney, in a small town in western Wisconsin. Roger Hartman didn't approve of his brother-in-law's drinking, sexual orientation, and criminal record. In retrospect, it was a recipe for disaster.

When Todd Brecht was tried for the murder of his sister's husband, he took the stand and argued that his shooting of a rifle into Roger Hartman's back was an accident. In arguing that this version of events was incredible, the prosecutor told the jury several times that Brecht had never told anyone this explanation before his testimony at trial. But to the extent that this referred to Brecht's post-arrest silence, after receiving *Miranda* warnings, the prosecutor's comments to the jury violated due process. Brecht was convicted; on appeal, he claimed that this constitutional violation justified a new trial.

When the Supreme Court ruled on his claim in 1993, eight years after the shooting, it was the fifth court to answer the question whether the constitutional violation had sufficiently impacted the trial to require reversal of the murder conviction, or whether the error was "harmless." Each court had reversed the one before it.[1] The Supreme Court upheld Brecht's conviction, in an opinion agreeing with the analysis of Seventh Circuit Judge Frank Easterbrook below.[2] But the questions surrounding "the riddle of harmless error"[3] continued, beyond the case of Todd Brecht and a prosecutor whose constitutional missteps were forgiven at a trial in which he was no doubt singularly focused on punishing his boss's killer.

* * *

The harmless-error rule has been called "probably the most cited rule in modern criminal appeals."[4] The Chief Judge of the Second Circuit has referred to harmless error as "one of the most important doctrines in appellate decisionmaking," and posited that harmless-error principles may determine the outcome of more criminal appeals than any other doctrine.[5] A

---

[1] State v. Brecht, 405 N.W.2d 718 (Wis. Ct. App. 1987) (reversing conviction), *rev'd*, 421 N.W.2d 96 (Wis. 1988); Brecht v. Abrahamson, 759 F. Supp. 500 (W.D. Wis.) (granting writ), *rev'd*, 944 F.2d 1363 (7th Cir. 1991).

[2] Brecht v. Abrahamson, 507 U.S. 619 (1993).

[3] This phrase is taken from the title of ROGER J. TRAYNOR, THE RIDDLE OF HARMLESS ERROR (1970).

[4] William M. Landes & Richard A. Posner, *Harmless Error*, 30 J. LEGAL STUD. 161, 161 (2001).

[5] Hon. John M. Walker, Jr., *Harmless Error Review in the Second Circuit*, 63 BROOK. L. REV. 395, 395 (1997). Chief Judge Walker also quoted one of his predecessors, former Chief Judge Jon Newman, for the related point that "assessment of harmlessness is probably the single most recurring issue presented" in habeas corpus challenges to state court convictions. *Id.* at n.2 (quoting Peck v. United States, 102 F.3d 1319, 1327 (2d Cir. 1996) (en banc) (Newman, J., concurring)). Indeed, two of the most prominent judges in the history of the Second Circuit and of twentieth century American law, Learned

leading treatise on federal appellate practice calls harmless error "probably the most far-reaching doctrinal change in American procedural jurisprudence since its inception."[6]

In 1970, in what remains the seminal work on harmless error, Roger Traynor—then Chief Justice of the California Supreme Court—pointed to the "obvious need of guidelines to control appellate discretion in the evaluation of error."[7] But in the wake of the decision in Todd Brecht's case, *Brecht v. Abrahamson*, harmless-error doctrine remains a bit of a mess. Thirty-five years after Justice Traynor identified the problem, the significance of the doctrine has grown, while the need for further guidelines remains strong.

It is difficult to overstate the stakes. The criminal justice system has been shaken by increasing evidence of wrongful convictions, and most of the constitutional errors being assessed in harmless-error analysis are central to the truth-determining process of criminal trials. Getting harmless-error determinations right, then, is central to accurate determinations of guilt—an area about which there is evident cause for concern.

The conventional wisdom on harmless-error doctrine is that there are two different and irreconcilable approaches that judges use in determining harmless error which are reflected in two coexisting lines of Supreme Court cases. Much of the scholarship on harmless error, as well as the ongoing debate within the Supreme Court, focuses on the difference in these two approaches. But this debate overstates the difference between the two approaches and obscures the shared normative ideal at the heart of harmless-error doctrine. Contrary to conventional assumptions underlying debates on harmless error, I argue that by using tort law, these two approaches can be reconciled in a way that increases the overall accuracy of harmless-error analysis.

This Article proceeds in four parts. Part II is a brief overview of harmless-error doctrine in the context of habeas challenges to state criminal convictions, focusing on the nature of the inquiry and the doctrinal deadlock described above. Part III is an empirical analysis of the post-*Brecht* cases in the federal courts of appeals. To search for a way out of the doctrinal deadlock, I started with a relatively straightforward question: what has happened to harmless-error analysis since *Brecht*? To answer this question, I reviewed and, with the help of a research assistant, coded all of the 315 harmless-error analyses on habeas review in published opinions over the last decade. Despite the different approaches, a common thread emerges:

Hand and Jerome Frank, sparred over the issue of harmless error in a series of cases in the 1940s. *See* Charles S. Chapel, *The Irony of Harmless Error*, 51 OKLA. L. REV. 501, 522–23 & nn.123–24 (1998) (citing and summarizing these cases).

[6] 2 STEVEN A. CHILDRESS & MARTHA S. DAVIS, FEDERAL STANDARDS OF REVIEW § 7.03, at 7 (2d ed.1986).

[7] TRAYNOR, *supra* note 3, at 15.

the language and logic of causation is everywhere, as courts struggle to assess the causal impact of an error on the jury's verdict. Contrary to the conventional view, judges using different approaches to harmless-error are actually trying to answer the same fundamental question: did the error *cause* the conviction?

Part IV proposes reconceptualizing harmless-error analysis as a determination of causation in a constitutional tort claim and using this reconception to provide a way out of the doctrinal morass. By turning to tort-law debates about what it means to cause harm, I grapple with the question: what does it mean for an error to *cause* a conviction? Indeed, it appears that different conceptions of causation—described in tort law as the "but for" versus the "substantial factor" tests—account for many of the differences in harmless-error outcomes in the federal courts. I argue that a hybrid approach can better serve the normative ideal of determining factual causation at criminal trials while avoiding appellate fact-finding that violates the Sixth Amendment guarantee of a trial by jury.

In Part V, I turn to prescription: how should judges implement this proposed reconception of harmless error? First, I observe the difference between the harmless-error cases and the social science literature with respect to the dominant model of how jurors use evidence to reach a verdict. I argue that in the absence of direct evidence about the impact of errors on particular jurors, drawing inferences from empirical research on juries as part of the harmless-error analysis is both appropriate and desirable. And in focusing on the issue of causation, judges and lawyers should turn their attention to evidence of influence on the jury, as opposed to simply weighing the evidence of guilt.

II. THE STILL UNSOLVED RIDDLE OF HARMLESS ERROR: DOCTRINAL DEADLOCK, UNGUIDED DISCRETION

> "The expansive code of constitutional criminal procedure that the Supreme Court has created in the name of the Constitution is like the grapes of Tantalus, since the equally expansive harmless error rule in most cases prevents a criminal defendant from obtaining any benefit from the code."[8]
>
> — Judge Richard Posner

A. Harmless Error on Habeas: Where the Rubber Hits the Road in Constitutional Criminal Procedure

The Bill of Rights reflects a number of values, including that of furthering the truth—in this context, convicting the guilty and acquitting the innocent. But our society's confidence in the criminal justice system and its

[8] United States v. Pallais, 921 F.2d 684, 692 (7th Cir. 1990) (Posner, J.).

truth-furthering safeguards has been shaken in the decade since *Brecht* was decided. DNA and other forms of forensic evidence have led to the overturning of scores of convictions in recent years, a rethinking of the death penalty in many states, and a renewed focus on police interrogations, eyewitness identification, and other staples of criminal prosecution.[9] The exoneration of innocent people who have served years in prison continues, and while many have been released, others no doubt remain imprisoned while the actual perpetrators go free.[10] The overwhelming majority of the exonerees were wrongfully convicted at state criminal trials.

For institutional and historical reasons, many have long questioned state courts' commitment to the protection of federal procedural safeguards in criminal cases.[11] As a result, Congress passed, just after the Civil War, the Habeas Corpus Act of 1867 to provide a federal forum for state prisoners, and more recently, the writ of habeas corpus has served as a general post-conviction remedy for state convictions tainted by constitutional violations.[12] These constitutional violations are important, for the Rehnquist Court at least, because they indicate that the truth-furthering function of constitutional criminal procedure has been hampered, increasing the risk of inaccurate determinations of guilt.[13]

In many ways, federal habeas petitions are a crapshoot.[14] The overwhelming majority of petitions are filed without counsel, but studies have demonstrated that "the availability of professional representation is the single most important predictor of success in federal habeas corpus."[15] Be-

---

[9] For a summary of some of these developments, see Steven A. Drizin & Richard A. Leo, *The Problem of False Confessions in the Post-DNA World*, 82 N.C. L. REV. 891 (2004); Daniel Givelber, *Meaningless Acquittals, Meaningful Convictions: Do We Reliably Acquit the Innocent?*, 49 RUTGERS L. REV. 1317 (1997); Samuel R. Gross, *The Risks of Death: Why Erroneous Convictions Are Common in Capital Cases*, 44 BUFF. L. REV. 469 (1996).

[10] For an ongoing account of exonerations around the country, see http://www.innocenceproject.org.

[11] This suspicion, of course, is what led to the passage of the law extending the scope of federal habeas corpus to state prisoners in the aftermath of the Civil War. *See* Fay v. Noia, 372 U.S. 391, 415–18 (1963); Daniel J. Meltzer, *Harmless Error and Constitutional Remedies*, 61 U. CHI. L. REV. 1, 28 (1994) ("The existence of federal habeas corpus jurisdiction itself reflects doubts that state courts, left to their own devices, would adequately enforce federal constitutional norms."); Larry W. Yackle, *Explaining Habeas Corpus*, 60 N.Y.U. L. REV. 991, 1031–32 (1985).

[12] *See* Yackle, *supra* note 11, at 1031.

[13] *See* Tom Stacy, *The Search for the Truth in Constitutional Criminal Procedure*, 91 COLUM. L. REV. 1369, 1372 (1991).

[14] Of the federal habeas petitions filed, it is estimated that between seventy and eighty percent are state prisoners who were convicted after trial. *See* Daniel J. Meltzer, *Habeas Corpus Jurisdiction: The Limits of Models*, 66 S. CAL. L. REV. 2507, 2524 & n.105 (1993). During 2000, 17 out of every 1000 state inmates filed habeas corpus petitions. John Scalia, *Prisoner Petitions Filed in U.S. District Courts, 2000, with Trends 1980–2000*, U.S. Dep't of Justice, Bureau of Justice Statistics, at 2, *available at* http://www.ojp.usdoj.gov/bjs/pub/pdf/ppfusd00.pdf (last modified Feb. 5, 2002).

[15] Richard Faust et al., *The Great Writ in Action: Empirical Light on the Federal Habeas Corpus Debate*, 18 N.Y.U. REV. L. & SOC. CHANGE 637, 707 (1991). Undoubtedly, the potential strength of the

cause of the stakes involved for the defendant (usually a long prison sentence or the death penalty) and the length of time it takes to get habeas petitions adjudicated, most habeas cases involve people convicted after being tried for murder.[16] Although less than 5% of state felony cases are resolved by trial each year,[17] one out of three defendants facing murder charges go to trial, compared to no more than 10% of defendants charged with other offenses.[18]

In a subset of these cases, a federal district or appellate judge determines that constitutional error occurred during the process that led to the petitioner's incarceration. In such cases, the truth-furthering safeguards were impaired, providing an indicator that the petitioner may have been innocent. The question then becomes whether the court should remedy error by overturning the conviction. Here the "harmless error" doctrine kicks in, and the rubber hits the road for constitutional criminal procedure and the accurate determination of guilt.

In this Article, I focus on these important cases in which a judge finds that constitutional error has infected the trial and must determine whether the error is harmless. In conducting harmless-error analysis, the Supreme Court has emphasized that "the central purpose of a criminal trial is to decide the factual question of the defendant's guilt or innocence . . . ."[19] This focus on the truth-determining function of the criminal trial animates the Court's harmless error jurisprudence.[20] Despite other possible reasons for remedying constitutional violations at criminal trials, judges will only remedy the constitutional violation if they believe it has impacted the jury's determination of guilt or innocence. The success of harmless-error analysis,

---

claims is related to the defendant's ability to get appointed counsel by the court, or to otherwise get a lawyer to take his case, but it is doubtful that the relationship is particularly robust.

[16] *See* Roger A. Hanson & Henry W.K. Daley, *Federal Habeas Corpus Review, Challenging State Court Criminal Convictions*, U.S. Dep't of Justice, Bureau of Justice Statistics, at v (Sept. 1995), *available at* http://www.ojp.usdoj.gov/bjs/pub/pdf/fhcrcscc.pdf (reviewing nearly half of the federal habeas petitions adjudicated in 1992 and finding that most involved challenges for prisoners "primarily convicted of violent offenses and given correspondingly severe sentences"); *see also id.* at 12 (noting that most offenders convicted of felonies are sentenced to five years or less, but the time required to exhaust state remedies takes on average almost five years).

[17] Approximately one in three cases ended with a dismissal, and 65% of felony cases were resolved by guilty pleas. *See* Brian Ostrom et al., *Examining the Work of State Courts*, Nat'l Ctr. for State Courts 44 (2003), *available at* http://www.ncsconline.org/D_Research/csp/2003_Files/2003_Main_Page.html.

[18] Gerard Rainville & Brian A. Reaves, *Felony Defendants in Large Urban Counties, 2000*, U.S. Dep't of Justice, Bureau of Justice Statistics, at 26 (Dec. 2003), *available at* http://www.ojp.usdoj.gov/bjs/pub/pdf/fdluc00.pdf. In 2000, an estimated 4% of the cases adjudicated within one year went to trial. Fifty-seven percent of these trials were bench trials, and 43% were jury trials. An estimated 78% of all trials ended with a guilty verdict, and 22% with an acquittal. Bench trials (81%) were more likely to result in a conviction than jury trials (74%). *Id.*

[19] Rose v. Clark, 478 U.S. 570, 577 (1986) (quoting Delaware v. Van Ardsall, 475 U.S. 673, 681 (1986)).

[20] Tom Stacy & Kim Dayton, *Rethinking Harmless Constitutional Error*, 88 COLUM. L. REV. 79, 80–81 (1988) (criticizing this emphasis).

then, is a critical part of maintaining a reasonably accurate criminal justice system.

In harmless-error analysis, judges are asked to determine the impact that the error—whether it be improperly admitted evidence, improperly excluded evidence, or faulty jury instructions—had on the jury's verdict. As Justice Traynor put it, review for harmless error "requires the most painstaking examination of the record and the most perceptive reflections as to the probabilities of the effect of error on a reasonable trier of fact."[21]

B. The Doctrinal Backdrop—and Current Deadlock—on Harmless Error

Doctrinally, the nature and source of harmless-error analysis is less than clear. The statute authorizing federal habeas jurisdiction over state court prisoners[22] says nothing about harmless error, although the Supreme Court has turned in part to the statute governing harmless error in federal criminal appeals, and the case law surrounding it, to interpret harmless error in the habeas context.[23] Absent a specific statutory source, however, some scholars have referred to the doctrine as "constitutional common law."[24] But, since harmless error does not address whether a constitutional violation actually occurred, other scholars consider harmless error a part of the law of remedies.[25] We are left wondering whether harmless error is part of the law of constitutional criminal procedure, remedies, both or neither. This matters because interpretive questions arise constantly, with the need to turn to some purpose or background doctrine.

From 1867 to 1967, courts applied the harmless-error doctrine in criminal cases only to nonconstitutional errors. Constitutional errors were per se reversible, without evaluating the impact of the error on the jury's verdict. This changed with the Supreme Court's 1967 decision in *Chapman v. California*.[26] In *Chapman*, the Court declared that constitutional errors

---

[21] TRAYNOR, *supra* note 3, at 30; *see also* Gray v. Klauser, 282 F.3d 633, 653 (9th Cir. 2002) (observing that judges conducting harmless-error analysis "are asked to make a judgment about the likely reasoning patterns" of jurors).

[22] 28 U.S.C. § 2254 (2000).

[23] *See, e.g.*, O'Neal v. McAninch, 513 U.S. 432 (1995).

[24] Craig Goldblatt, *Harmless Error as Constitutional Common Law: Congress's Power to Reverse* Arizona v. Fulminante, 60 U. CHI. L. REV. 985, 1009–12 (1993); Henry P. Monaghan, *Harmless Error and the Valid Rule Requirement*, 1989 SUP. CT. REV. 195, 200 & n.30.

[25] Richard H. Fallon, Jr. & Daniel J. Meltzer, *New Law, Non-Retroactivity, and Constitutional Remedies*, 104 HARV. L. REV. 1733, 1770–73 (1991) (arguing that harmless-error doctrine emerges from a "balancing calculus familiar to the law of remedies: if the risks of prejudice, though not nonexistent, are small, the burdens of retrial are not warranted").

[26] 386 U.S. 18 (1967); *see also* Jeffrey O. Cooper, *Searching for Harmlessness: Method and Madness in the Supreme Court's Harmless Constitutional Error Doctrine*, 50 U. KAN. L. REV. 309, 313 (2002) (describing the development of harmless error in the twentieth century as "steady expansion"). For a comprehensive look at the history and development of harmless-error doctrine, see *id.* at 314–24; Chapel, *supra* note 5, at 515–30; Martha S. Davis, *Harmless Error in Federal Criminal and Habeas Ju-*

would not necessitate overturning a conviction if the error was "harmless beyond a reasonable doubt."[27] After *Chapman*, the use of the harmless-error doctrine increased significantly, doubling within three years in the federal courts of appeals.[28] According to one estimate, in the decade following *Chapman* approximately 10% of all criminal appellate cases throughout the country were determined by a finding of harmless constitutional error.[29] But as applied to constitutional error in criminal trials, the harmless-error doctrine has frequently been criticized.[30]

If the 1967 *Chapman* decision opened the door slightly to maintaining convictions tainted by constitutional error, the 1993 *Brecht* decision opened the door even wider. Previously, the state had been required to demonstrate that the error was "harmless beyond a reasonable doubt" to avoid a new trial or sentencing proceeding, but *Brecht* held that the state was only required to show that the error did not have a "substantial and injurious effect or influence in determining the jury's verdict."[31] In this way, the Court broadened the range of constitutional errors that would not be remedied on collateral review to beyond the merely trivial.

The five-to-four decision was controversial. Justice Stevens's concurrence provided the crucial fifth vote, and it emphasized his belief that the change in standard was "less significant than it might seem."[32] Justice O'Connor's dissent disagreed, arguing that "by tolerating a greater probability that an error with the potential to undermine verdict accuracy was harmful, the Court increases the likelihood that a conviction will be pre-

---

risprudence: The Beast That Swallowed the Constitution, 25 T. MARSHALL L. REV. 45, 46–50, 57–77 (2000).

[27] 386 U.S. at 24.

[28] *See* Donald A. Winslow, Note, *Harmful Use of Harmless Error in Criminal Cases*, 64 CORNELL L. REV. 538, 545–46 & n.36 (1979). The proportion of harmless error cases remained at approximately 2% until 1986, when the percentage dropped to 1.58% and held there through the mid-1990s. *See* Harry T. Edwards, *To Err Is Human, but Not Always Harmless: When Should Legal Error Be Tolerated?*, 70 N.Y.U. L. REV. 1167, 1180–81 & n.52 (1995).

[29] *See* Steven H. Goldberg, *Harmless Error: Constitutional Sneak Thief*, 71 J. CRIM. L. & CRIMINOLOGY 421 (1980).

[30] *See, e.g.*, David R. Dow & James Rytting, *Can Constitutional Error Be Harmless?*, 2000 UTAH L. REV. 483; Bennett L. Gershman, *The Gate Is Open but the Door Is Locked—Habeas Corpus and Harmless Error*, 51 WASH. & LEE L. REV. 115, 132 (1994) (criticizing the harmless error rule's "standardless and ad hoc application by appellate judges who purport to be making precise quantitative and qualitative calculations of the impact of errors based on the 'cold black and white of the printed record'" (quoting United States v. Grunberger, 431 F.2d 1062, 1067 (2d Cir. 1970))); Rex R. Perschbacher & Debra Lyn Bassett, *The End of Law*, 84 B.U. L. REV. 1, 39 (2004) (arguing that the harmless error doctrine is now applied to mistakes "for which the adjective 'harmless' seems highly questionable").

For the purposes of this Article, I accept the harmless-error doctrine as necessary and, more importantly, entrenched. Moreover, I accept harmless-error doctrine's premise that judicial remedies for constitutional violations at criminal trials should be limited to cases where the judge is uncertain that the determination of legal guilt was accurate.

[31] 507 U.S. 619, 638 (1993).

[32] *Id.* at 643 (Stevens, J., concurring).

served despite an error that actually affected the reliability of the trial."[33] The weight of scholarly reaction was critical, with one commentator concluding that *Brecht* "effectively locks the door to meaningful habeas review of most constitutional trial violations" and "strikes a heavy blow to effective federal oversight of state constitutional violations."[34]

After *Brecht*, what is the precise question that the judge must answer? In *Brecht*, the question was whether the error had a "substantial and injurious effect or influence in determining the jury's verdict."[35] This formulation relied in turn on *Kotteakos*, where Justice Rutledge said:

> [I]f one cannot say, with fair assurance, after pondering all that happened without stripping the erroneous action from the whole, that the judgment was not substantially swayed by the error, it is impossible to conclude that substantial rights were not affected. The inquiry cannot be merely whether it was enough to support the result, apart from the evidence affected by the error. It is rather, even so, whether the error itself had substantial influence. If so, or if one is left in grave doubt, the conviction cannot stand.[36]

But since *Brecht*, different verbal formulations have continued to proliferate in the Court. In *Sullivan v. Louisiana*, Justice Scalia explained that harmless-error review "looks . . . to the basis on which 'the jury actually rested its verdict.'"[37] He explained further: "The inquiry . . . is not whether, in a trial that occurred without the error, a guilty verdict would surely have been rendered, but whether the guilty verdict actually rendered in *this* trial was surely unattributable to the error."[38] In *O'Neal*, Justice Breyer said that the judge should ask: "Do I, the judge, think that the error substantially influenced the jury's decision?"[39] But in *Neder v. United States*, Justice Rehnquist put the nature of the inquiry somewhat differently: "Is it clear beyond a reasonable doubt that a rational jury would have found the defendant guilty absent the error?"[40]

---

[33] *Id.* at 653 (O'Connor, J., dissenting).

[34] *See* Gershman, *supra* note 30, at 132–33. Gershman pointed to the state and federal judges' different assessments of harm in *Brecht* as "a striking example of the essential absurdity of harmless error review." *Id.* at 132. Along similar lines, Liebman and Hertz observed that "it is difficult to avoid Justice O'Connor's suspicion that the Court's goal is not improvement in the administration of justice but, instead, 'denying [habeas corpus] relief whenever possible.'" James S. Liebman & Randy Hertz, Brecht v. Abrahamson*:* *Harmful Error in Habeas Corpus Law*, 84 J. CRIM. L. & CRIMINOLOGY 1109, 1156 (1994) (quoting *Brecht*, 507 U.S. at 656 (O'Connor, J., dissenting)).

[35] Kotteakos v. United States, 328 U.S. 750, 776 (1946).

[36] *Id.* at 764–65.

[37] 508 U.S. 275, 279 (1993) (quoting Yates v. Evatt, 500 U.S. 321, 404 (1991)).

[38] *Id.*

[39] 513 U.S. 432, 436 (1995).

[40] 527 U.S. 1, 18 (1999). Though this case was a direct appeal from a federal criminal conviction, the difference in how the inquiry is framed remains instructive and applicable to the habeas context. Scalia retorted in a footnote in his dissent that "[t]he jury has the right to apply its own logic (or illogic) to its decision to convict or acquit." *Id.* at 35 n.2.

There are subtle but real differences among these formulations. Indeed, scholars tend to agree that there are two very different approaches that judges use in determining harmless error.[41] Under the error-based approach, the focus of the court is on the likely impact of the error on the jury in the actual trial that took place. Under the guilt-based approach, the court considers a hypothetical trial conducted without the constitutional error, and asks whether the defendant would have nonetheless been convicted. The Supreme Court has used both approaches while rarely discussing the distinctions between them.[42]

Scholars have weighed in, mostly on the side of the error-based approach,[43] and almost all with the view that the two approaches are irreconcilable.[44] At this point, scholars even disagree about the current state of the law. After *Brecht*, Liebman and Hertz thought that the issue was decided in favor of the error-based approach, a reaction that other scholars shared after the Supreme Court's 1993 decision in *Sullivan v. Louisiana*.[45] Later cases,

---

[41] *See* Cooper, *supra* note 26, at 325 (noting that since *Chapman*, the two approaches "have battled for supremacy").

[42] As one scholar put it, "[i]n its harmless error jurisprudence, the Supreme Court has endorsed both general approaches to measuring harm." Gregory Mitchell, *Against "Overwhelming" Appellate Activism: Constraining Harmless Error Review*, 82 CAL. L. REV. 1335, 1341 (1994). Mitchell also observed that the "relatively short history of the harmless error test has been one of considerable confusion." *Id.* at 1342.

For examples of the Supreme Court using a "guilt-based" approach, see *Arizona v. Fulminante*, 499 U.S. 279, 310 (1991), *United States v. Hasting*, 461 U.S. 499, 510–11 (1983), *Brown v. United States*, 411 U.S. 223 (1973), *Schneble v. Florida*, 405 U.S. 427 (1972), *Milton v. Wainwright*, 407 U.S. 371 (1972), and *Harrington v. California*, 395 U.S. 250 (1969). For examples of the Supreme Court using an "error-based" approach, see *O'Neal v. McAninch*, 513 U.S. 432 (1995), *Brecht v. Abrahamson*, 507 U.S. 619 (1993), *Fulminante*, 499 U.S. at 296–300; *id.* at 313 (Kennedy, J., concurring), *Satterwhite v. Texas*, 486 U.S. 249, 259 (1988), and *Chapman v. California*, 386 U.S. 18 (1967).

[43] *See, e.g.*, Edwards, *supra* note 28, at 1192–94; Martha A. Field, *Assessing the Harmlessness of Federal Constitutional Error—A Process in Need of a Rationale*, 125 U. PA. L. REV. 32–36 (1976); Mitchell, *supra* note 42, at 1365–68.

[44] *See, e.g.*, John H. Blume & Stephen P. Garvey, *Harmless Error in Federal Habeas Corpus After* Brecht v. Abrahamson, 35 WM. & MARY L. REV. 163, 176 (1993) (describing the two approaches as "analytically distinct"). *But see* Cooper, *supra* note 26, at 332 (describing the "conceptual overlap" between the two approaches); Stephen A. Saltzburg, *The Harm of Harmless Error*, 59 VA. L. REV. 988, 1008 & n.64 (1973) (arguing that the two tests are "substantially alike"). In the scholarship and commentary, the "guilt-based" approach is also referred to as the "overwhelming evidence" test, *see, e.g.*, Cooper, *supra* note 26; Field, *supra* note 43, and the "hypothetical trial" approach, *see* Liebman & Hertz, *supra* note 34. The "error-based" approach is also referred to as the "effect-on-the-verdict" approach, *see, e.g.*, Edwards, *supra* note 28, and the "actual trial" approach, *see* Liebman & Hertz, *supra* note 34. For more extensive discussions of the difference between the two approaches, see Blume & Garvey, *supra* note 44, at 176–82, Edwards, *supra* note 28, at 1185–1209 (discussing the two approaches, and preferring the "effect-on-the-verdict approach" over the "guilt-based approach"), Field, *supra* note 43, at 16–36 (referring to these two approaches as "opposite extremes," but also considering a third approach), and Mitchell, *supra* note 42, 1341–47 (comparing the two approaches, and also considering hybrid approaches).

[45] 508 U.S. 275 (1993); *see also* Mitchell, *supra* note 42, at 1339 (concluding that after *Sullivan*, the guilt-based approach "should be considered an impermissible test").

particularly those applying harmless error in the federal criminal context, have led some commentators to conclude that *Sullivan* was a mere blip, and that Chief Justice Rehnquist's guilt-based approach had emerged victorious.[46] Others view the two approaches as uneasily coexisting, with one scholar going so far as to say that judges have a choice in which approach to use.[47] One wonders if Justice Rehnquist, who frequently uses the guilt-based approach, and Justice Scalia, who is fiercely partial to the error-based view, would give the same answer if asked about the current state of the law.

In part the different approaches simply reflect differing levels of comfort with a certain amount of appellate fact-finding, which harmless error necessarily entails. Justice Rehnquist shrugs his shoulders, it seems: as he stated in *Neder*, the erroneous admission or exclusion of evidence inevitably affects "the jury's deliberative process in ways that are, strictly speaking, not readily calculable."[48] Scalia, though, cannot tolerate such invasion into the province of the jury; if judicial fact-finding is needed to uphold the verdict, then the conviction simply cannot stand. As he reminded his colleagues with outraged use of italics in his *Neder* dissent, "*the Constitution does not trust judges to make determinations of criminal guilt.*"[49]

Fundamentally, the harmless-error inquiry is an empirical, if unanswerable, one: what impact did the error have on the actual jury's verdict? That is, the jurors sit through the trial, listen to the opening statements, and receive various pieces of evidence, all the while making a judgment about whether they think the defendant is guilty or innocent. The constitutional error has a tangible impact on the mind of the jurors. For example, after hearing an improperly admitted co-defendant's confession, four jurors may be significantly more likely to convict the defendant, four other jurors may be absolutely convinced that the defendant is guilty, and the remaining four jurors may give the confession little to no weight. Actual impact on the verdict is the focus of the harmless-error analysis.[50]

---

[46] *See* Cooper, *supra* note 26, at 324 (concluding, after reviewing Supreme Court doctrine, that "a clear, if narrow, majority of the Court supports the overwhelming evidence standard" while acknowledging that "the question cannot be regarded as settled").

[47] Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. CHI. L. REV. 511, 575 (2004) (describing harmless-error analysis as an area where legal decisionmakers "have a choice as to the nature of their task" because of the two approaches).

[48] Neder v. United States, 527 U.S. 1, 18 (1999).

[49] *Id.* at 32 (Scalia, J., dissenting).

[50] *See* Mitchell, *supra* note 42, at 1357 ("An appellate court deciding a harmless error case must make a determination that is largely factual in nature."); Stacy & Dayton, *supra* note 20, at 128–29 n.195 (noting that "[l]anguage in several Supreme Court decisions suggests that harmless error analysis entails an empirical evaluation of the likely influence the error had on the jury's decision to convict" (citing cases)). *But see* Davis, *supra* note 26, at 89–90 ("When the issue arises, it seems likely that the Supreme Court will determine that, despite its fact-like characteristics, harmless error is at least a mixed question and perhaps pure law.").

This cognitive task is quite different than most undertaken by judges. Judges perform a variety of cognitive tasks in deciding a case: they interpret statutes, determine the state of legal doctrine, and apply legal doctrine to a particular set of facts. They also make retrospective evaluations of the reasonableness of people's behavior, but generally limit that analysis to the parties in a case.[51] Certainly, in bench trials where judges are the finders of fact, they frequently must determine state of mind, such as the intent of a criminal defendant. But federal district and appellate judges are rarely asked to determine the state of mind of jurors in a state criminal trial. The task of the judge in harmless-error determinations is to read the minds of twelve jurors—people the habeas judge has never met, and about whom he has virtually no information.[52]

So regardless of the precise question, how are judges to determine the answer? This question of the content of the harmless-error analysis—once you get past the "guilt-based" or "error-based" question—is touched upon in *Delaware v. Van Ardsall*, where the Court specified the factors a court should consider in evaluating the harm of a Confrontation Clause violation.[53] With respect to other types of constitutional violations, though, judicial discretion in determining harmlessness is largely unguided. In Part III, I take a closer empirical look at how federal appellate judges have actually been making this determination.

III. HOW ARE JUDGES DETERMINING HARMLESSNESS?: AN EMPIRICAL ANALYSIS OF THE POST-*BRECHT* CASES

Part III is an empirical analysis of the post-*Brecht* cases in the federal courts of appeals, attempting to answer the question: what has happened to harmless-error analysis since *Brecht*?[54] Or put differently, how have courts

---

[51] For example, when determining negligence in a tort case, a judge determines what a reasonable person in the defendant's position would do and then compares the defendant's behavior to that of the reasonable person.

[52] As the *Kotteakos* Court put it, "[t]he crucial thing is the impact of the thing done wrong on the minds of other men, not on one's own, in the total setting." Kotteakos v. United States, 328 U.S. 750, 767 (1946). For skeptical views on this task, see Cooper, *supra* note 26, at 312, indicating that the determination of harmlessness "necessarily rests on fiction" and involves a task that appellate judges are "fundamentally unqualified to perform," and Sam Kamin, *Harmless Error and the Rights/Remedies Split*, 88 VA. L. REV. 1, 21 (2002), arguing that the impossibility of the cognitive task of harmless error means that "there is simply no way that consensus can be found regarding the results of hard cases" and that the conclusions obtained from harmless-error analysis "can be no better than science fiction."

[53] 475 U.S. 673, 681–82 (1986).

[54] In asking this question, I do not attempt to answer the question that divided commentators in the immediate aftermath of *Brecht*: whether the standard would be significantly different than the *Chapman* standard. Rather, I take the view offered in Justice O'Connor's *Brecht* dissent that both standards "require[] an exercise of judicial judgment that cannot be captured by the naked words of verbal formulae," echoed in Justice Stevens's concurrence that "[i]n the end, the way we phrase the governing standard is far less important than the quality of the judgment with which it is applied." 507 U.S. 619, 656 (1993) (O'Connor, J., dissenting); *id.* at 643 (Stevens, J., concurring). I therefore look at how that "judicial

been determining which constitutional errors should be remedied? These post-*Brecht* cases conduct the harmless-error analysis in the context of federal habeas review of state court convictions infected by constitutional violations.

A. Methodology

The data come from a sample of 263 published habeas opinions from the federal courts of appeals, decided from May 1993 through July 2004, in which the majority opinion directly addressed the issue of whether an alleged error was harmless or harmful. Each error discussed in a given case generated a separate analysis, resulting in 315 analyses overall. Of these, 287 are *Brecht* harmless-error analyses, with the remainder consisting of analyses of "prejudice" for claims where the existence of a constitutional violation itself depends on the effect on the verdict.[55] Several variables were coded for each analysis including the type of error, the approach used by the court to arrive at its decision, and the reason for determining that the error was harmless or harmful.[56] Methodologically, this analysis takes the

---

judgment" is being exercised in cases applying *Brecht*, under a standard that we know "will permit more errors to pass uncorrected." *Id.* at 656 (O'Connor, J., dissenting).

[55] These include ineffective assistance of counsel claims and claims that the prosecution failed to disclose exculpatory evidence. *See* Edwards, *supra* note 28, at 1178 (summarizing the kinds of constitutional claims that incorporate the harmlessness inquiry into the determination of whether an error has even occurred). I have included these analyses as well because the inquiry is essentially the same as that undertaken in a *Brecht* harmless-error analysis.

[56] Each separate analysis was coded for several variables. There were four summary case statistics including: date the case was decided, crime, length of sentence the petitioner received, and time since the trial. In addition, the decision type was coded as being en banc, unanimous, concurrence or dissent. The analyses were coded for whether the defendant testified, and whether the defendant had a lawyer or was pro se. The error was coded for the time in which it took place: pretrial, trial or sentencing, and if during the trial, whether it was early, late, or in the middle of the trial. The type of error was also coded as being included evidence, excluded evidence or other, and for which constitutional right it violated (Fifth Amendment *Miranda* violation, Sixth Amendment confrontation error, Sixth Amendment right to counsel error, Fourteenth Amendment jury instruction error, Fourteenth Amendment prosecution error, Fourteenth Amendment right to present a defense error, or other type of error).

The harmless error analysis was coded according to the standard used, whether the court placed the burden on the petitioner or the state, whether the court used a test for determining harm, whether the court used a guilt-based or verdict-based approach, whether the jury asked questions about the evidence, whether the jury instructions were mentioned, whether the length of the deliberations were mentioned, whether the court decided based on the weight of the evidence or the plausibility of competing narratives, whether the summation/opening was mentioned, and whether the defense theory of the case was mentioned. The reasons for finding an error harmless were coded using the following categories: the quantity of the evidence against the petitioner, the quality of the evidence against the petitioner, the presence of an ameliorating instruction, the evidence was merely cumulative, the error was irrelevant, or the evidence was insignificant. The reasons for finding the error harmful or for remanding on harm were that the other evidence was weak, the errors were cumulatively significant, the improper evidence was significant, or the error was itself significant.

judicial reasoning presented in written opinions seriously as an indicator, at least, of the actual reasoning used to reach decision.[57]

This empirical inquiry into lower court application of harmless-error analysis since *Brecht* has several components. For example, after *Brecht*, there was confusion in the lower courts about whether the petitioner or the state bore the burden of demonstrating harm. This confusion was clarified in a 1995 case called *O'Neal v. McAninch*, which held that when the judge was in "grave doubt" on the harmlessness of the error, the petitioner should be granted relief.[58] Are courts following *O'Neal*? Are there certain kinds of errors that are more likely to be held harmful than others? What factors are courts using, regardless of the approach, to actually determine whether an error is harmful? Are courts using predominantly the guilt-based or the error-based approach? Finally, does there appear to be any common ground between these two approaches, or are we stuck in doctrinal deadlock?

B. Summary of Findings

Most of the errors (72%) took place during trial.[59] Of the sentencing errors, nearly all (seventy-five out of seventy-seven) were in capital cases. In 80% of the analyses, the crime was murder,[60] and 54% were murder trial analyses. Improperly included evidence comprised 52% of the errors, improperly excluded evidence 18% and other error types, such as improper jury instructions, the remainder.[61] As Table 1 shows, the most common violations were errors in instructing the jury and Confrontation Clause violations, each with roughly one-fifth of the errors. The rest of the errors included *Miranda* violations, ineffective assistance of counsel, prosecutorial misconduct, and impairment of the right to present a defense.[62]

---

[57] My sample, and method of analysis, falls somewhere between an in-depth, doctrinal analysis of a handful of Supreme Court cases and a large-scale, empirical analysis that involves running multivariate regressions on a large database of cases. The number of harmless-error analyses is large enough to permit some generalization, but small enough to actually be able to read all the cases and have a textured sense of what is going on.

[58] 513 U.S. 432 (1995). Baseball fans know this as the "tie goes to the runner" rule.

[59] In 227 of the 315 analyses the error took place during the trial. Of the remaining errors, nineteen (6%) were pretrial and seventy-seven (24%) were sentencing errors.

[60] Of the 315 analyses, 253 were for trials where the petitioner was accused of murder.

[61] One hundred sixty-four of the analyses were improperly included evidence errors, while fifty-seven were improperly excluded evidence and ninety-four (30%) were other types of errors.

[62] Forty-three (14%) of the analyses were *Miranda* violations; sixty-five (21%) were confrontation clause errors; twenty-six (8%) were right to counsel errors; seventy (22%) were jury instruction errors; thirty-one (10%) were prosecution errors; twenty-seven (8%) were right to present a defense violations; and fifty-three (17%) were other types of errors.

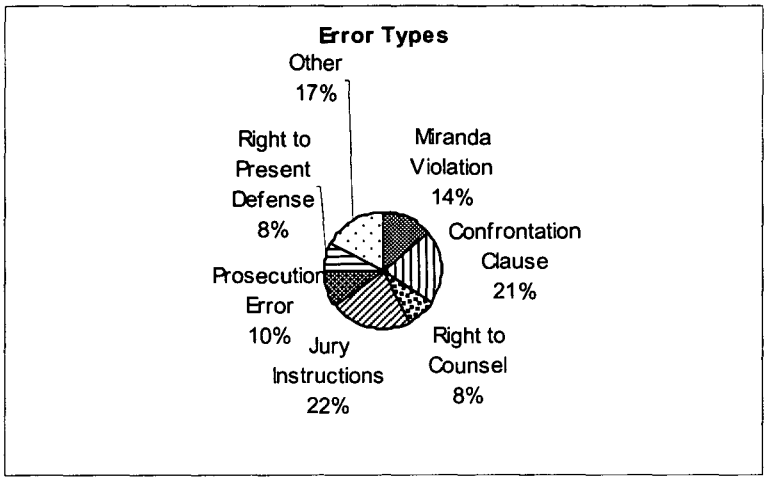

TABLE 1

In nearly two out of three analyses, the court found the error to be harmless.[63] Overall, trial errors were more likely to be held harmless (68%), as compared to sentencing errors (49%), which were almost all death penalty cases.[64] The nature of the crime also seemed to impact whether or not an alleged error was found to be harmless or harmful: the court found 68% of the errors harmless in murder analyses and only 44% in non-murder analyses.[65] Indeed, three out of four (76%) errors during murder trials were held harmless.[66]

There remains little in the way of guidance for how to determine harmlessness. Less than 20% of the analyses used a test for determining harm, although analyses that cited a test for determining harm were more likely to rule that the error was harmful (45%, compared to 32% overall).[67] When

[63] In 199 of the analyses, the court ruled the error harmless. In 100 of the analyses the court found the error to be harmful, and in sixteen of the cases the court chose to remand on harm. Because of the number of unpublished opinions not included in this analysis, this no doubt understates the overall percentage of harmless-error analyses that resulted in a finding of harmlessness.

[64] One hundred fifty-five (68%) of the 227 trial errors were found harmless, sixty-four (28%) were harmful and the court remanded in eight (4%) of the cases. In contrast, the sentencing errors were found harmless in only thirty-eight (49%) of seventy-seven total and harmful in thirty-two (42%). The court remanded seven (9%) of the sentencing errors.

[65] In 172 (68%) of the 253 murder analyses the court found the error harmless, while in seventy cases (28%) the error was found harmful, and in eleven (4%) the court remanded on harm. In contrast, the court found just twenty-seven (44%) of the non-murder errors harmless, thirty (48%) harmful, and remanded on harm in five (8%) of the non-murder analyses.

[66] Of the 170 analyses that were for murder trials, 130 (76%) of the errors were found harmless, thirty-seven (22%) of the errors were found harmful, and in three (2%) of the analyses the court remanded on harm.

[67] A test was cited in sixty-two (20%) of the 315 analyses. The error was found harmless in thirty-one (50%) of those analyses, harmful in twenty-eight (45%), and the court remanded on harm in three

the violation at issue was a Confrontation Clause violation, the courts frequently used a test for determining harmlessness, with 41% of the Confrontation Clause harmless-error analyses citing a test, primarily the Supreme Court's analysis in *Delaware v. Van Ardsall*.[68] Of the rest of the analyses of other types of constitutional errors, only 14% employed any test at all.[69]

It is unclear how consistent these decisions are with Supreme Court precedent, in part because the Supreme Court precedent is itself inconsistent in certain areas. Nonetheless, there is cause for concern. Even after *O'Neal* ruled otherwise, of the analyses that referenced who bears the risk of uncertainty on harmlessness, more than one in four improperly placed the burden on the petitioner.[70] Moreover, the extent to which appellate courts weighed the evidence indicates that Sixth Amendment boundaries may have been crossed.

(5%) of these cases.

[68] 475 U.S. 673, 681–82 (1986). In *Van Ardsall*, the Court indicated that:

> Whether such an error is harmless in a particular case depends upon a host of factors, all readily accessible to reviewing courts. These factors include the importance of the witness's testimony in the prosecution's case, whether the testimony was cumulative, the presence or absence of evidence corroborating or contradicting the testimony of the witness on material points, the extent of cross-examination otherwise permitted, and, of course, the overall strength of the prosecution's case.

Id. at 684. Of the sixty-four confrontation clause cases, twenty-six (41%) cited a test—mostly this one—for determining harm. Only thirty-six of the remaining cases cited a test for determining harm (14% of the non-confrontation cases).

[69] Other harmless error tests included those used to evaluate post-arrest silence in violation of *Doyle v. Ohio*, *see, e.g.*, Sallahdin v. Gibson, 275 F.3d 1211, 1230–31 (10th Cir. 2002), those that use the same analysis for whether an erroneous jury instruction constitutes a due process violation, *see, e.g.*, Garceau v. Woodford, 275 F.3d 769, 777 (9th Cir. 2001) (examining the "quality, significance, and pervasiveness" of the evidence related to the erroneous instruction), those that evaluate the effect of improper prosecutorial comments, *see, e.g.*, Whitehead v. Cowan, 263 F.3d 708, 728 (7th Cir. 2001) (declaring that "prejudice" analysis on the merits of whether there is a constitutional violation, not harmless error analysis per se (citing Darden v. Wainwright, 477 U.S. 168 (1986))); Maurino v. Johnson, 210 F.3d 638, 647 (6th Cir. 2000) (citing Angel v. Overberg, 682 F.2d 605, 608 (6th Cir. 1982) (en banc)), those used to evaluate failure to provide defendant with expert psychiatric assistance in capital sentencing proceeding where future dangerousness is at issue, *see, e.g.*, Tuggle v. Netherland, 79 F.3d 1386, 1393 (4th Cir. 1996) (outlining six-factor test), those that examine whether the alleged introduction to the jury of extrinsic evidence affected the verdict, *see* Lawson v. Borg, 60 F.3d 608, 610 (9th Cir. 1995), those used to evaluate improper admission of codefendant's confession, *see, e.g.*, Samuels v. Mann, 13 F.3d 522, 526–27 (2d Cir. 1993), and those that evaluate the erroneous admission of evidence more generally, *see, e.g.*, Wray v. Johnson, 202 F.3d 515, 526 (2d Cir. 2000).

[70] *See, e.g.*, Dillard v. Roe, 244 F.3d 758, 774 (9th Cir. 2001) ("[W]e may not grant habeas relief unless defendant can establish that as a result of the state trial court's error, he suffered 'actual prejudice.'"); Doan v. Brigano, 237 F.3d 722, 738 (6th Cir. 2001) (stating that "[t]he burden in this analysis is on the habeas petitioner, and he has failed to show that the juror misconduct" influenced the verdict); Sassounian v. Roe, 230 F.3d 1097, 1110 (9th Cir. 2000) (Silverman, J., dissenting in relevant part) ("The question is whether the petitioner has carried his burden of proving that [the error] had a *substantial and injurious effect* on the verdict."). Of the 268 post-*O'Neal* cases, 160 (60%) of the analyses did not mention who the burden was on, 21 (8%) said the burden was on the state, 30 (11%) said the burden was on the petitioner, and 57 (21%) of the cases properly cited *O'Neal*.

Indeed, courts frequently appear to improperly supplement their error analyses by weighing the evidence themselves.[71] In the analyses where the court weighed the evidence in determining the error's harmlessness, the error was found harmless in 79% of analyses, while in those where the court considered the effect of the error on the two sides' narratives, the error was found harmless in just 31% of all analyses.[72] Moreover, the reason for determining harmless error was most often that the quality and quantity of the evidence against the petitioner was overwhelming (48%), with the error being irrelevant to the prosecution's case (15%) or the evidence being insignificant (17%) also frequently cited.[73]

---

[71] *See* Whitney v. Horn, 280 F.3d 240, 257 (3d Cir. 2002) (rejecting argument that erroneous jury instruction caused harm on the grounds that no "reasonable jury could have had any doubt about whether the defendant was too inebriated to form the intent to kill"); Fortini v. Murphy, 257 F.3d 39, 52 (1st Cir. 2001) (reasoning that neither eyewitness identifications nor confession evidence were particularly "reliable" or "compelling"); Barrett v. Acevedo, 169 F.3d 1155, 1165 (8th Cir. 1999) (en banc) (finding that the prosecution's expert testimony on critical issue in the case was "more persuasive" and "simply makes more sense than the testimony of the other experts," and concluding that "we have no difficulty finding that the jury believed [prosecution's expert] over the others"); Lyons v. Johnson, 99 F.3d 499, 504 (2d Cir. 1996) (finding harm and noting that "the identifications provided by the prosecution's eyewitnesses were shaky from the start"); Thompson v. Borg, 74 F.3d 1571, 1583 (9th Cir. 1996) (Reinhardt, J., dissenting) ("The conflicting theories and evidence at trial presented a difficult case for the fact finder."); Levasseur v. Pepe, 70 F.3d 187, 195 (1st Cir. 1993) ("Was the properly admitted evidence so strong that it overwhelmed the impact of the [error]?").

[72] In 148 (79%) of the 189 analyses in which the evidence was weighed, the court found the error harmless, in thirty-seven (20%) the court found the error harmful, and in four (2%) of the analyses the court remanded on harm. In contrast, the court found the error harmless in just sixteen (31%) of the fifty-two analyses where the competing narratives were considered. As Professor Mitchell pointed out in his analysis, the difference in outcomes between the two approaches may not be attributable to the approaches themselves, but rather the judges may choose the particular approach in part based on their predilection to reverse or affirm convictions. Mitchell, *supra* note 42, at 1352.

[73] Ninety-five (48%) of the harmless analyses were based on the fact that the quality and/or quantity of the evidence against the petitioner was overwhelming. In thirty-one (15%) of the harmless analyses the reason given was that the error was irrelevant, and in thirty-three (17%) the reason given was that the evidence was insignificant. Other reasons cited included that there had been an ameliorating instruction given to the jury or that the evidence was cumulative to other evidence admitted.

The most often cited reason why the error was found harmful was that the error was significant. This was the reason given in 45% of cases where the error was harmful. The evidence that was the subject of the error being significant and the other evidence being weak were also often cited reasons why the error was found to be harmful.

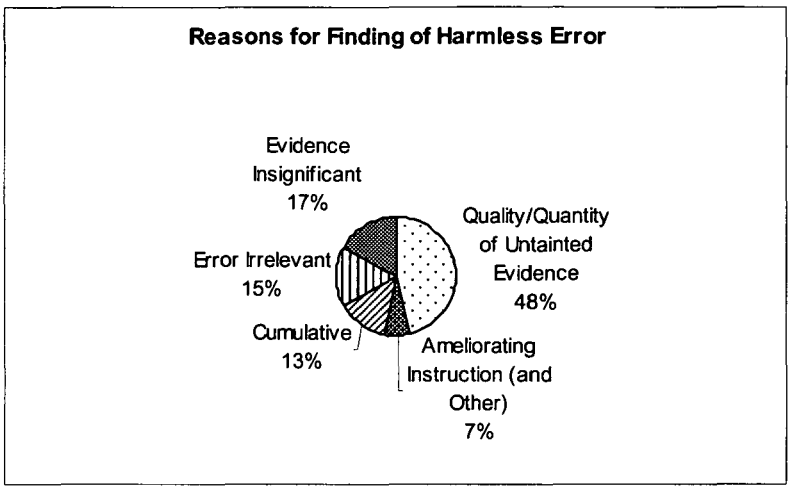

TABLE 2

But the notion that the evidence can be characterized objectively as "overwhelming," reflected in many of the 48% of cases referenced above in Table 2, is undermined in part by judges frequently having different views on the strength of such evidence.[74]

Throughout the cases, empirical assumptions abound about the impact of certain kinds of evidence on jurors and appear to be influential in the harmless-error analysis. Many courts assume, for example, that jurors follow instructions given to them by the judge, and can ignore both lawyers'

---

[74] *See, e.g., Barrett*, 169 F.3d at 1164, 1171 (stating that "[t]his was not a close case," while the dissent quoted three judges of the Iowa Supreme Court describing the case as "obviously close"); Agard v. Portuondo, 117 F.3d 696, 705 (2d Cir. 1997) ("On direct review, the Appellate Division characterized the evidence of [defendant's] guilt as 'overwhelming.' Having reviewed the entire record of the trial, we cannot agree with that characterization."); Kyles v. Whitley, 5 F.3d 806, 820 (5th Cir. 1993) (King, J., dissenting) (arguing that there is by no means "overwhelming evidence" of defendant's guilt, as the majority indicated; Supreme Court reversed in a five-to-four decision, 514 U.S. 419, 454 (1995), with majority arguing that this is not the "massive" case the dissent describes); *see also* Mitchell, *supra* note 42, at 1351 (finding that disagreement within the same case about an error's harmful or harmless nature was common).

A friend points out that plenty of people thought the evidence against O.J. Simpson was "overwhelming," but the jury acquitted. The task on harmless-error analysis is to predict the effect of the error on the jury, not what the judges think of the evidence. Moreover, before one can characterize the evidence as "overwhelming" or anything else, one has to have a theory about whether the evidence is to be viewed in the light most favorable to the prosecution, the defense, or somewhere in between. The proper approach is not clear from either the case law or scholarship. *See, e.g.*, Rossetti v. Curran, 80 F.3d 1, 18 (1st Cir. 1996) (stating that though the other evidence was "impressive, if taken most favorably to the government," courts doing harmless-error analysis do not look "only at the government's best case but rather at the evidence as a whole").

arguments to the contrary[75] and their own intuitions.[76] One court assumed, not unreasonably, that the state's misrepresentation of evidence was influential.[77] But courts rarely rely on actual social science research about the effects of different kinds of evidence, argument, or instructions on jurors.[78]

Indeed, the harmless-error opinions tend to provide little sense of the actual trial dynamics in each case. Only one in three of the analyses reference the opening statements or summations, which frame the case for the jury.[79] Less than 40% mention the defense theory of the case, either as part of the harmless-error analysis or otherwise.[80] In other words, many of the cases lack a textured sense of the centrality of the error, or the issue it bears upon, to the competing narratives that are being put before the jury.

C. The Effects of Doctrinal Deadlock: Is There a Way Out?

As expected, the guilt-based and error-based approaches certainly seem to lead to very different outcomes. As Table 3 indicates, analyses where a guilt-based approach was used by the court during its analysis found the error harmless 93% of the time, while those using an error-based approach found the error harmless only 47% of the time.

---

[75] *See, e.g.*, Caldwell v. Bell, 288 F.3d 838, 845 (6th Cir. 2002) (Norris, J., dissenting) (complaining that the majority ignores the fact that in determining the effect that a constitutional instruction has had on a verdict, courts are "bound by the presumption that 'juries follow their instructions'" (quoting Washington v. Hofbauer, 228 F.3d 689, 706 (6th Cir. 2000))).

[76] *See, e.g.*, *Whitney*, 280 F.3d at 257 (arguing that "expecting jurors' 'common sense' judgment to prevail over the court's instructions would conflict with the presumption that juries follow their instructions"). *But see* McCracken v. Gibson, 268 F.3d 970, 977 (10th Cir. 2001) (holding that an unconstitutional "presumed not guilty" instruction had "little, if any" effect on the credibility determination central to the decision of guilt).

[77] *See, e.g.*, Gall v. Parker, 231 F.3d 265, 313 (6th Cir. 2000) ("Misrepresenting facts in evidence can amount to substantial error because doing so 'may profoundly impress a jury and may have a significant impact on the jury's deliberations.'" (quoting Donnelly v. Dechristoforo, 416 U.S. 637, 646 (1974))).

[78] Cooper, *supra* note 26, at 331 (noting that there is "little suggestion" that appellate judges conducting harmless error review rely on the "substantial literature" on jury decisionmaking). This kind of research has been used in several isolated instances. *See, e.g.*, Strickler v. Greene, 527 U.S. 263, 305 (1999) (Souter, J., dissenting) (relying in part on empirical study for the proposition that guilt-phase evidence and arguments will often significantly affect jurors' choice of sentence (citing William J. Bowers et al., *Foreclosed Impartiality in Capital Sentencing: Jurors' Predispositions, Guilt-Trial Experience, and Premature Decision Making*, 83 CORNELL L. REV. 1476, 1486–96 (1998))); Rice v. Wood, 77 F.3d 1138, 1142–43 & n.4 (9th Cir. 1996) (en banc) (Kozinski, J.) (citing a study that found conviction rates in robbery and burglary cases were 40% to 180% higher where there is a confession and comparing this strong evidence of causal impact with the "mere dram of authority supporting the view that defendant suffers *any* concrete harm whatever by being absent when the jury returns" its verdict, the issue in the case at bar).

[79] The summation/opening was mentioned in ninety-seven (31%) of the analyses.

[80] The defense theory of the case was mentioned in just 110 (38%) of the analyses.

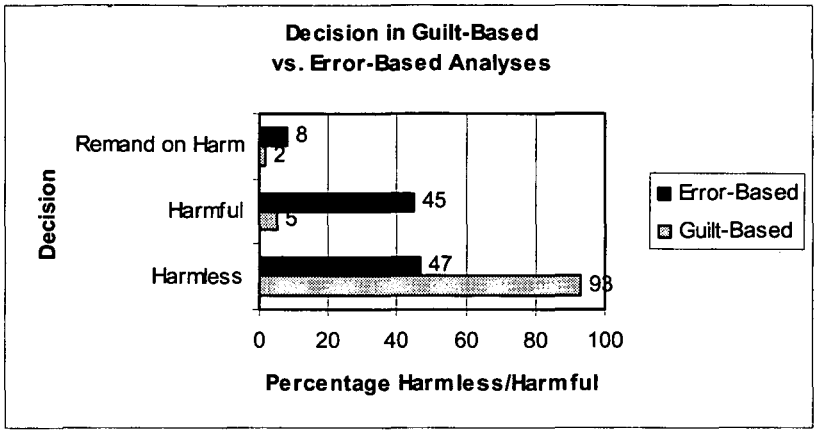

TABLE 3

In cases where the majority and dissent differ on the harmlessness of the error, it frequently seems that they are talking about two entirely different trials, or that they simply have very different perspectives on the same trial. Most of the time, one side is using the guilt-based approach, and the other the error-based approach. Is there any principled common ground, or is this just results-oriented jurisprudence at work, with judges choosing the approach to use based on their desired outcome, in turn determined by their ideology or view of factual guilt?

To search for a way forward, it makes sense to return to first principles, starting with the language of *Brecht.* In determining whether the "error" brought about "harm," according to the Court in *Brecht,* the judge is to determine whether the error had "substantial and injurious effect or influence in determining the jury's verdict."[81] So to determine the connection between the "error"—not a mistake but a deviation by the judge or prosecution from constitutionally mandated duties to a criminal defendant—and the "harm" (the conviction), one must look to effect's necessary antecedent, "cause." A breach of duty, resulting harm, and an inquiry into the causal connection—harmless-error analysis starts to sound less like the law of constitutional criminal procedure, or even remedies, and more like tort law. In Part IV, I reconceptualize harmless-error as a constitutional tort, and consider whether this can bring conceptual clarity and doctrinal consensus.

[81] Brecht v. Abrahamson, 507 U.S. 619, 638 (1993).

IV. BRINGING THE LENS OF TORT LAW TO HARMLESS ERROR

A. Habeas as a Constitutional Tort

The universe of "constitutional torts" is generally thought to consist of those claims brought under 42 U.S.C. § 1983, the civil rights statute. This is sensible, as tort claims are normally claims for damages, which § 1983 authorizes. Bringing a writ of habeas corpus, of course, is not asking for compensation; rather, the writ seeks the remedy of freedom.[82] Nonetheless, the notion of a habeas claim as a tort claim is both conceptually coherent and practically useful.[83]

After all, the basic claim at the heart of every habeas petition is that the petitioner's rights were violated in some fashion at trial, which caused his conviction. So when a prisoner brings a habeas claim against the state, all the elements of a tort claim are present: (1) duty (of the state to provide petitioner with a criminal trial that conforms to constitutional procedure); (2) breach of duty (the constitutional violation); (3) causation (the error tilting the jury toward guilt); and (4) harm (conviction). In this context, harmless error is an affirmative defense to the constitutional tort claim, a characterization that has considerable support from two main strands in the case law: the cases culminating in *O'Neal* that place the burden of uncertainty of harmlessness on the state, and the cases holding that the state has waived any harmless-error defense if not raised in a timely fashion.[84]

Though the tort law analogy in harmless-error analysis is largely unexplored,[85] it did play a role in the 1995 Supreme Court case *O'Neal v. McAninch*.[86] *O'Neal* turned on who bore the burden of demonstrating harm in a federal habeas case: was it the petitioner's burden to persuade the judge that the error had a "substantial and injurious" effect on the verdict, or was it the state's burden to persuade the judge that the error was harmless? The critical issue in the case centered on causation.

---

[82] *See* Fallon & Meltzer, *supra* note 25, at 1804 ("Many constitutional tort actions . . . involve requests for injunctions."); Daniel J. Meltzer, *Harmless Error and Constitutional Remedies*, 61 U. CHI. L. REV. 1, 29 (1994) ("An appellant attacking an adverse judgment is seeking affirmative relief against a past wrong, not unlike a plaintiff suing the government for damages in a constitutional tort action.").

[83] *Cf.* United States v. Johnson, 380 F.3d 1013 (7th Cir. 2004) (Posner, J.) (arguing that the independent-source and inevitable-discovery doctrines in Fourth Amendment law are "easily collapsed into the familiar rule of tort law that a person can't complain about a violation of his rights if the same injury would have occurred even if they had not been violated").

[84] Another way of thinking about the habeas claim and the harmless-error defense is with reference to burden-shifting on the issue of causation. If the petitioner raises a claim that his constitutional rights were violated at trial, a determination by a judge that there was constitutional error shifts the burden of persuasion on causation to the state to negate the presumption that the error caused the harm. The harmless-error inquiry then focuses on causation: whether the error caused the conviction.

[85] *But see* Brandon L. Garrett, *Innocence, Harmless Error and Federal Wrongful Conviction Law*, 2005 WIS. L. REV. (forthcoming) (comparing the harmless-error inquiry to the causation element in a § 1983 wrongful conviction claim).

[86] 513 U.S. 432 (1995).

The state argued, supported by the Solicitor General as *amicus curiae*, that the language of the federal habeas statute with respect to state prisoners—"in custody in violation of the Constitution or laws or treaties of the United States"[87]—meant that a "causal link" is required "between the violation and the custody."[88] The habeas petitioner is bringing a civil action as a plaintiff, and according to basic tenets of civil procedure and evidence, should bear the risk of non-persuasion. Siding with this argument in his dissent, Justice Thomas cited to the *Restatement (Second) of Torts*, quoting the statement that "[t]he burden of proof that the tortious conduct of the defendant has caused the harm to the plaintiff is upon the plaintiff."[89] Justice Thomas pointed out that by discussing "*harmful* errors as opposed to mere errors," the majority "implicitly agrees that causation is necessary."[90]

Justice Thomas is clearly right that treating only harmful errors as worthy of a remedy necessarily implies the need for a causal connection between the error and the conviction. However, there is no reason why the burden necessarily must be on the petitioner to demonstrate causation. And after considering the "basic purposes underlying the writ of habeas corpus,"[91] the Court held that if "in the judge's mind, the matter is so evenly balanced that he feels himself in virtual equipoise as to the harmlessness of the error,"[92] then the error will be considered harmful, and the conviction overturned.

Though an important clarification of the law (the circuits were split on this issue before *O'Neal*), *O'Neal* got somewhat lost between *Brecht*'s application of a more prosecution-friendly standard to constitutional errors in 1993, and the Anti-Terrorism and Effective Death Penalty Act's ("AEDPA") significant reshuffling of the habeas landscape in 1996. As a result, the broader implications of the tort law analogy described in *O'Neal* have been ignored by judges and scholars.[93] Regardless of who bears the risk of uncertainty, both the majority and the dissent implicitly seemed to agree: *the harmless-error analysis is the causation inquiry in the habeas version of a constitutional tort claim.* As the Court in *Brecht* explains the

---

[87] 28 U.S.C. § 2254(a) (2000).

[88] 513 U.S. at 446.

[89] *Id.*

[90] *Id.*

[91] *Id.* at 442.

[92] *Id.* at 435.

[93] Writing before *O'Neal* was decided, Professors Fallon and Meltzer, in a footnote, had considered the possibility that harmless error analysis could be seen as "a species of causation doctrine at the post-conviction stage, rather than as a remedial doctrine." Fallon & Meltzer, *supra* note 25, at 1772 & n.222. But they rejected the suggestion on several grounds, primarily because harmless error analysis had a more stringent standard for constitutional errors—a distinction with no particular relation to causation and because violations of some rights were subject to automatic reversal. Since their article, however, *Brecht* lowered the standard for constitutional errors, and the number of rights subject to harmless error analysis has continued to expand, undercutting these reasons.

nature of the inquiry, the judge must ask whether the error had a "substantial and injurious effect" on the jury's verdict.[94] In tort law terms, did the error *cause* the petitioner's conviction?

B. The Doctrinal Deadlock: Reflecting Different Tort-Law Approaches to Factual Causation

But what does it mean for an error to cause a conviction? Understanding what constitutes "cause-in-fact," otherwise referred to as "factual causation" or "actual causation," has vexed torts scholars and judges for years.[95] For much of the first two centuries of American law, the "but for" test was the *sine qua non* of factual causation.[96] In the first *Restatement of Torts*, the "substantial factor" test was adopted primarily for the proximate cause inquiry, but was soon applied to cause-in-fact with many jurisdictions following suit.[97] In recent years, however, the pendulum has swung the other way, with the "substantial factor" test roundly criticized as indeterminate and unhelpful. As the draft *Restatement (Third) of Torts* puts it, "[t]he substantial-factor test has not . . . withstood the test of time, proving confusing and being misused."[98] Accordingly, the draft *Restatement (Third) of Torts* section on factual causation eliminates any reference to the "substantial factor" test for most types of cases—a major change in doctrine. The "but for" test is once more in ascendance.[99]

The "but for" causation inquiry is the essence of counterfactual reason-

---

[94] This phrase was borrowed from *Kotteakos v. United States*, 328 U.S. 750, 776 (1946).

[95] For an overview of the debates on factual causation, *see* KENNETH S. ABRAHAM, THE FORMS AND FUNCTIONS OF TORT LAW (2d ed. 2002); ARNO C. BECHT & FRANK W. MILLER, THE TEST OF FACTUAL CAUSATION IN NEGLIGENCE AND STRICT LIABILITY CASES (1961); H.L.A. HART & TONY HONORE, CAUSATION IN THE LAW (2d ed. 1985); ROBERT E. KEETON, LEGAL CAUSE IN THE LAW OF TORTS (1963); Wex S. Malone, *Ruminations on Cause-In Fact*, 9 STAN. L. REV. 60 (1957).

[96] *See* RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL HARM (BASIC PRINCIPLES) § 26 cmt. b (tentative draft no. 2, 2002) (explaining that "but-for" test means that "an act is a factual cause of an outcome if, in the absence of the act, the outcome would not have occurred") [hereinafter RESTATEMENT (THIRD)].

[97] *See* Richard W. Wright, *Causation in Tort Law*, 73 CAL. L. REV. 1735, 1781–84 (1985) (describing the origin of the substantial-factor test as a guide for resolving proximate cause issues and its subsequent application to factual causation through the influential Prosser treatise on torts); John D. Rue, Note, *Returning to the Roots of the Bramble Bush: The "but for" Test Regains Primacy in Causal Analysis in the American Law Institute's Proposed Restatement (Third) of Torts*, 71 FORDHAM L. REV. 2679, 2687–93 (2003) (describing the rise of the "substantial factor" test); *see also* RESTATEMENT (THIRD), *supra* note 96, § 26 reporters' note cmt. j (noting that in first two versions of the *Restatement of Torts*, the substantial-factor test appears also to be doing "proximate-cause duty").

[98] *See* RESTATEMENT (THIRD), *supra* note 96, § 26 cmt. j. In virtually all cases, the *Restatement* drafters decided that the "substantial-factor" test provided "nothing of use in determining whether factual cause exists." *Id.* § 26 reporters' note cmt. j.

[99] *See id.* § 26 reporters' note cmt. j (noting that the substantial-factor test has "few supporters" among commentators, and citing the critics); David A. Fischer, *Causation in Fact in Omission Cases*, 1992 UTAH L. REV. 1335, 1338 (noting that the but-for test "is the most widely accepted test for determining cause in fact").

ing. As one scholar has described it, the but-for test "instructs the fact finder to re-create an imaginative past, in which the fact finder eliminates the tortious act and plays out an alternative (counterfactual) history."[100] Harmless-error cases using the guilt-based approach employ similar reasoning, determining whether, if not for the error, the defendant would have been convicted. On the other hand, the error-based approach asks not a counterfactual question but a historical one—essentially, was the error a "substantial factor" in the jury's verdict?[101]

Indeed, the debate among tort scholars between those who would emphasize a "but for" approach to causation, as opposed to those who prefer a "substantial-factor" formulation, mirrors the difference between the guilt-based and error-based approaches. Under the guilt-based approach, the defendant will be denied a remedy if the state can persuade the judge that the defendant would have been convicted despite the constitutional error. On the other hand, under the error-based approach, as long as the petitioner can persuade the judge that the error was a "substantial factor" contributing to the conviction, the judge will order a new trial without asking whether the defendant would have been convicted absent the error.

The difference between the two approaches to causation is highlighted by the difference between the majority and dissent in a Fourth Circuit death penalty case eventually decided en banc, *Cooper v. Taylor*.[102] In this murder trial, the key prosecution evidence was three confessions—two of which the jury heard through the testimony of the officers who allegedly heard them, and the third a tape-recorded confession that was played for the jury while they followed along with the transcript. On habeas review, the third confession was judged improperly admitted, as it was taken after the defendant requested the presence of a lawyer.[103]

For the majority, the "obvious power" of the two other confessions, along with the other "overwhelming evidence" of guilt, led to the conclusion that the guilty verdict "could not have more fairly represented the facts of record."[104] In other words, take away the improperly admitted confession, and the defendant is still clearly found guilty based on the properly admitted evidence. Indeed, this is quite consistent with the counterfactual "but for" test of causation reflected in the majority of the post-*Brecht* cases that employ a "guilt-based" approach to harmless error.

Rather than reason counterfactually, and consider the hypothetical trial without the error, the dissenting judges looked historically at the actual trial

---

[100] Robert N. Strassfeld, *If . . . : Counterfactuals in the Law*, 60 GEO. WASH. L. REV. 339, 346 (1992).

[101] *But see* Cooper, *supra* note 26, at 334 (arguing that both approaches require judges to engage in "problematic counterfactual reasoning").

[102] 103 F.3d 366 (4th Cir. 1996) (en banc).

[103] *Id.* at 367–68.

[104] *Id.* at 370–71.

that did take place, and asked whether the improperly admitted confession had a "substantial and injurious" influence on the verdict. In other words, they applied the "substantial factor" test for factual causation, where the question of what would have happened if the jury had only considered properly admitted evidence is irrelevant. Consistent with such an analysis, the dissenting judges observed that the taped and erroneously admitted confession was "the undeniable centerpiece of the state's case."[105] Judge Hamilton, one of the dissenting judges, went so far as to say that:

> [I]f the erroneous admission of the taped confession in Cooper's trial did not have substantial and injurious effect or influence on the jury's verdict, then no evidence erroneously admitted could ever be found to have such an effect or influence where there is independent evidence of guilt in the record.[106]

Judge Motz, another dissenter, took on the guilt-based "but for" approach quite directly, arguing that the majority "applies the wrong legal analysis—whether there was sufficient evidence of guilt without the erroneously admitted taped confession—to arrive at the wrong result."[107] Judge Motz described the approach taken by the majority as "excis[ing]" the defendant's taped confession ("but for") and considering only the properly admitted evidence.[108] With this approach, Judge Motz asserted, the majority "unequivocally demonstrates its lack of understanding of harmless-error review."[109]

Because the prosecution relied heavily on the taped confession at trial (referencing it fifteen different times during summation), the difference between the majority and dissent in understanding what factual causation means—or, more precisely, how to apply it in the harmless-error context—clearly accounts for the different conclusions reached by the majority and dissent in this case. Though the majority and dissent in *Cooper* put in stark relief the difference between the "guilt-based" approach and the "error-based" approach to harmless error, and their close cousins the "but for" test and the "substantial factor" test for causation, these differences are just beneath the surface throughout other post-*Brecht* cases.

C. A Hybrid Approach to Factual Causation in Criminal Trials

Having identified the normative ideal of determining cause-in-fact with respect to the constitutional error, the remaining question for harmless-error doctrine is a pragmatic one: what is the best way to determine factual cau-

---

[105] *Id.* at 374 (Hamilton, J., dissenting).

[106] *Id.* at 375.

[107] *Id.*

[108] *Id.* at 376.

[109] *Id.* at 382. Judge Motz also argued that the majority failed to follow the proper mode of analysis in "the most basic manner: the majority never once discusses the effect the error had on the jury at the trial that Cooper actually received." *Id.* at 376. Instead, she argues later, the majority "conducts its own hypothetical trial." *Id.* at 379.

sation when looking at constitutional errors in criminal trials? Professor Jane Stapleton, a leading torts scholar, describes the common ground of cause-in-fact in a way that is useful here: that the breach of duty "played a role in the history of the transition to that injury."[110] In this section, drawing on lessons from tort theory and doctrine, I argue that a hybrid of the two approaches, though closer to the "substantial-factor" end of the spectrum, should be used to determine whether the constitutional error played such a role.

If we view harmless-error analysis as the question of factual causation in criminal trials, then perhaps we ought to follow the latest *Restatement* and employ a "but for" approach to harmless error—whether we call it the "overwhelming evidence" approach, guilt-based approach, or "hypothetical outcome" approach. The "substantial factor" test has outlived its usefulness, or so the *Restatement* drafters tell us.[111]

The problem is that in cases of overdetermined causation, or multiple sufficient causes, the "but for" inquiry does not work very well in identifying legally relevant causal factors.[112] In such cases, scholars and judges agree, the "substantial factor" test works much better than the "but for" test in determining cause-in-fact,[113] and indeed convictions at criminal trials fit

---

[110] Jane Stapleton, *Legal Cause: Cause-in Fact and the Scope of Liability for Consequences*, 54 VAND. L. REV. 947, 958 (2001). But Stapleton is quick to point out that the rub arises when figuring out how to test this, referring to the traditional but-for test as "notoriously inadequate for this purpose," particularly in the case of overdetermined events. *Id.* Stapleton proposes a "targeted but-for test," which works by taking all the factors existing at the time of the transition to injury, removing one at a time, and considering whether the further removal of the factor in which we are interested leaves a set of remaining factors that would not have produced the injury. *Id.* at 959. If so, Stapleton says, then that factor played a role in the history of the transition to that injury. Stapleton describes this approach as building on the "necessary element of a sufficient set" test formulated by H.L.A. Hart and Tony Honore, and further developed by Richard Wright. *Id.* She also points out that this kind of testing is consistent with much scientific experimental design. *Id.* at 960 n.44.

[111] *See supra* note 98 and accompanying text.

[112] *See* RESTATEMENT (THIRD), *supra* note 96, § 27 reporters' note cmt. a ("There is near-universal support recognizing the inappropriateness of the but-for standard for factual causation when multiple sufficient causal sets exist."); David A. Fischer, *Successive Causes and the Enigma of Duplicated Harm*, 66 TENN. L. REV. 1127, 1129 (1999) ("In multiple sufficient cause cases, the 'but for' test cannot identify which event caused an injury because each of the multiple forces alone was sufficient to cause the injury."); Rue, *supra* note 97, at 2687 ("In the hard cases, however, where cause-in-fact is not intuitively clear, the 'but for' test can be ineffective.").

[113] *See* RESTATEMENT (THIRD), *supra* note 96, § 26 cmt. j (explaining that the "primary function" of the substantial factor test was to permit the fact finder "to decide that factual cause existed when there were overdetermined causes—each of two separate causal chains sufficient to bring about the plaintiff's harm, thereby rendering neither a but-for cause"); Fischer, *supra* note 112, at 1130 ("In lieu of 'but for' causation in multiple sufficient cause cases . . . courts require the jury to find that the tortfeasor's conduct was a 'substantial factor' in producing the harm."); Rue, *supra* note 97, at 2693 ("The strongest argument for the 'substantial factor' doctrine can be found in cases of multiple sufficient causes."); *see also* Zuchowicz v. United States, 140 F.3d 381, 389 n.6 (2d Cir. 1998) (Calabresi, J.) ("Many courts long ago abandoned the requirement of *but for* cause in situations where, since the negligence of any one of several defendants was sufficient to cause the harm, the negligence of none was its *necessary*

precisely in that category. In other words, the only category of cases where "substantial factor" is the best test to use in determining factual causation, according to leading commentators, is one that just happens to include harmless-error analysis.

Consider the paradigmatic example of "overdetermined harm," or multiple sufficient causes—the classic "two-fires" case familiar to first-year torts alums.[114] This case, credited with first adopting the substantial-factor test,[115] involved two separate fires that joined together and burned the plaintiff's property. Either of the fires was sufficient to cause the harm. A strict application of the "but-for" test would result in no liability for the defendant railroad that negligently started one of the fires. But this result of course seems anomalous, because it would mean that the innocent plaintiff would bear the full cost of the harm, despite at least one tortious actor. Moreover, it seems to make little sense to say that neither fire caused the damage.[116] As a result, the Minnesota Supreme Court in the two-fires case used the "substantial-factor" test as a substitute for the "but-for" test to determine cause-in-fact in this circumstance, leading to liability for the negligent railroad.[117]

Harmless-error analysis presents just such a "multiple sufficient causes" scenario. Consider first the extreme rarity of courts overturning convictions based on sufficiency of evidence grounds. Assume that courts are discharging the responsibility seriously in denying nearly all such claims, and ask why that is. Intuitively, the answer is that there is almost always more than enough evidence to convict; the question is whether the jury credits enough of it to get over the "beyond a reasonable doubt" threshold. When observed from a factual causation perspective, then, criminal trials are almost always overdetermined causation cases. The ones that are barely determined don't get brought to trial.

Rarely is any one piece of evidence necessary to satisfy an element of the crime. If there are three witnesses, and two pieces of physical evidence, removing one may still result in conviction. "But for" causation in the strict sense will not be satisfied. But that does not mean that that one piece of evidence did not "cause" the conviction. The jury may have discredited some witnesses' testimony, and credited others. But we would not say that the particular piece of evidence did not factually cause the conviction simply because the conviction *could have* occurred without it.[118] Moreover, a

---

cause.").

[114] Anderson v. Minneapolis, St. P. & S. S. M. Ry. Co., 179 N.W. 45 (Minn. 1920).

[115] *See* RESTATEMENT (THIRD), *supra* note 96, § 26 reporters' note cmt. j.

[116] *See id.* § 27 cmt. c ("Multiple sufficient causes are also factual causes because we recognize them as such in our common understanding of causation, even if the but-for standard does not.").

[117] *See id.* § 26 reporters' note cmt. j ("In the instance of multiple sufficient causes, however, the substantial-factor test can be useful because it substitutes for the but-for test in a situation in which the but-for test fails to accomplish what law demands.").

[118] Another way of getting at the same idea would be to consider the harm as a type of "lost oppor-

strict "but-for" test in this circumstance would provide perverse incentives for prosecutors who think they have strong cases (i.e., harmful error-proof) to commit or fail to avoid constitutional violations.[119]

Determining factual causation in a criminal trial with reference to any particular piece of evidence, then, presents the familiar problem from tort law of multiple sufficient causes. Strict application of the "but-for" test would lead to the anomalous result that perhaps none of the individual pieces of evidence, even if *all actually playing* a significant role in the jury's decision to convict, would be considered harmful, because the jury *could* have convicted on the basis of the remaining evidence. Like tort law, the "substantial-factor" test appears to fare better than the "but-for" test in determining the causal relationship in which we are interested.

But harmless-error analysis is also a particular kind of "multiple sufficient causes" case, where torts plaintiffs get somewhat less sympathy from the courts when they cannot satisfy the "but-for" test. It is a case where one of the "sufficient causal sets" is innocent, not tortious. That is, the prosecution's properly admitted evidence appropriately contributed to, and probably also "caused," the conviction.[120] Returning to the two-fires analogy, assume that one of the fires started from natural causes, not negligent conduct. Courts typically still find liability in this circumstance, though not as often as when both causes are tortious.[121] And one of the reasons for finding liability in a typical tort claim like this is absent in harmless-error analysis: when choosing between a tortious defendant and innocent plaintiff to bear the cost of the harm, equitable principles weigh in favor of liability. Here, though, society wants the petitioner to bear the harm (serve his prison sentence) if, absent the error, the jury would have convicted. Otherwise, petitioner would get an undeserved "windfall," in tort-law terms.[122] The case for simply using the "substantial-factor" test, then, is less strong than was initially apparent from the multiple sufficient cause analogy.

Perhaps an even more apt analogy than the two-fires case is the mixed-motive analysis used in constitutional tort and Title VII employment discrimination cases.[123] Under the regime established by *Mount Healthy City*

---

tunity" or "lost chance" of demonstrating innocence or being acquitted. This is analogous to tort cases where the harm is a lost opportunity for cure of a medical condition. *See id.* § 26 cmt. n.

[119] *Cf.* Guido Calabresi, *Concerning Cause and the Law of Torts: An Essay for Harry Kalven, Jr.*, 43 U. CHI. L. REV. 69, 86–87 (1976) (arguing that in multiple sufficient cause cases, it is "at the very least, doubtful whether blind adherence to the requirement that the victim prove a *but for* relationship serves the purposes of market deterrence").

[120] It has long been accepted that more than one factor may be considered the "cause" of harm. *See* RESTATEMENT (THIRD), *supra* note 96, § 26 cmts. c, d.

[121] *See id.* § 27 reporters' note cmt. d.

[122] *See* Stephen D. Sugarman, *Rethinking Tort Doctrine: Visions of a Restatement (Fourth) of Torts*, 50 UCLA L. REV. 585, 606 (2002) (pointing out that in such circumstances, holding the negligent party liable means putting the victim back "in a better position than he/she would have been in absent the defendant's negligence").

[123] I am grateful to Rebecca White and Cindy Estlund for suggesting this analogy in the Title VII

School District Board of Education v. Doyle,[124] a plaintiff alleging First Amendment-based retaliation must show that his constitutionally protected conduct was a "substantial factor" in the adverse decision taken against him. At that point, the burden shifts to the defendant, who can still avoid liability by showing that "it would have reached the same decision . . . even in the absence of the protected conduct."[125] This analytical framework was subsequently applied to other constitutional tort claims and adopted in the Title VII context as well.[126] But after the Supreme Court relied on *Mount Healthy* to justify a "but for" test for causation to determine Title VII liability,[127] Congress changed the rule as part of the Civil Rights Act of 1991. In that Act, Congress created a rule holding the defendant liable if the discriminatory reason was the "motivating factor" in the adverse employment action. However, Congress also allowed the employer to use the "same decision" affirmative defense—essentially disproving "but for" causation—to avoid a variety of remedies, including reinstatement.[128]

The nature of the employer's defense here is the same as the harmless-error defense used by the state—just as the employee would have been fired anyway, despite the discriminatory motive—the defendant would have been convicted anyway, despite the constitutional error. The task of the judge is also similar. In both cases, the judge has the task of discerning the relative importance of the innocent factor (proper evidence in harmless-error analysis; permissible reason for adverse employment action in Title VII analysis) and the tortious factor (constitutional violation in harmless-error analysis; impermissible motive such as race or gender in Title VII analysis, or protected speech in First Amendment retaliation cases) in bringing about the harm.

Based on these similarities, one could argue that courts should import this "modified but-for" approach to mixed-motive causation to the harmless-error context. But it is a more uncertain prospect to determine the state

---

context.

[124] 429 U.S. 274 (1977).

[125] *Id.* at 287; *see also* Vill. of Arlington Heights v. Metro. Hous. Dev. Corp., 429 U.S. 252, 270 n.21 (1977) (noting in dicta that *Mt. Healthy* approach would apply in equal protection context). The *Mt. Healthy* approach to causation in constitutional tort law is discussed in Michael Wells, *Three Arguments Against* Mt. Healthy*: Tort Theory, Constitutional Torts, and Freedom of Speech*, 51 MERCER L. REV. 583 (2000). *See also* Thomas A. Eaton, *Causation in Constitutional Torts*, 67 IOWA L. REV. 443 (1982).

[126] *See* Mark C. Weber, *Beyond* Price Waterhouse v. Hopkins*: A New Approach to Mixed Motive Discrimination*, 68 N.C. L. REV. 495, 522 (1990). For an overview of mixed-motive doctrine, see Peter Siegelman, Un-Muddling Mixed-Motives?: Protection for the Non-Exemplary Worker (draft of October 12, 2004), and for a critique of mixed-motive analysis, see Susan Bisom-Rapp, *Of Motives and Maleness: A Critical View of Mixed Motive Doctrine in Title VII Sex Discrimination Cases*, 1995 UTAH L. REV. 1029, 1074–76 (arguing that mixed-motive analysis undetrmines both the compensation and deterrence objectives of antidiscrimination legislation).

[127] Price Waterhouse v. Hopkins, 490 U.S. 228 (1989).

[128] Pub. L. No. 102-166, Sec. 107 (codified at 42 U.S.C. § 2000e-5(g)(2)(B)).

of mind of jurors, as compared to decisionmakers in the workplace, and therefore more difficult to determine whether the outcome would have been the same absent the impermissible factor. In an employment discrimination case, the fact-finder can listen to witness testimony and make direct credibility judgments. Whereas in harmless-error analysis, the jurors are neither deposed nor put on the stand, and their thought processes forever remain a source of pure speculation.

In my view, there is a better way to combine the two approaches to factual causation in this context: specifically, I propose a hybrid approach that employs historical reasoning, like the "substantial-factor" test, but considers the strength of the prosecution's case as a factor, like the "but-for" test.

For several reasons, a judge's harmless-error analysis ought to be historical, like the substantial-factor test, not counterfactual, like the "but-for" test. First, the plain language of *Brecht* and *Kotteakos* seems to support this kind of historical reasoning by asking the empirical question of the "effect" the error "had" on the verdict. Justice Breyer puts the historical question most plainly in *O'Neal* when he asks: "Do I, the judge, think that the error substantially influenced the jury's decision?"[129]

Second, Sixth Amendment values weigh heavily in favor of historical, not counterfactual, analysis. Under "but for" analysis, the judge must literally run a hypothetical trial to determine whether, under the counterfactual scenario, the defendant is still convicted. Though the historical inquiry is not without its own uncertainty, the question is what impact the error had on the actual jury that decided the defendant's case in accordance with his Sixth Amendment right to trial by jury.[130] One might argue that inevitably, the harmless-error analysis involves some appellate weighing of the evidence—if nothing else, to determine whether the error was "substantial" and "injurious," as *Brecht* and *Kotteakos* mandate—and so in terms of Sixth Amendment offensiveness the two tests differ in degree, but not in kind.[131] But this understates the difference in the reasoning involved in the two tests of factual causation.

Third, the empirical uncertainty surrounding causal relationships in criminal trials argues for historical reasoning, which though speculative, is

---

[129] O'Neal v. McAninch, 513 U.S. 432, 436 (1995).

[130] *See* Richard W. Wright, *Causation, Responsibility, Risk, Probability, Naked Statistics, and Proof: Pruning the Bramble Bush by Clarifying the Concepts*, 73 IOWA L. REV. 1001, 1041 (1988) (criticizing scholars who assume that "the causal inquiry is concerned with all the hypothetical scenarios that might have happened, rather than with determining the causal processes at work in the one scenario that did happen"). As Professor Leon Green put it, referring to the but-for test, "tests of this character have the same vice as any 'if,' or any analogy. They take the eye off the ball." Leon Green, *The Causal Relation Issue in Negligence Law*, 60 MICH. L. REV. 543, 556 (1962).

[131] Though disagreeing with the guilt-based approach, Chief Judge Harry Edwards of the D.C. Circuit acknowledged that "more often than not, we review the record to determine how we might have decided the case; the judgment as to whether an error is harmless is therefore dependent on our judgment about the factual guilt of the defendant." Edwards, *supra* note 28, at 1171.

grounded in the actual trial, as opposed to counterfactual reasoning, which imagines a hypothetical trial. As discussed in more detail below,[132] cognitive and experimental psychology, behavioral decision theory, evidence scholarship, and other disciplines have provided models for how jurors process information and make decisions. But none of the models have tremendous predictive power. That is, we have little ability to hypothetically go back and change one variable in the cauldron of evidence presented to the jury, and have any confidence in what the impact of that change might be. This level of uncertainty is a magnitude or two greater than the factual uncertainty presented in a typical tort case, for example, because our causal generalizations about the impact of evidence on juries are so poorly developed in comparison to cases of physical injury.[133] Relying on counterfactual reasoning under conditions of such empirical uncertainty carries too high a risk of inaccurate determinations of harm.

Finally, as we saw in the two-fires case, counterfactual reasoning is particularly poor in identifying factual causes in "overdetermined causation" cases. As a result, one leading torts theorist, Professor Stephen Perry, recently proposed a "historical worsening" test for harm in tort law.[134] Under the counterfactual test, the relevant interest subsequent to a particular event is compared to a hypothetical condition of that interest if the event had never occurred.[135] Perry's historical worsening test, however, is more straightforward: it simply compares the relevant interest subsequent to a particular event to the condition of the interest prior to the event. If the interest has been affected adversely, then harm has been caused.[136] In the harmless-error context, the relevant interest is the probability of conviction. If that probability increased as a result of the constitutional error, the error would be harmful under this type of historical worsening analysis.

Nonetheless, strict application of a "historical worsening" test does not fully capture the *Brecht* inquiry and the question of factual causation. Applied literally, it would mean that any worsening in the probability of conviction would lead to a conclusion of factual cause. But such a result is

---

[132] *See infra* Part IV.

[133] As Richard Wright has pointed out, causal analysis relies upon and is limited by "our empirical knowledge of the conditions that actually existed on the particular occasion and the possibly applicable causal generalizations." Wright, *supra* note 130, at 1042. Indeed, the limits to our empirical knowledge in both of these areas is what makes the harmless-error inquiry so difficult.

[134] *See* Stephen Perry, *Harm, History, and Counterfactuals*, 40 SAN DIEGO L. REV. 1283 (2003).

[135] *See, e.g.*, JOEL FEINBERG, *Wrongful Life and the Counterfactual Element in Harming*, *in* FREEDOM AND FULFILLMENT: PHILOSOPHICAL ESSAYS 3 (1992).

[136] In significant part, Perry reaches the conclusion that the historical worsening test is a superior alternative to determining harm by considering the two fires case. Under the counterfactual test, as we have discussed, one would reach the anomalous result that neither of the two fires caused harm to the property owner. To avoid such a result, Perry proposes historical worsening. *See* Perry, *supra* note 134, at 1286–87. *But see* John C.P. Goldberg, *Rethinking Injury and Proximate Cause*, 40 SAN DIEGO L. REV. 1315, 1320 (2003) (arguing that the two tests "overlap considerably, especially in the tort context").

inconsistent with *Brecht*'s description of the magnitude of the effect as "substantial and injurious," and risks an undesirable result in this situation of "one innocent, one tortious" sufficient causes.

Here is where an important aspect of the "but-for" test, or guilt-based approach to harmless error, comes in. To be sure, a significant part of assessing whether there was "substantial and injurious" effect is performed by looking at the error itself, but the judge should also assess the probability of conviction before the error. If the "innocent" cause—the properly admitted prosecution evidence—actually accounts for the conviction, then we ought not to be worried about accurate determination of guilt, and the conviction should be upheld. If the error occurred late in the trial, for example, when the jurors may have already come to a view of the case, then it may be right to say that the error did not "play a role in the history of the transition to the injury," to use Stapleton's terms. The problem is that in many cases, it is frequently difficult to assess the strength of the prosecution's case from the perspective of the jury.

Nonetheless, some assessment of the strength of the prosecution's case is appropriate and usually necessary. Indeed, even those that support the "error-based" approach have occasionally conceded as such.[137] The impact of the error necessarily depends in part on the strength of the prosecution's case versus the strength of the defense case, and how the error plays into that choice for the jury. So evidence of guilt cannot and should not be ignored. The existing evidence of guilt can and should be considered as *a* factor (but not *the* only factor) to make inferences about the causal effect of the error.

By combining this aspect of the guilt-based approach to harmless error, or the "but-for" analysis in torts, with the historical reasoning employed in the error-based approach, or the "substantial-factor" test in tort law, we can

---

[137] *See* Cooper, *supra* note 26, at 328–29 (pointing out that some judges may be using the strength of the evidence as a "proxy" for a conclusion that the jury was unaffected by the error); Edwards, *supra* note 28, at 1187 (acknowledging that the "presence of massive evidence of the defendant's guilt surely is one factor for a court to consider" in determining harmless error). Even Justice Brennan's dissent in *Harrington v. California* acknowledged that some appellate assessment of the remaining evidence was inevitable, arguing that the "focus of appellate inquiry should be on the character and quality of the tainted evidence as it relates to the untainted evidence and not just on the amount of untainted evidence." 395 U.S. 250, 256 (1969) (Brennan, J., dissenting). In several post-*Brecht* cases, judicial assessment of the *weakness* of the evidence led to a finding of harm. *See* Gray v. Klauser, 282 F.3d 633, 653 (9th Cir. 2002) ("Given the paucity in the physical and eyewitness testimony, one can infer that the evidence concerning Gray's anger toward his wife and motive for murdering her . . . was of great importance in the jury's deliberations."); Thomas v. Hubbard, 273 F.3d 1164, 1181 (9th Cir. 2002) (finding of harm based in part on the weakness of the case, specifically empirical assumption that "in a case in which the State's evidence consists largely of the uncorroborated testimony of a person who himself had both a motive and the opportunity to commit the crime, there is a greater likelihood that any error will be prejudicial"); Gall v. Parker, 231 F.3d 265, 320 (6th Cir. 2000) (concluding that, because the State had "no evidence to rebut [defendant's] showing of insanity," "the Confrontation Clause violation likely stood in the way of an acquittal for reason of insanity").

move past the doctrinal deadlock, and judges can more accurately achieve the shared normative ideal in harmless-error analysis of determining whether the error caused the conviction.

V. DEMONSTRATING CAUSATION: PRESCRIPTION FOR DOCTRINE AND PRACTICE

Thus far, causation has played both a positive and normative role. The positive account is that different understandings of the necessary causation for harmless error—"but for" or "substantial factor"—help explain different approaches taken in harmless-error analysis. And I have proposed a hybrid approach to factual causation to be used going forward.

In this section, I move from the theoretical to the practical, and explain what judges should look at in doing this type of analysis. This section addresses the central problem underlying harmless-error analysis that has been ignored by scholars struggling with the morass of standards and approaches: *How exactly do you figure out whether an error has caused a conviction*? In other words, settling on the analytic framework of causation is a useful and critical step, but it is not enough. To maximize our chances of accurate determinations of harm, we must turn our attention to what I call the "causal laws of criminal trials."

A. The Causal Laws of Criminal Trials

A reading of the post-*Brecht* cases reveals that harmless-error analysis is dependent on a series of assumptions about causal laws involving the impact of evidentiary and other factors on jury decisionmaking. But a look at research on jury decisionmaking reveals that these assumptions are remarkably ungrounded in empiricism. In this section, I provide a brief overview of the research on how jurors process information and reach decisions,[138] and conclude that the dominant view in the harmless-error cases of how jurors use pieces of evidence to reach a decision on guilt or innocence contradicts the findings of the empirical literature. This difference, I argue, calls into serious question the accuracy of harmless-error analysis. I then consider a few specific examples of empirical assumptions in the post-*Brecht* cases.

Determining factual causation, and therefore harm, requires a theory of how jurors make decisions, and, more specifically, how jurors use various pieces of evidence to reach a judgment on the defendant's guilt.[139] But as

---

[138] A full evaluation of the different models that are used to explain juror decisionmaking is beyond the scope of this paper.

[139] In this paper, I focus exclusively on the initial vote that jurors bring into to the deliberation room. Because research shows that the distribution of the individual jurors' first-ballot votes predicts the result in 90% of cases, this is a useful and legitimate focus of analysis for determining causation (and therefore harm) at criminal trials. *See* Dennis J. Devine et al., *Jury Decision Making: 45 Years of Empirical Research on Deliberating Groups*, 7 PSYCHOL. PUB. POL'Y & L. 622, 690 (2001) ("There are

Dow and Rytting have pointed out, "harmless error doctrine is not perspicuous even with respect to the process by which the hypothetical reasonable juror couples evidentiary propositions with a causal law in order to generate a verdict."[140]

Indeed, all available evidence indicates that judges are not very good at determining the impact of various pieces of evidence on jurors.[141] There has been little empirical research on this topic, and one of the aims of this Article is to lay the groundwork for more such research. In one of the few studies on this subject, researchers at the University of New Mexico found that lawyers (a proxy for judges) and laypersons (a proxy for jurors) varied widely in their assessments of the probative value of various pieces of evidence.[142]

Moreover, various lessons from cognitive psychology and behavioral decision theory point in pessimism's direction. First, hindsight bias is likely to be influential here—that is, the fact that a judge knows, considering the case retrospectively, that the defendant has been found guilty introduces a natural bias that this outcome was inevitable, regardless of the error.[143] The hindsight bias problem also relates to a base rate problem. Every criminal defendant that appellate judges see is guilty, a convicted criminal before the law.[144] In the experience of appellate judges, then, 100% of defendants that stand trial are convicted. But murder trials, which comprise the overwhelming majority of habeas cases, result in significantly

---

compelling data from numerous studies indicating that the verdict favored by the majority of the jury at the beginning of deliberation will be the jury's final verdict about 90% of the time.").

[140] Dow & Rytting, *supra* note 30, at 510. Richard Wright describes "causal generalizations" as "incompletely specified causal laws that list only some, not all, of the abstract antecedent conditions that would be found in the fully specified causal laws." Richard W. Wright, *The Efficiency Theory of Causation and Responsibility: Unscientific Formalism and False Semantics*, 63 CHI.-KENT. L. REV. 553, 559 (1987). Because we live "in the real world of imperfect knowledge," we must use causal generalizations to try to judge singular instances of causation. *Id.*

[141] Goldberg, *supra* note 29, at 430 (noting that one of the problems with appellate fact-finding is that "the appellate court is likely to be wrong").

[142] *See* Lee E. Teitelbaum et al., *Evaluating the Prejudicial Effect of Evidence: Can Judges Identify the Impact of Improper Evidence on Juries?*, 1983 WIS. L. REV. 1147, 1183 (concluding that "[t]he high incidence of significant differences between lawyers' and laypersons' evaluations of prejudice leaves it even more doubtful that appellate judges can tell when an item of proof was sufficiently weighty to move the jury from nonpersuasion to belief beyond a reasonable doubt").

[143] *See* Michael J. Saks & Robert F. Kidd, *Human Information Processing and Adjudication: Trial by Heuristics*, 15 L. & SOC'Y REV. 123, 144–45 (1981) (describing the hindsight bias).

[144] *See* Cooper, *supra* note 26, at 343 (noting that "the fact of the jury's guilty verdict is likely to affect the appellate judge's perception of the weight of the evidence against the defendant"); Kate Stith, *The Risk of Legal Error in Criminal Cases: Some Consequences of the Asymmetry in the Right to Appeal*, 57 U. CHI. L. REV. 1, 48 (1990) (noting the possibility of appellate prejudice against defendants who "have been 'found' to be criminals"); *see also* Saks & Kidd, *supra* note 143, at 150–51 (noting that the research on heuristics suggests that errors "are massively in the direction of being seduced by case-specific information and failing to employ base-rate information").

more acquittals than most trials.[145] Of course, another common cognitive bias—the attribution error—could cut the other way. Judges might see the error, and the resulting conviction, and be naturally inclined to see a causal connection where such a connection may be weak.[146]

Existing models in the evidence scholarship and literature on juror decisionmaking fall into roughly two categories: a probabilistic or algebraic model, versus the "story model." Under the algebraic or probabilistic model, various pieces of evidence are presented at trial. After each piece of evidence is presented, jurors assign it a certain weight based on the credibility of the evidence and its probative value bearing on guilt.[147] Under this model, based on Bayes' theorem of probability, each juror starts with an initial assessment of the probability that the defendant is guilty, then has an updated assessment after receiving additional evidence.[148] At the end of the trial, each piece of evidence can be added up, as in an algebraic equation, based on the juror's assessment of probative direction (guilt or innocence) and force, and if the probabilistic assessment of guilt exceeds the juror's understanding of guilt "beyond a reasonable doubt," (let us say 90%, for example), then the juror will vote for a guilty verdict.

A visual representation of this model—and a metaphor commonly seen in judicial opinions and scholarship—might be that of a scale where evidence supporting guilt is on one end, and evidence supporting innocence is on the other. Such a visual representation is, of course, behind the verbal formulation of "weighing the evidence," the dominant paradigm in the

---

[145] In 1996, 5% of murder defendants in the seventy-five largest counties were tried and acquitted, compared with only 1% of defendants for all serious felonies. William J. Stuntz, *Plea Bargaining and Criminal Law's Disappearing Shadow*, 117 HARV. L. REV. 2548, 2563 n.37 (2004) (citing BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS—2000, at 463 tbl.5.53 (Kathleen Maguire & Ann L. Pastore eds., 2001)). Stuntz argues that the relatively high acquittal rate is because prosecutors generally pursue every murder case they can, in part because of the public pressure to punish a homicide. *Id.*

[146] Other aspects of retrospective evaluation of the effect of various pieces of evidence on jurors are problematic, particularly the difference between the "cold record" and seeing the witnesses in person. Although social science research indicates that people tend to do a poor job of assessing witness credibility, they do use nonverbal cues to assess credibility and the probative value of the evidence. *See, e.g.*, Lance Stockwell & David C. Schrader, *Factors that Persuade Jurors*, 27 U. TOL. L. REV. 99, 109–10 (1995) (indicating that "nonverbal immediacy" was the most important factor distinguishing between winning and losing lawyers and witnesses). But these nonverbal cues are inaccessible to judges, who are therefore likely to guess incorrectly as to how much force jurors assigned a particular witness's testimony.

[147] For an early exposition of this view, see John Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN L. REV. 1065 (1968), and for a rejection of it, see Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1350–54 (1971).

[148] *See* Thomas Bayes, *An Essay Towards Solving a Problem in the Doctrine of Chances*, *in* PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY OF LONDON (1763), *available at* http://www.stat.ucla.edu/history/essay.pdf (last visited Jan. 20, 2005); Daniel J. Kornstein, *A Bayesian Model of Harmless Error*, 5 J. LEGAL STUD. 121, 125–27 (1976) (describing the Bayesian model of determining guilt).

harmless-error analyses.[149]

This Bayesian model is perhaps most clearly reflected in the majority opinion in *Cooper v. Taylor*, the en banc Fourth Circuit death penalty case discussed in Section III.B. In that case the majority concluded that the improperly admitted confession was harmless because the evidence in the case was "totally one-sided."[150] The majority then made the following baseball analogy:

> By analogy, the jury witnessed the government score 14 runs with its evidence and the defense score none. If, for the sake of argument, we were required to invalidate what we would expect Cooper to characterize as a government grand-slam home run, the remaining 10-0 score would still have left the jury's verdict the same. . . . [T]he dissent would somehow urge if the grand slam home run were disqualified and the resulting score were reduced 10-0, the guilty verdict is per se adversely affected. That, however, is not the law.[151]

Under this model, each piece of inculpatory evidence is added to the "score" on the prosecution's side, much like the variables intended to represent pieces of evidence in the Bayesian algebraic model. If the margin is sufficiently large (call it a prosecutorial blowout), then the jury should find guilt beyond a reasonable doubt.

But the dissenting judges took issue with this model, arguing that the majority's method of tallying up the score, besides trivializing the inquiry, ignored the "central role the taped confession played in the trial."[152] In other words, the dissenting judges took issue with the very nature of how the majority thought about the impact of evidence on the verdict.

Indeed, the current consensus in social science research is that this Bayesian model does a poor job of describing how jurors actually make decisions.[153] Rather, the "story" model does a better job of describing jurors' information processing, and how evidence affects the jurors' ultimate decision. First conceptualized by Bennett and Feldman, this model describes what jurors do during trials as "story construction."[154] That is, jurors hear

---

[149] *See supra* Part II; *see also, e.g.*, Hill v. Brigano, 199 F.3d 833, 847 (6th Cir. 1999) ("In light of the great weight of evidence against the defendant, we do not believe the introduction of these limited statements . . . had any significant influence on the jury's decision making process.").

[150] 103 F.3d 366, 370 (4th Cir. 1996).

[151] *Id.* at 370–71.

[152] *Id.* at 374 (Hamilton, J., dissenting).

[153] *See, e.g.*, Ronald J. Allen, *Factual Ambiguity and a Theory of Evidence*, 88 NW. U. L. REV. 604, 604 (1994) (explaining that the conventional Bayesian approach has been undermined by developments in cognitive psychology and research into jury decisionmaking which "made it rather plain that virtually no one thinks as the conventional legal theory requires"); Shari Seidman Diamond & Neil Vidmar, *Jury Room Ruminations on Forbidden Topics*, 87 VA. L. REV. 1857, 1861 (2001) (observing that Bayesian and averaging models have enjoyed "limited success as descriptions of how jurors actually decide cases," but that the story model comports better with empirical evidence).

[154] W. LANCE BENNETT & MARTHA S. FELDMAN, RECONSTRUCTING REALITY IN THE COURTROOM: JUSTICE AND JUDGMENT IN AMERICAN CULTURE 117 (1981) (pointing out that facts or evidence occupy

witnesses, receive physical evidence, and as they are listening, attempt to fit the evidence into a narrative that explains the events described. Seen this way, the adversarial trial is a battle of competing narratives, and the jurors' task is to assess the "relative plausibility" of each side's narrative.

Two cognitive psychologists, Nancy Pennington and Reid Hastie, have actually tested this theory and found it has significant explanatory power.[155] And the resulting shift among cognitive and experimental psychologists has influenced the direction of evidence theorists as well. Ronald Allen, a leading evidence scholar, has led the charge of "Bayesioskeptics," using Pennington and Hastie's work to develop a theory of "relative plausibility," or competing narratives, as a more realistic representation of jurors' cognitive processing.[156] On the Supreme Court, Justice Souter has indicated interest in the competing narratives model, and how it bears on the impact of evidence on the jury.[157]

Some courts conducting harmless-error analysis do explicitly consider the impact of the error on the "relative plausibility" of the defendant's narrative as compared to the prosecution's version of events.[158] But by and large, harmless-error doctrine and scholarship has not caught up to this

---

an inherently ambiguous position in trials, and that "[w]hat makes a fact or piece of evidence meaningful in a particular case is its contextual role in the stories that make up the case").

[155] Nancy Pennington & Reid Hastie, *Explaining the Evidence: Tests of the Story Model for Juror Decision Making*, 62 J. PERSONALITY & SOC. PSYCHOL. 189, 202 (1992) (concluding that research supports the claim that "stories are the mediating mental structures that cause decisions in the juror's judgment task"); Nancy Pennington & Reid Hastie, *A Cognitive Theory of Juror Decision Making: The Story Model*, 13 CARDOZO L. REV. 519, 536 (1991) (citing research showing that "story structures differed systematically for jurors choosing different verdicts") [hereinafter Pennington & Hastie, *A Cognitive Theory*].

[156] *See, e.g.*, Ronald J. Allen & Brian Leiter, *Naturalized Epistemology and the Law of Evidence*, 87 VA. L. REV. 1491, 1527–28 (2001) ("The critical insight of the relative plausibility theory is that legal fact finding involves a determination of the comparative plausibility of the parties' explanations offered at trial rather than a determination of whether discrete elements are found to a specific probability.").

[157] *See* Strickler v. Greene, 527 U.S. 263, 307 (1999) (Souter, J., dissenting) (arguing that undisclosed impeachment evidence could have led to a sentence other than death because of the "undeniable narrative force" of the witness's testimony (citing E. LOFTUS & J. DOYLE, EYEWITNESS TESTIMONY: CIVIL AND CRIMINAL 5 (1997) ("Research redoundingly proves that the story format is a powerful key to juror decision making"))); Old Chief v. United States, 519 U.S. 172, 187–89 (1997).

[158] *See* Caldwell v. Bell, 288 F.3d 838, 844 (6th Cir. 2002) ("We believe the instruction did particular damage by undermining Caldwell's alternative theory of the killing based on the claim of 'provocation.'"); Thomas v. Hubbard, 273 F.3d 1164, 1181 (9th Cir. 2001) (concluding that errors were harmful because they bore in part on defense's ability to "attack the prosecution's theory of the case" and limited defendant's opportunity to present evidence "in support of his principal defense"); Mach v. Stewart, 137 F.3d 630, 634 (9th Cir. 1998) (concluding that error bore on critical issue in sexual misconduct case of "whether the jury chose to believe the child or the defendant"); Taylor v. Singletary, 122 F.3d 1390, 1396 (11th Cir. 1997) (defendant's inability to introduce codefendant's testimony "essentially precluded him from putting on a defense"); Lyons v. Johnson, 99 F.3d 499, 504 (2d Cir. 1996) (finding harm when defendant was prevented from putting another man who may have physically resembled the defendant, and who two defense eyewitnesses said was "the actual perpetrator," before the jury).

trend. The Bayesian model remains dominant in harmless-error cases.[159] In determining harm, judges in the post-*Brecht* cases essentially take away the variable representing the error, asking the question: if you subtract that variable, does the jurors' probabilistic assessment of guilt still exceed the reasonable doubt threshold? This Bayesian approach is not an unreasonable way of thinking about the causal effect of a particular piece of evidence, but it doesn't quite get at the way jurors reason. Rather, in determining harmlessness, judges ought to move beyond considering the evidence's probative force in a vacuum, and instead consider the extent to which the evidence bears on the relative plausibility of the prosecution and defense versions of events.[160]

In short, in light of social science research which indicates that the "story model" more accurately describes the way jurors process information and reach a verdict, harmless-error analyses overuse the probabilistic or algebraic model, likely leading to mistakes in the harmless-error determination.

B. General Causation: Drawing Inferences from Empirical Research

Besides decisionmaking models like those described above, judges necessarily rely on a set of causal generalizations to determine the impact of an error on the verdict. In the following section, I provide a few specific examples of causal generalizations from the post-*Brecht* cases whose em-

---

[159] In *Arizona v. Fulminante*, Justice Rehnquist's opinion describes "trial error" subject to harmless-error analysis as evidence that can be "quantitatively assessed in the context of other evidence presented in order to determine whether its admission was harmless beyond a reasonable doubt." 499 U.S. 279, 308 (1991). The use of the term "quantitatively assessed" seems to imply a weighing or algebraic model, as opposed to the "relative plausibility" model of a trial. *See also* Stith, *supra* note 144, at 44 n.112 (noting that consideration of the strength of the remaining evidence of guilt is consistent with a Bayesian approach to harmless error).

[160] *See* Gray v. Klauser, 282 F.3d 633, 651 (9th Cir. 2002) (reasoning that excluded evidence was harmful because it "would have provided the jury with a possible answer to the critical question that arises whenever the defendant's defense is that he did not commit the crime: Then who did?"); Agnew v. Leibach, 250 F.3d 1123, 1133 (7th Cir. 2001) (concluding that error was not harmless based in large part on the fact that improper testimony related to the central issue in the case: "whether a robbery occurred or whether [defendant] was simply defending himself in a fight"); Maurino v. Johnson, 210 F.3d 638, 650 (6th Cir. 2000) (Holschuh, J., dissenting) (arguing that prosecutorial misconduct designed to undermine defendant's credibility went right to central issue in the case—intent—and "the defendant's own testimony regarding his intent was the foundation of his defense"); Barrett v. Acevedo, 169 F.3d 1155, 1170 (8th Cir. 1999) (en banc) (Gibson, J., dissenting) (arguing that error was not harmless because it was relevant to "perhaps the most crucial issue at trial"—whether the victim was murdered (prosecution's version) or committed suicide (defendant's version)); Scoggin v. Kaiser, 186 F.3d 1203, 1208 (10th Cir. 1999) (finding improper jury instruction on theft of merchandise harmless because "[p]etitioner's defense turned not on the nature of the property taken" but on his claim of misidentification); Buehl v. Vaughn, 166 F.3d 163, 177 (3d Cir. 1999) ("In this case, the nature of the prosecution's evidence and Buehl's defense rendered the error in this instruction harmless."); Harris v. Warden, 152 F.3d 430, 439 (5th Cir. 1998) (reasoning that an improper jury instruction on intent was harmless where intent was not in dispute).

pirical grounding appears to be in question.

1. The Impact of "Cumulative" Evidence.—One reason commonly given by judges for a determination of harmlessness is that the improperly admitted evidence was "cumulative" in light of other, properly admitted evidence—or, alternatively, that the improperly excluded evidence would have been cumulative.[161] On its face, the proposition seems straightforward and uncontroversial enough. Ostensibly, because the jury has already received evidence tending to prove what this evidence demonstrates, the error could not have had a "substantial and injurious effect" on the verdict. Upon closer look, however, this logical turn is less straightforward than it might appear. There is an empirical assumption embedded in this logic—that evidence that is "cumulative" necessarily has a negligible causal impact on the jury's assessment of guilt. That is, that if two witnesses have already testified that they saw Jim Jones shoot a gun on the night of June 16, then the third witness's testimony to the same effect is of little value to the jury.

But social science research indicates that this assumption may not be valid. For example, research on cascaded inferences by Schum and Martin indicates that "corroboratively or cumulatively redundant testimony" tends to be overvalued or "double counted."[162] Schum and Martin's subjects were asked to read transcripts in invented felony cases, and assess how strongly each piece of evidence supported the guilt or innocence of the defendant. The authors had the subjects assess the evidence in three different ways. When presented with evidence that was redundant in some sense, the subjects in Schum and Martin's experiments, using two out of those three methods, were consistent in assigning the same probative weight to the second (and corroborative or cumulative) item of testimony, as they did to the first.

In other words, though the evidence was "cumulative," it still had a

---

[161] *See* Braun v. Powell, 227 F.3d 908, 921 (7th Cir. 2000) (reasoning that the prosecution's failure to inform the jury of terms of cooperating witness's plea agreement was harmless because witness "was cross-examined for approximately a week, and the jury heard extensive evidence demonstrating his lack of credibility"); Evans v. Lock, 193 F.3d 1000, 1003–04 (8th Cir. 1999) (reasoning that improper identification testimony was "cumulative" of proper identification testimony by another witness); Bryson v. Ward, 187 F.3d 1193, 1206 (10th Cir. 1999) (finding harmlessness based in part on fact that excluded videotape "did not tend to establish any facts . . . that had not already been presented to the jury"); Pyles v. Johnson, 136 F.3d 986, 996 (5th Cir. 1998) (reasoning that juror's daytime visit to the crime scene was harmless because it was "largely duplicative" of evidence presented at trial); Craig v. Singletary, 127 F.3d 1030, 1040 (11th Cir. 1997) (reasoning that admission of confession was harmless because a properly admitted confession "overlapped the first" and was even more damaging to the defendant). *But see* Stapleton v. Wolfe, 288 F.3d 863, 868 (6th Cir. 2002) (ruling that admission of accomplice statement was harmful because the multiple accomplice statements "reinforced and corroborated each other" (quoting *Fulminante*, 499 U.S. at 299)).

[162] *See* David A. Schum & Anne W. Martin, *Formal and Empirical Research on Cascaded Inference in Jurisprudence*, *in* INSIDE THE JUROR 136, 165 (Reid Hastie ed., 1993). The concern that jurors "double count" redundant testimony is drawn from Richard O. Lempert, *Modeling Relevance*, 75 MICH. L. REV. 1021, 1057 (1977).

causal impact that was significant, or at least no less significant than if the evidence was the first of its kind. Two eyewitnesses with motives to lie and shifty demeanors means that the third eyewitness, disinterested and confident on the stand, might well have a "substantial and injurious effect" on the minds of the jurors even if the content of his testimony is exactly the same as that of the first two witnesses.[163] No doubt, there will be circumstances where the "cumulative" nature of the evidence will diminish its impact on the jury, but it is by no means a truism, as it appears to be treated in many judicial opinions.

2. *The Framing of Choices.*—Another example of a causal generalization with dubious empirical grounding is the effect of improper (but not chosen) options that were presented to the jury, and is illustrated in the Tenth Circuit death penalty case, *Hale v. Gibson*.[164] In *Hale*, the defendant was charged with murder and kidnapping. The prosecution told the jury that both counts carried the death penalty, and urged its imposition on both counts, but the kidnapping charge actually did *not* carry a possible death sentence. The jury returned a life sentence on the kidnapping count and a death sentence on the murder count. The reviewing court reasoned that the erroneous instruction on the kidnapping count was harmless because the jury "was given a full range of possible sentences and chose a permissible sentence under Oklahoma law—life in prison."[165] Furthermore, the court argued that there was no evidence that "the jury was influenced to give a life sentence simply because they were given the impermissible choice of giving a death sentence."[166]

The court was correct in a narrow sense. There was no evidence of direct or specific causation here—that is, evidence that the choice of a death sentence on the kidnapping charge influenced this jury to choose a life sentence. But as to general causation, there is plenty of evidence—ignored by this court—that cuts the other way. Social science research has shown that the same option is frequently evaluated more favorably when it is presented as intermediate in a set of options under consideration, compared to when it is at an extreme.[167]

---

[163] The Schum and Martin experiment assumed the same level of credibility for each witness; in an actual trial, of course, jurors will credit the testimony of a particular witness to a greater or lesser extent depending on a variety of factors, with the extent impossible to discern from an appellate judge's reading of the transcript. *See also* Devine et al., *supra* note 139, at 685 (describing study that found that the number of trial witnesses was positively related to the likelihood of conviction); Lempert, *supra* note 162, at 1054 (pointing out that the testimony of four eyewitnesses to a crime, though it could be termed cumulative, may have "substantial probative value since the probability of the same mistake being made by four witnesses with such different reasons to err appears small").

[164] 227 F.3d 1298 (10th Cir. 2000).

[165] *Id.* at 1325.

[166] *Id.*

[167] *See* Mark Kelman et al., *Context-Dependence in Legal Decision Making*, *in* BEHAVIORAL LAW & ECONOMICS 61 (Cass R. Sunstein ed., 2000). This phenomenon is referred to as a "compromise ef-

Hale v. Gibson is a perfect example of that phenomenon. A jury faced with a choice of 20 years in prison, life in prison, and death on the kidnapping charge may have been influenced to choose a life sentence in part because it was an intermediate option. If the choices were 10 years in prison, 20 years in prison, and life in prison, which would have been more consistent with the law, the jury might well have made a different choice simply because of the framing of the options.[168] This appears to be another case where the causal generalization that is determining harm lacks a strong empirical foundation.

3. Limiting or Ameliorating Instructions.—When an error occurs at trial, the judge, or the lawyers, often recognize the error when it happens. In such a case, or in the case where evidence is admitted but only for a specific purpose, judges frequently give the jury limiting or ameliorating instructions, telling them to disregard the error.

When the resulting conviction is challenged, reviewing judges frequently point to such instructions in concluding that the error did not have a substantial and injurious effect on the verdict.[169] The empirical assumption of causal relations here is straightforward: any prejudicial impact of the evidence is minimized by the judge's instructions. Such an assumption is consistent with—indeed a subset of—the broader maxim that jurors are presumed to follow the judge's instructions. The only problem is this: the bulk of research on this topic indicates that these sorts of instructions do not work and often backfire.[170] In other words, limiting or ameliorating instruc-

---

fect." *See also* Saks & Kidd, *supra* note 143, at 140–42 (describing the related phenomena of "anchoring"). As a practical matter, this would not have made a difference in this case (which may ultimately account for the outcome of this issue) because the jury imposed the death sentence on the murder conviction.

[168] *Cf.* Devine et al., *supra* note 139, at 670 (summarizing studies suggest that "allowing juries the opportunity to convict the defendant on a lesser charge has a substantial impact on their verdicts").

[169] *See* Payton v. Woodford, 258 F.3d 915, 918 (9th Cir. 2001) (concluding that no harm was done, based in part on the assumption that the prosecutor's incorrect statement of the law "was ameliorated by the court's admonishment"); McGhee v. Yukins, 229 F.3d 506, 514 (6th Cir. 2000) ("The trial court instructed the jurors to consider each defendant's confession only against that defendant, and the jury is presumed to have followed that instruction."); Thompson v. Borg, 74 F.3d 1571, 1576 (9th Cir. 1996) (ruling that judge's admonition to disregard extraneous evidence of prior plea "adequately told" jurors to decide the case on the evidence); Nethery v. Collins, 993 F.2d 1154, 1159 (5th Cir. 1993) ("Nor are we prepared to say that this assumed error was not corrected by the court's curative instruction."); *see also* Devine et al., *supra* note 139, at 686 ("Courts have thus implicitly accepted the notion that jurors can and do heed the direction of the judge, but social scientists have been more skeptical and have sought to determine empirically if jurors do in fact disregard inadmissible evidence."). *But see* Maurino v. Johnson, 210 F.3d 638, 650 (6th Cir. 2000) (Holschuh, J., dissenting) (concluding that "the standard cautionary instruction was inadequate to 'unring the bell'"); Hill v. Turpin, 135 F.3d 1411, 1419 (11th Cir. 1998) ("Nor can we say that the trial court's valiant and well-intentioned attempt to remedy the . . . error through curative instructions eliminated the taint created by the prosecutor.").

[170] *See* Devine et al., *supra* note 139, at 666 ("In general, limiting instructions have proven to be ineffective and have even been associated with a paradoxical increase in the targeted behavior."); Phoebe C. Ellsworth & Alan Reifman, *Juror Comprehension and Public Policy: Perceived Problems and Pro-*

tions by the judge frequently *increase,* rather than decrease, the weight that jurors give to the evidence. Contrary to notions of causal effect that appear in judicial harmless-error analyses, the presence of such instructions ought to weigh on the side of harmfulness, not harmlessness.

What are we to take from these examples? One conclusion might be to avoid making any such assumptions at all. But this seems difficult, if not impossible, in determining causal impact. Relying on such causal generalizations is necessary and appropriate—in tort-law terms, this is the "general causation" inquiry—but the generalizations themselves must be more empirically grounded than is currently the case.

In fact, there is a clear analog in tort law in toxic tort cases, where evidence of causation in specific instances is difficult to demonstrate.[171] Consider cases where workers allege that exposure to asbestos caused them to develop cancer. There simply is no medical test that can demonstrate the causal link with any degree of certainty, so courts have allowed epidemiological evidence—that is, studies of the effect of certain chemicals on populations, as opposed to direct medical evidence that the specific plaintiff was harmed by the chemical—to allow for an inference of specific causation.[172] In such cases, evidence of general causation can help make up for lack of proof of specific causation, which simply cannot be expected given the circumstances.[173]

---

posed Solutions, 6 PSYCHOL., PUB. POL. & L. 788, 796 (2000) (same); Saul M. Kassin & Holly Sukel, *Coerced Confessions and the Jury: An Experimental Test of the "Harmless Error" Rule*, 21 LAW & HUM. BEHAV. 27, 43 (1997) (finding in study that mock jury participants appeared to rely on confession even when the judge ruled it inadmissible and admonished the jury to disregard it and citing other studies with similar findings); William S. Laufer, *The Rhetoric of Innocence*, 70 WASH. L. REV. 329, 364–65 (1995) (summarizing empirical evidence on jurors' lack of understanding of instructions); *see also* Thompson v. Borg, 74 F.3d 1571, 1581 (9th Cir. 1996) (Reinhardt, J., dissenting) (arguing that instructing a jury to ignore prior acts information in judging the offense at trial "is to ask human beings to act with a measure of dispassion and exactitude well beyond mortal capacities" (citations omitted)); *cf.* Yohn v. Love, 76 F.3d 508, 524 (3d Cir. 1996) ("Because of all the controversy over the tape throughout the trial, the jury must have believed it was an important piece of evidence.").

[171] *See* Bert Black & David E. Lilienfeld, *Epidemiologic Proof in Toxic Tort Litigation*, 52 FORDHAM L. REV. 732 (1984); Mark Geistfeld, *Scientific Uncertainty and Causation in Tort Law*, 54 VAND. L. REV. 1011, 1013–14 (2001) (noting that because many potentially hazardous substances have not been epidemiologically studied, "evidentiary gaps" exist with tort law that raise questions about special evidentiary rules for establishing causation).

[172] *See* RESTATEMENT (THIRD), *supra* note 96, § 28 cmt. a ("Courts generally permit juries to infer specific causation from a group study when the study finds that exposure to the agent causes at least a doubling in the incidence of disease in a group exposed to the agent compared to a group that was not exposed."); *see also id.* § 28 reporters' note cmt. c ("Applying the results of group studies to assess the probability of causation in an individual has become accepted."). I have previously discussed the analogy of epidemiological evidence in tort cases to prove causation in the context of workplace injuries. *See* Jason M. Solomon, Note, *Fulfilling the Bargain: How the Science of Ergonomics Can Inform the Laws of Workers' Compensation*, 101 COLUM. L. REV. 1140, 1171–72 (2001).

[173] *See, e.g.*, David W. Robertson, *The Common Sense of Cause in Fact*, 75 TEX. L. REV. 1765, 1774–75 (1997) ("The central idea is that when a defendant has engaged in conduct that we consider to be wrongful in major part because such conduct often leads to the kind of harm the plaintiff has suf-

Similarly, in retrospectively determining the effect of error on jurors, a variety of factors, including the time lag since the trial and the impermissibility of direct inquiry into the jurors' states of mind, makes direct evidence of causal impact rare. As a result, drawing inferences from studies on the impact of different kinds of evidence on jurors is appropriate, and preferable to mere speculation which, as the examples above demonstrate, is likely to result in inaccurate determinations of harm.

C. Specific Causation: Looking to Evidence of Influence

Harmless-error analysis has a methodological ideal to match the normative ideal of accurate determinations of factual causation. Methodologically, if judges could go back in time and read the minds of the jurors to determine the effect of the error upon the jurors' assessment of guilt, that would be the way to determine whether the error was harmless. If the error was "substantial" (non-negligible) and "injurious" (in the direction of guilt), then the error caused constitutional harm. But until time travel or neuroscience improves, we are stuck with trial transcripts and causal generalizations. In the face of this dilemma, and in the absence of much guidance from scholars or the Supreme Court, many judges have quite understandably taken to looking solely at the untainted evidence of guilt. Although this is an important part of determining factual causation, as I argued earlier, other evidence—evidence of influence—is underutilized by judges and is an even better indicator of causation in this context.

"Evidence of influence" is composed of several different factors that provide indirect or circumstantial evidence of: (1) the closeness of the case; (2) the centrality of the error to the jury's consideration of the case; and (3) the impact (substantial or not) of the error on the minds of the jury. These factors are occasionally but infrequently considered by judges in doing harmless-error analysis, as my empirical analysis of the post-*Brecht* cases revealed.

First, judges ought to look at the length of jury deliberations. This is frequently knowable simply from looking at the trial transcripts. Moreover, the prosecutor in the case can be asked to submit an affidavit when the claim is adjudicated in federal district court on how long the deliberations lasted. Increased attention to this factor will no doubt induce state trial judges and lawyers to ensure that the length of deliberations is reflected in the record. The length of deliberations is likely to be a much better indicator of the closeness of the case, than appellate judges' weighing of the evidence based on the trial transcript and the parties' briefs. If the deliberations lasted one day versus one week, that ought to be a relevant factor in determining whether the error caused the conviction. But in the post-*Brecht* cases, the length of jury deliberations was only mentioned in

---

fered, we are rightfully impatient with the defendant's claim that plaintiff cannot prove that the conduct caused the harm on this occasion.").

thirteen percent of cases.[174] This kind of analysis could also apply to referencing what happened in a previous trial, if the evidence presented was similar, for evidence of the closeness of the case.[175]

Second, judges ought to look at what the jury asked for during jury deliberations to see on what issues and evidence the jury was focused.[176] For example, if the jury asked for readbacks of testimony at issue in harmless-error analysis, that should presumptively be strong evidence of causation.[177] On the other hand, if they asked to see three critical pieces of testimony, and the erroneous piece was not one of them, that should weigh on the side

---

[174] *But see* Parker v. Gladden, 385 U.S. 363, 365 (1966) (emphasizing the fact that the jury deliberated for twenty-six hours in finding prejudice); French v. Jones, 282 F.3d 893, 902 (6th Cir. 2002) (finding that improper supplemental jury instruction caused harm when during three days of deliberations, jury sent out three "increasingly emphatic notes" saying that they were deadlocked); Karis v. Calderon, 283 F.3d 1117, 1140 (9th Cir. 2002) (noting that "even with the weak mitigation evidence that was presented, the jury was out for three days before rendering its verdict" in finding prejudice from ineffective assistance); Murtishaw v. Woodford, 255 F.3d 926, 968, 973 (9th Cir. 2001) (determining harm from erroneous jury instruction by relying in part on fact that jury took over two days to deliberate, twice asked for the definition of first-degree murder, asked to review testimony on defendant's alleged intoxication, and asked for readback of critical witness's testimony); Sassounian v. Roe, 230 F.3d 1097, 1110 (9th Cir. 2000) ("Lengthy deliberations preceding the misconduct and a relatively quick verdict following the misconduct strongly suggest prejudice." (citation omitted)); Tuggle v. Netherland, 79 F.3d 1386, 1393 (4th Cir. 1996) (including "any indications that the jury was hesitant or entertained doubt in reaching its sentencing determination" as a factor in determining harm during capital sentencing proceeding).

[175] *See* Newman v. Hopkins, 192 F. 3d 1132, 1137 (8th Cir. 1999) ("It is of more than passing significance that [defendant's] first trial ended in a mistrial because the jury could not reach a unanimous verdict."); Bonner v. Holt, 26 F.3d 1081, 1083–84 (11th Cir. 1994) (determining key evidence of influence in finding harm was that there were two sets of jury deliberations in which "the only difference in the evidence before the jury was the evidence admitted in error"); Kyles v. Whitley, 5 F.3d 806, 832–34 (5th Cir. 1993) (King, J., dissenting) (arguing that confidence in guilty verdict was undermined in part because defendant's first jury, hearing "essentially identical" evidence to that offered at the second trial, deadlocked on the question of guilt and stating "the fact that one or more jurors at [defendant's] first trial were not convinced beyond a reasonable doubt of his guilt is significant in assessing the force of [defendant's] case or, alternatively, the weaknesses in the State's case"). *But see* Kooce v. Pepe, 99 F.3d 469, 476 (1st Cir. 1996) ("[Defendant] cites no authority for his premise that the mistrial in his first trial should shade [the court's] reasoning in this case.").

[176] *See, e.g.*, Wray v. Johnson, 202 F.3d 515, 530 (2d Cir. 2000) (finding of harm from improper "showup identification" based in part on the fact that one of jury's "few requests during deliberations" was for a rereading of this testimony); Thompson v. Borg, 74 F.3d 1571, 1584 (9th Cir. 1996) (Reinhardt, J., dissenting) (noting that the jury "deliberated for three days and continuously grappled with the self-defense issues—issues central to Thompson's primary defense theory—sending numerous questions and requests for assistance to the trial court"); Lawson v. Borg, 60 F.3d 608, 612 (9th Cir. 1995) (noting that the jury's questions to the court indicated that "use of force was the subject of extensive deliberations"); Castro v. Oklahoma, 71 F.3d 1502, 1516 (10th Cir. 1995) (finding, in a review of a denial of funds for expert psychiatrist during capital sentencing proceeding, that the jury question of "Exactly what is meant by a life sentence?" indicated that it was "not self-evident" to the jury whether death penalty was warranted).

[177] The length of the testimony could also be used in certain circumstances to draw inferences about its impact. *See, e.g.*, Vanderbilt v. Collins, 994 F.2d 189, 199 (5th Cir. 1993) (concluding that the fact that witness's testimony "was about four times that of the states' other five witnesses combined" necessarily suggests that testimony would have "substantial impact" on the jury).

of harmlessness.

Third, the opening and summations ought to play a critical role. According to the "story model," the competing narratives—presented most coherently in the opening and closing statements—are influential in determining the jury's verdict. Though the research is mixed on how much sway opening and closing statements have over the jury in and of themselves, they are likely to frame the way the jury thinks about the case.[178] Viewing the trial as one of competing narratives, the opening statements and summation would play a critical role in determining the role that the error played in each side's story.[179] Whether the evidence or issue that is constitutional error is highlighted in the openings and closings ought to be a key factor in determining whether the error caused harm.

Attention to the placement of the evidence within the trial could also be used as evidence of influence, something rarely done in current harmless-error analysis.[180] Social science research indicates that evidence presented at the beginning and end of the trial is likely to be more influential than evidence presented in the middle of the trial.[181] And the research on coherence-based reasoning gives further support to the importance of evidence at the beginning of trial.[182] Finally, in appropriate circumstances, the

---

[178] *See* Shari Seidman Diamond et al., *Criminology: Juror Reactions to Attorneys at Trial*, 87 J. CRIM. L. & CRIMINOLOGY 17, 28 (1996) (describing the mixed evidence on opening statements).

[179] *See, e.g.*, Karis v. Calderon, 283 F.3d 1117, 1140 (9th Cir. 2002) (relying in part on the fact that the prosecutor "repeatedly stressed the absence of any testimony of mitigation throughout the penalty phase argument" in finding of prejudice from ineffective assistance of counsel); Ghent v. Woodford, 279 F.3d 1121, 1131 (9th Cir. 2002) (finding of harm based in part on prosecution's heavy reliance on witness' testimony during both opening and closing arguments); Moon v. Head, 285 F.3d 1301, 1317 (11th Cir. 2002) (ruling in a case where *Brecht* standard was clearly outcome-determinative that admission of a conviction subsequently vacated was harmless because the prosecution "confined its comments" regarding the vacated convictions to approximately five pages in the forty-six page transcript of the closing argument at capital sentencing); Wray v. Johnson, 202 F.3d 515, 529 (2d Cir. 2000) (noting that in summation, prosecutor "launched directly and repeatedly" into improper identification of defendant at precinct); Woods v. Johnson, 75 F.3d 1017, 1032 (5th Cir. 1996) (finding harmlessness in part because "[t]he complained of references" constitute "less than a tenth" of prosecutor's sentencing argument).

[180] *But see* Strickler v. Greene, 527 U.S. 263, 302 (1999) (Souter, J., concurring in part and dissenting in part) (arguing the credibility of witness's testimony was particularly important because witness was the first to describe defendant in any detail, "thus providing the frame for the remainder of the story the prosecution presented to the jury"); *Ghent*, 279 F.3d at 1131 (inferring the importance of psychiatrist's testimony from the fact that "the State reordered its proof in the special circumstances retrial so as to make [the psychiatrist] its second witness in its case-in-chief (instead of using him only as a rebuttal witness, as it did at the first trial)").

[181] *See* Pennington & Hastie, *A Cognitive Theory*, *supra* note 155, at 542 (presenting study demonstrating that presentation order of evidence affects verdict decisions dramatically); Michael J. Saks, *What Do Jury Experiments Tell Us About How Juries (Should) Make Decisions?*, 6 S. CAL. INTERDISC. L.J. 1, 25 (1997) ("If nothing else is clear from the order effects research, it is that information presented in the middle has less impact than information presented at the beginning or the end.").

[182] *See, e.g.*, SAUL M. KASSIN & LAWRENCE S. WRIGHTSMAN, THE AMERICAN JURY ON TRIAL: PSYCHOLOGICAL PERSPECTIVES 134–37 (1988) (explaining that considerable research supports the idea that initial impressions of jurors persevere even when presented with contrary evidence); Laufer, *supra*

District Court can further develop the record on whether errors influenced the jury,[183] although inquiries into the jury's deliberative process are prohibited.

VI. CONCLUSION

As the empirical analysis in Part III revealed, harmless-error analysis is breaking down in the midst of doctrinal deadlock over the proper approach, preventing headway on the even more difficult task of arriving at an answer in determining harm. With the confusion in habeas and harmless-error jurisprudence after *Brecht* and the Anti-terrorism and Effective Death Penalty Act of 1996, the central inquiry—how to determine whether an error is harmless—has gotten lost in the shuffle. This Article has tried to shift the focus back to that important question.

The question of factual causation is the shared normative ideal at the heart of both doctrinal approaches. To make this causal determination, I hope that courts will use a hybrid approach that uses the historical (and not counterfactual) reasoning of the "substantial factor" test, while retaining some evaluation of the strength of the prosecution's case—the centerpiece of the "but for" analysis—as one of several factors to weigh in assessing harm. Under such an approach, the judge would also consider three factors, broadly defined: (1) general causation, using empirical research on how various types of evidence impact jurors; (2) specific causation—that is, evidence of influence on the jury such as the length of deliberations, or requests for the readback of testimony; and (3) consideration of the relevance of the error to the competing narratives presented to the jury, given the research supporting the "story model" in explaining how jurors reach a verdict.

Further empirical research is needed, both on jury decisionmaking and on judges' abilities to assess the impact of various errors on jurors. Through such research, scholars and members of the bar can determine whether any mistakes in assessment by judges are systematic and can be corrected. But the Federal Judicial Center, bar associations, and others can and should take steps now to make the existing empirical research on juries more widely available to judges and their staffs, as well as to criminal appellate lawyers and pro se petitioners.

Despite the overwhelming percentage of criminal cases that are now resolved by pleas, and the scholarly attention that has followed, these harm-

---

note 170, at 399–400 (same).

[183] *See, e.g.*, Fullwood v. Lee, 290 F.3d 663, 683 (4th Cir. 2002) (remanding case for determination of whether jury was improperly made aware of information regarding defendant's prior death sentence); Hegler v. Borg, 50 F.3d 1472, 1474–75 (9th Cir. 1995) (noting that district court judge examined court reporter and four jurors to determine effect of defendant's absence during readback of testimony); Duckett v. Godinez, 67 F.3d 734, 749 (9th Cir. 1995) (remanding case for evidentiary hearing on whether defendant's shackles were visible to the jury and "how onerous" they were during trial).

less-error cases remain a critical slice of our criminal justice system. They are the cases where the defendant has decided to make the prosecution prove its case, perhaps because he is innocent, and where a judge has determined that the truth-furthering constitutional protections have broken down. Here is one identifiable class of cases, then, where we risk wrongfully convicted men and women remaining imprisoned, while the actual perpetrators go free. In light of the stakes, it is worth our attention to see whether the law of wrongs can help solve the riddle of harmless error.