

U.C.L.A. Law Review

“Sociological Gobbledygook”: *Gill v. Whitford*, *Wal-Mart v. Dukes*, and the Court’s Selective Distrust of “Soft Science”

Kyle Reinhard

ABSTRACT

In the last decade, the Supreme Court has heard a number of high-profile cases holding potential for major progressive reform which hinge crucially on the Justices accepting, or at least not rejecting, the findings of social science experts. Time and again, however, the Court has found ways to ignore, minimize, or misconstrue the work of highly-qualified social scientists, with Chief Justice Roberts even stating last Term that a plaintiff’s model looked to him like “sociological gobbledygook.” This Comment explores why the judiciary is especially skeptical of plaintiffs whose conduit theories of liability rely on cutting—or bleeding—edge social science, and how methodological aspects of these “soft” sciences frustrate deeply entrenched legal norms and judicial expectations for proper expertise and proof.

Because the *Daubert* “gatekeeper” stage is an imperfect forum, in which judges regularly scrutinize complex proof they lack the epistemic competence to evaluate independently, courts must turn from actual substance of testimony to secondary “proxy” indicia for its reliability. Utilizing proxy knowledge in this manner rewards the traditional at the expense of the new and grants undue weight to proxies with no correlation to a proof’s validity, such as its conformance with individual judges’ preexisting cultural agendas. These issues all negatively and disproportionately impact the social sciences. This institutional aversion halts desirable legal development, as a number of long-vexing systemic harms can only capably be measured with social science, and if such proof is not embraced, important rights will be left perpetually unvindicated.

This Comment argues not for an uncritical acceptance of all social science, but instead for the judiciary to acknowledge and respect it more fully as a source of knowledge. In conclusion, it calls for courts to embrace the trial setting as a means for developing more certain answers to the types of highly challenging questions that social science attempts to answer. All told, this approach will lead to desirable, common sense outcomes rather than “law based on self-induced blindness, on flagrant contradiction of known fact.”



AUTHOR

J.D., UCLA School of Law, Class of 2021. Sincere thanks to Professor Richard Re, Dean Jennifer Mnookin, and Cheryl Kelly Fischer for incredibly thoughtful contributions and feedback; Professor Sunita Patel, for encouragement and mentorship through the Library Research Assistant Program; and Bethanie John, who listens to my ideas even when they do not make sense yet.

TABLE OF CONTENTS

INTRODUCTION..... 702

I. “WELCOME TO WAL-MART”: AN INTRO TO *DUKES*..... 710

II. SOCIAL SCIENCE AND THE PROBLEMS OF FOREIGN EXPERTISE AND UNFAMILIAR PROOF 719

 A. Shortcuts to Knowledge: The Judicial System’s Reliance on Secondary Proxies
 for “Actual” Understanding..... 721

 B. The Impact of Culture as a Proxy in Judicial Decisionmaking..... 724

 C. The Judiciary’s Normative Expectations for Good Experts and Expertise..... 732

 1. The Archetypal Expert 732

 2. The Perfect Proof..... 734

III. TRANSLATING BETWEEN THE LANGUAGES OF SCIENCE AND LAW..... 742

CONCLUSION: CONSTRUCTING “CERTAINTY” THROUGH TRIAL 752

INTRODUCTION

Few matters on the U.S. Supreme Court's recent docket have sparked more in the progressive imagination than the partisan gerrymandering case *Gill v. Whitford*.¹ As just the fifth case argued before the Court in its 2017 Term, *Gill* attracted months of media and scholarly attention prior to decision on June 18, 2018, with many hoping it could turn what seemed a growing tide of lopsided voting maps cropping up nationwide.²

The case concerned the Wisconsin state legislature's "virtually without historical precedent" 2010 redistricting plan,³ which in 2012 resulted in Democratic candidates winning only 39 percent of the available seats in the state legislature despite having received 53 percent of the votes cast.⁴ Although the Court had considered numerous partisan gerrymandering claims in the past, no previous case commanded a majority for the question of the appropriate test

-
1. 138 S. Ct. 1916 (2018). As the case worked its way up the appellate ladder, rumblings about *Gill*'s generational potential began to emerge, and by the time the Court heard arguments, the buzz was considerable. E.g., Steven D. Schwinn, *Gill v. Whitford*, in *DIV. OF PUB. EDUC., AMERICAN BAR ASS'N, PREVIEW OF THE UNITED STATES SUPREME COURT CASES 26, 28 (2017)* ("It's hard to overstate the potential significance of this case, especially given today's political climate."); Russ Feingold, *Will the Fate of America's Democracy Be Decided By This US Supreme Court Case?*, *GUARDIAN* (Oct. 3, 2017), <http://www.theguardian.com/commentisfree/2017/oct/03/america-democracy-us-supreme-gerrymandering> [<https://perma.cc/YWS2-AYQ8>] (arguing that the case "could transform US politics"); *OT2017 #1: "A Very Nice Office"* at 37:41, *FIRST MONDAYS* (Oct. 2, 2017), <http://www.firstmondays.fm/episodes/2017/10/2/ot2017-1-a-very-nice-office> [<https://perma.cc/W4SL-V75G>] ("Gotta be the biggest case of the sitting, maybe the biggest case of the term Maybe the biggest case of your life!").
 2. In recent years, partisan gerrymandering has drawn increased attention and has been labeled a "second generation voting barrier" now frustrating the spirit, if not the letter, of the Voting Rights Act of 1965, Pub. L. No. 89-110, 79 Stat. 438 (1965). See *Shelby County v. Holder*, 570 U.S. 529, 566 (2013) (Ginsburg, J., dissenting) ("[D]espite this progress, 'second generation barriers constructed to prevent minority voters from fully participating in the electoral process' continu[e] to exist . . ."). Along with the map at issue in *Gill*, some other recent, high-profile cases have included disputes over North Carolina's state legislative district (*Common Cause v. Rucho*, 279 F. Supp. 3d 587 (M.D.N.C. 2018)); Maryland's 6th congressional district (*Benisek v. Lamone*, 266 F. Supp. 3d 799 (D. Md. 2017)); and Georgia House districts 105 and 111 (*Georgia State Conference of the NAACP v. Georgia*, 312 F. Supp. 3d 1357 (N.D. Ga. 2018)).
 3. Expert Report of Simon Jackman at 63, *Whitford v. Nichol*, 151 F. Supp. 3d 918 (W.D. Wis. 2015) (No. 3:15-CV-00421), 2015 WL 10091020 (assessing the current Wisconsin State Legislative Districting Plan).
 4. Barry C. Burden, *Everything You Need to Know About the Supreme Court's Big Gerrymandering Case*, *WASH. POST* (Oct. 1, 2017, 11:03 AM), <https://www.washingtonpost.com/news/monkey-cage/wp/2017/10/01/everything-you-need-to-know-about-the-supreme-courts-big-gerrymandering-case> [<https://perma.cc/S232-T3Q6>].

to apply.⁵ Given its easy-to-digest facts coupled with the emergence of increasingly sophisticated social science metrics for measuring partisan vote dilution, *Gill* seemed a prime opportunity for the Court to at long last accept a justiciable⁶ test for policing partisan gerrymandering: the Efficiency Gap Method (EG).⁷

At oral arguments, however, Chief Justice Roberts proved unimpressed. In a sequence foreshadowing the Court's ultimate holding eight months later, Roberts described the gut-level flaws inherent to tests like EG, which seek to generate legal significance, as it were, from out of a calculator:

[I]f you're the intelligent man on the street and the Court issues a decision, and let's say, okay, the Democrats win, and that person will say: "Well, why did the Democrats win?" And the answer is going to be because EG was greater than 7 percent, where EG is the sigma of party X wasted votes minus the sigma of party Y wasted votes over the sigma of party X votes plus party Y votes.

-
5. See, e.g., *League of United Latin Am. Citizens v. Perry*, 548 U.S. 399, 414 (2006) (noting that the "disagreement over what substantive standard to apply" that was evident in *Bandemer* "persists"); *Davis v. Bandemer*, 478 U.S. 109, 116–17 (1986) (plurality opinion) (agreeing that the case was justiciable but splintering over the proper standard to apply).
 6. In 2004, the Court declared partisan gerrymandering claims nonjusticiable given that "no judicially discernible and manageable standards for adjudicating such claims exist." *Vieth v. Jubelirer*, 541 U.S. 267 (2004). However, in a concurrence which left mathematicians and academics scrambling for the next decade, Justice Kennedy opined "[t]hat no such [workable] standard has emerged in this case should not be taken to prove that none will emerge in the future." *Id.* at 311 (Kennedy, J., concurring). Given his seeming openness to a test, courting Justice Kennedy's swing vote may have seemed the most realistic path forward for voting rights advocates. See Jessica Kwong, *How Anthony Kennedy May Kill the Gerrymander With a Novel Approach to Democracy*, NEWSWEEK (Oct. 5, 2017, 3:18 PM), <http://www.newsweek.com/why-supreme-court-case-voting-rights-could-change-america-676871> [https://perma.cc/QTW3-TY98].
 7. The Efficiency Gap Method (EG), created by political scientist Eric M. McGhee and law professor Nicholas O. Stephanopoulos, debuted in a 2015 article, coauthored by the pair, and offered a potentially "manageable" test. See generally Nicholas O. Stephanopoulos & Eric M. McGhee, *Partisan Gerrymandering and the Efficiency Gap*, 82 U. CHI. L. REV. 831 (2015) [hereinafter Stephanopoulos & McGhee, *Partisan Gerrymandering and the Efficiency Gap*]. See also *Vieth*, 541 U.S. at 267, for a discussion of manageability in this context. To help assess whether a districting plan rises to the level of unconstitutional partisan gerrymandering, EG compares each party's number of "wasted votes" (meaning those which do not contribute to victory either because they exceeded what the party needed to win a given district or because they were cast in a district the party did *not* win) so as to capture, in a single figure, the extent to which partisan interests affected "one party's voters more than the other's." Nicholas O. Stephanopoulos & Eric M. McGhee, *The Measure of a Metric: The Debate Over Quantifying Partisan Gerrymandering*, 70 STAN. L. REV. 1503, 1506 (2018) [hereinafter Stephanopoulos & McGhee, *The Measure of a Metric*].

And the intelligent man on the street is going to say that's a bunch of baloney.

...

[Y]ou're going to take these—the whole point is you're taking these issues away from democracy and you're throwing them into the courts pursuant to, and it may be simply my educational background, but I can only describe as sociological gobbledygook.⁸

With this description, Roberts hit the nail squarely on the head: it *was* simply his educational background.⁹ But we surely cannot fault the Chief Justice for his failure to grasp the complexities of a metric which, admittedly, was the product of several doctorates' worth of expertise. Why Roberts viewed EG's academic makeup as a negative, however, is not entirely clear. After all, four years earlier in *Shelby County v. Holder*, Roberts had rested his invalidation of Section 4(b) of the Voting Rights Act (VRA)¹⁰ on social science evidence that was only a fraction less

8. Transcript of Oral Argument at 37–40, *Gill v. Whitford*, 138 S. Ct. 1916 (2018) (No. 16-1161), 2017 WL 4517131.

9. With “only” a Juris Doctorate and Bachelor of Arts degree in history, Roberts’s educational background, standing alone, could not have prepared him to independently evaluate the methodological soundness of a statistical formula like EG. Richard A. Serrano et al., *Roberts Was Ready at Every Turn*, L.A. TIMES (July 25, 2005), <https://www.latimes.com/archives/la-xpm-2005-jul-25-na-roberts25-story.html> [<https://perma.cc/QV2V-87L8>]. For further discussion, see *infra* Part II.

10. See *Shelby County v. Holder*, 570 U.S. 529, 531, 556 (2013). In *Shelby County*, Roberts declared the “coverage formula” of Section 4(b) of the VRA unconstitutional, removing the sole mechanism by which jurisdictions became subject to the “preclearance” measures of Section 5 of the Act, which prevents problem jurisdictions from implementing changes to voting systems without first receiving approval from the U.S. Attorney General. *Id.* at 537. In striking down Section 4(b), Roberts noted that “things have changed dramatically,” and claimed that “today’s statistics tell an entirely different story” than those which justified the Section’s original enactment. *Id.* at 531, 556.

Though statistical calculations seemed outside Roberts’ comfort zone in *Gill*, in *Shelby County*, he had not only relied on such data, but compiled it himself, and perhaps unsurprisingly, “some of the numbers Roberts included in his chart were wrong.” Ryan Gabrielson, *It’s a Fact: Supreme Court Errors Aren’t Hard to Find*, PROPUBLICA (Oct. 17, 2017, 8:00 AM), <http://www.propublica.org/article/supreme-court-errors-are-not-hard-to-find> [<https://perma.cc/2V79-YB4D>] (finding that “Roberts used numbers that counted Hispanics as white, including many Hispanics who weren’t U.S. citizens and could not register to vote, which had the effect of inaccurately lowering the rate for white registration”). For instance, as evidence of the supposed “dramatic[]” change obviating the need for Section 4(b), Roberts stated in *Shelby County* that African American voter registration now outpaced white voter registration in five of the six States originally covered by the preclearance requirements of Section 5, but due to Roberts’ flawed data collection, this point was false. See *Jurisdictions Previously Covered by Section 5*, U.S. DEP’T JUST., <https://www.justice.gov/crt/jurisdictions-previously-covered-section-5> [<https://perma.cc/6HPU-CERR>] (last updated Aug. 6, 2015). If the numbers had been tallied properly, then in Georgia, black voter registration would not have exceeded whites’ by 0.7

complex than EG.¹¹ Why, we might ask, was Roberts comfortable deferring to social science to overturn settled law in *Shelby County*, but so turned off by its use in *Gill*, another voting rights—but not VRA¹²—case?

Hints may lie in Roberts' colorful word choice. "Gobbledygook" is defined as "language marked by abstractions, jargon, euphemisms, and circumlocutions," and is synonymous with "claptrap," "nonsense," "garbage," "poppycrack," "crapola," and another evocative term leveled at EG by Roberts, "baloney."¹³ Ouch. That the word is an inherently funny one,¹⁴ onomatopoeic for the sounds of noisy turkeys,¹⁵ might obscure just how devastating a descriptor it is in *any* context, and much less when leveled by the Chief Justice at your shiny new legal tool. But here, gobbledygook's use seems to connote something even more damning for adventuresome plaintiffs: that the Court is leery of the unique dangers posed by social science expertise. What are the Justices to do, Roberts might well wonder, if an illegitimate or predatory "expert" weasels his way into the courtroom and attempts to blind them with junk science?¹⁶ Might EG be the work of such a charlatan? And if so, why should Roberts be deferential to what he might rightly worry is purposely inscrutable? Is such devil's advocacy a necessary

percent, but instead still trailed it by 4 percent; in Virginia, white voter registration would not have been just 10 percent higher than blacks', but 14.2. Gabrielson, *supra*.

11. See Adam Liptak, *A Case for Math, Not 'Gobbledygook,' in Judging Partisan Voting Maps*, N.Y. TIMES (Jan. 15, 2018), <http://www.nytimes.com/2018/01/15/us/politics/gerrymandering-math.html> [https://perma.cc/P2V4-7WSZ] (noting that "[t]he statistics in the recent gerrymandering cases were more complicated [than those in *Shelby County*], but not by much").
12. See *infra* Subpart II.C.2 for further discussion regarding this distinction.
13. *Gobbledygook*, MERRIAM-WEBSTER, <http://www.m-w.com/thesaurus/gobbledygook> [https://perma.cc/5PWE-UGJN] (last visited Oct. 30, 2018) (providing the definition); *Gobbledegook*, OXFORD LIVING DICTIONARIES, ENGLISH, <http://en.oxforddictionaries.com/thesaurus/gobbledegook> (last visited Oct. 30, 2018) [https://perma.cc/VCE9-MK34] (listing synonyms for the alternate spelling "gobbledegook").
14. See *How Funny Is This Word? The 'Snunkoople' Effect*, SCIENCE DAILY (Nov. 30, 2015), <http://www.sciencedaily.com/releases/2015/11/151130131847.htm> [https://perma.cc/3PEA-ZEBZ] (discussing "inherently funny" words and efforts to mathematically quantify how innately surprising, and therefore funny, words are by examining their entropy, "a mathematical measure of how ordered or predictable" they are).
15. See Maury Maverick, *The Case Against 'Gobbledygook'; Maury Maverick Assails Vague, Pompous, Repetitious English and the Two-Gun Word Bandits Who Use It*, N.Y. TIMES, May 21, 1944, at 11, 35–36. Congressman Maverick (1895–1954), a U.S. House Representative from Texas, debuted the word in a memorandum in which he cautioned colleagues against using overly verbose, bureaucratic language. *Id.* In this New York Times piece, Maverick wrote of the word: "People asked me how I got the word. I do not know. It must have come in a vision. Perhaps I was thinking of the old bearded turkey gobbler back in Texas who was always gobbledygobbling and strutting with ridiculous pomposity. At the end of his gobble there was a sort of gook." *Id.* at 11.
16. See generally Michael Rustad & Thomas Koenig, *The Supreme Court and Junk Social Science: Selective Distortion in Amicus Briefs*, 72 N.C. L. REV. 91 (1993).

safeguard then to prevent the potential problems of imperfect expertise from being magnified and compounded by the Highest Court's canonization?

Though this hesitancy is unfortunate for the *Gill* plaintiffs, it is by no means new. The legal literature through the decades is rife with skepticism toward experts, with an oft-quoted aphorism even stating that liars “are of three classes: liars, damned liars, and expert witnesses.”¹⁷ These beyond-damned liars, the story goes, are “partisans pure and simple,”¹⁸ “hired guns,”¹⁹ or, less politely, “whores.”²⁰ Because they “often are driven in their work by a need to answer a particular question related to the issues of a particular case, they sometimes face pressure to sacrifice appropriate methodology for the sake of expediency.”²¹ And moreover, given an inevitable lack of adequate—as Roberts termed it—“educational background” in the expert's field of study, judges and jurors are inevitably unable to intelligently evaluate the substance of evidence presented.²² “The trouble with all this,” Learned Hand wrote in his classic 1901 essay about expert witnesses, “is that it is setting the jury to decide, where doctors disagree.”²³

In *Gill*, the “doctors” did not even disagree.²⁴ Nonetheless, the Court remained unimpressed, passing on EG and punting the case back to the District

17. Paul Meier, *Damned Liars and Expert Witnesses*, 81 J. AM. STAT. ASS'N 269, 269 (1986) (noting that “[t]he origins of this aphorism are uncertain, but it appears in various forms in legal writing during the past century”). For another, amusingly-worded view, see *Earl M. Kerstetter, Inc. v. Commonwealth*, 171 A.2d 163, 165 (Pa. 1961) (“Expert opinion . . . is only an ordinary guess in evening clothes . . .”).
18. Roscoe Pound, *The Causes of Popular Dissatisfaction With the Administration of Justice*, 14 AM. LAW. 445, 448 (1906) (“It leads to exertion to ‘get error into the record’ rather than to dispose of the controversy finally and upon its merits. It turns witnesses, and especially expert witnesses, into partisans pure and simple.”).
19. See, e.g., Jennifer L. Mnookin, *Expert Evidence, Partisanship, and Epistemic Competence*, 73 BROOK. L. REV. 1009, 1026 (2008); Rachel F. Moran, *What Counts as Knowledge? A Reflection on Race, Social Science, and the Law*, 44 LAW & SOC'Y REV. 515, 531 (2010).
20. See Samuel R. Gross, *Expert Evidence*, 1991 WIS. L. REV. 1113, 1115 (“Experts in other fields see lawyers as unprincipled manipulators of their disciplines, and lawyers and experts alike see expert witnesses—those members of other learned professions who will consort with lawyers—as whores.”).
21. *Melendez-Diaz v. Massachusetts*, 557 U.S. 305, 318 (2009) (alteration in original).
22. Transcript of Oral Argument, *supra* note 8, at 40; Mnookin, *supra* note 19, at 1026; see also *infra* Subpart II.A for a detailed discussion of this issue.
23. Learned Hand, *Historical and Practical Considerations Regarding Expert Testimony*, 15 HARV. L. REV. 40, 54 (1901). For a more contemporary view, replete with characteristic Judge Posner pith, see *Rosen v. Ciba-Geigy Corp.*, 78 F.3d 316, 319 (7th Cir. 1996) (“[T]he courtroom is not the place for scientific guesswork, even of the inspired sort. Law lags science; it does not lead it.”).
24. That is, they did not disagree that gerrymandering had occurred, even if they would have tweaked or revised EG as the new standard. See generally Benjamin Plener Cover, *Quantifying Partisan Gerrymandering: An Evaluation of the Efficiency Gap Proposal*, 70 STAN. L. REV. 1131, 1131 (2018) (noting that though “the efficiency gap is a useful indicative measure of partisan

Court for the plaintiffs' lack of standing.²⁵ Partisan gerrymandering would instead remain a judicially unsolvable problem despite a swell of scholars, in a consensus bordering on unanimity, crying out that they had already known the solution for decades.²⁶ To crib Charles L. Black, Jr.'s influential essay discussing the legality of the segregation decisions, "it seems that . . . while no actual doubt exists as to what [*partisan gerrymandering*] is for and what kind of societal pattern it supports and implements, there is no ritually sanctioned way in which the Court, as a Court, can permissibly learn what is obvious to everybody else and to the Justices as individuals."²⁷

But why not now sanction a new ritual? Was the sociology here really, actually, such "gobbledygook"? We might have our doubts, as there exists a long record of Justices not only resisting "these gauzy sociological considerations" broadly, but of doing so in the voting rights field specifically,²⁸ and each time the excuse recurs, it grows a bit less persuasive. How is it that time and again such impeccably-credentialed social scientists, utilizing widely accepted methodologies

gerrymandering under the circumstances of cases like the one currently before the Court," there exist adjustments which would allow similar formulas to be more descriptively valuable in other circumstances). For their part, the Wisconsin State appellants also took it as given that political gerrymandering took place. They pointed to the fact that the practice was well-rooted in American history, offering examples dating back as far as 1789, when "the Federalist Party unnecessarily divided a county by lopping off towns into different districts to secure to their party the control of the new government provided for in the federal Constitution." Brief for Appellants at 6, *Gill*, 138 S. Ct. 1916 (No. 16-1161) (internal quotation marks omitted).

25. *Gill*, 138 S. Ct. at 1933, 1934; see also *infra* Subpart II.C.2 for further discussion regarding Plaintiffs' standing.
26. See Stephanopoulos & McGhee, *The Measure of a Metric*, *supra* note 7, at 1505 (noting that, for "several decades," there has been a virtual consensus among scholars that statistical formulas like EG can detect whether a partisan gerrymander has occurred); see also Transcript of Oral Argument, *supra* note 8, at 16 (Justice Sotomayor noting that "every single social science metric points in the same direction.").
27. Charles L. Black, Jr., *The Lawfulness of the Segregation Decisions*, 69 *YALE L.J.* 421, 426, 427 (1960) (swapping partisan gerrymandering for segregation).
28. *City of Mobile v. Bolden*, 446 U.S. 55, 66–67, 75 n.22 (1980) (plurality opinion) (holding that plaintiffs must demonstrate not just disparate impact, but also discriminatory intent when bringing claims under Section 2 of the VRA). In dissent, Justice Marshall argued impact alone should be sufficient, and that the plurality's approach turned a blind eye to an ugly and still recent past: "Unconstitutional vote dilution occurs . . . when a discrete political minority whose voting strength is diminished by a districting scheme proves that *historical and social factors* render it largely incapable of effectively utilizing alternative avenues of influencing public policy." *Id.* at 111 n.7 (Marshall, J., dissenting) (emphasis added). Writing for the plurality, Justice Stewart took aim at Marshall's dissent, stating "these gauzy sociological considerations have no constitutional basis." *Id.* at 75 n.22 (plurality opinion).

Partly in response to outcries over *Bolden*, Congress strengthened Section 2 in a 1982 amendment specifically incorporating many of Marshall's "gauzy considerations," and effectively overturning the decision. See SAMUEL ISSACHAROFF ET AL., *THE LAW OF DEMOCRACY: LEGAL STRUCTURE OF THE POLITICAL PROCESS* 641–42 (5th ed. 2016).

to offer the same sort of competent, informed analysis which more than passes muster in peer-reviewed scholarly journals, can so frequently have their conclusions dismissed by the judiciary? And what impact does this have on the development of the law?

In this Comment, I argue that the judiciary has an improper institutional aversion to what it misguidedly terms “sociological gobbledygook.” This tack shuts out scholars like EG creators McGhee and Stephanopoulos rather than embracing their specialized knowledge to fashion creative new solutions to long-vexing problems. Consequently, and in large part due to widespread cultural and cognitive forces—as well as procedural biases toward expert witnesses generally and social scientists specifically—courts routinely wave away soft science testimony worth considering, a practice that systematically deprives plaintiffs of a whole category of academic disciplines by which they may prove especially subtle, but no less pernicious, harms under the law. This Comment explores why this is the case.

Throughout, I primarily situate my analysis within the doctrines of (1) voting rights litigation and (2) Title VII systemic disparate treatment liability in the class action setting, two areas in which plaintiffs rely crucially, if not entirely, on social science research for building up causes of action. The separateness, and even dissimilarity, of these doctrines ideally will suggest that the issue recurs across a range of substantive areas I lack the space to fully detail. Indeed, as I demonstrate, it may be quite natural—even unavoidable—for the problem to arise in any context in which (a) social science is a necessary component of proof and (b) there is no express political will guiding or authorizing the judiciary to depart from its traditional approaches. Other fields that I discuss, though in lesser depth, include litigation surrounding the death penalty and marriage equality.

In Part I, I introduce Title VII class action case *Wal-Mart Stores, Inc. v. Dukes*²⁹ as a paradigmatic decision, in which a large number of the factors that have contributed to the modern Court’s treatment of social science expertise appear. Thereafter, I frequently return to *Wal-Mart* to further illustrate various features of social science testimony that play into the phenomenon, and to show how or why these features may be common in a variety of legal frames.

In Part II, I detail the challenges faced by plaintiffs bringing claims which require establishing—through social science evidence—propositions “unfamiliar” to the judiciary in either content or form. Given the system’s unique demands, judges must frequently make snap determinations about nuanced or difficult subject matter outside their epistemic competence. As a result, these judges must turn to

29. 564 U.S. 338 (2011).

secondary proxies for reliability rather than assessing true substance, a process I describe in Subpart II.A. In Subpart B, I link this notion of secondary knowledge in judicial decisions with perhaps the most powerful proxy: an individual judge or Justice's own personal beliefs, or "culture." Because social science is deeply enmeshed in, and arguably inseparable from, culture itself, research challenging—or in some way failing to conform with—an individual judge's preexisting cultural agenda regarding highly-charged topics like race, gender, sexuality, religion, politics, or others, will disproportionately face heightened skepticism and exclusion. In this Subpart, I contrast *Brown v. Board of Education*³⁰ with *Exxon Shipping Co. v. Baker*³¹ and *McCleskey v. Kemp*,³² two cases in which the Court's ultimate holding seems to have been foreordained by the majority Justices' preexisting cultural agendas.

In Subpart C, I move beyond substance and into form, to further describe how the exercise of utilizing proxy knowledge impacts social scientists from a procedural standpoint. Because proxy knowledge necessarily rewards the familiar, the unchallenging, and the traditional, and since social science experts frequently fail to conform to the judiciary's comfortable archetype of good experts and expertise, such experts more frequently face undeserved exclusion or bias than do other types. Features inherent to social science—such as its focus on aggregate behavior or harms suffered en masse, as opposed to discrete, individual moments of action or injury—frustrate judicial expectations for what proof is quote-unquote supposed to look like. As I show, these normative expectations are even embedded in doctrines like Article III standing and in the Federal Rules of Civil Procedure. Because some causes of action rely crucially on social science evidence, this creates a regime in which certain rights lack any judicially-endorsed methods of proof.

In Part III, I consider problems arising from attempts to "translate" between the frequently incompatible languages of science and law more broadly. In the process, I describe a variety of ways in which the methodological norms of soft science predictably lead to the Court privileging "hard" science over it. Moreover, I consider how a formalist method of legal interpretation not only coexists more peaceably with hard science proof, but in fact creates considerable tension with typical soft science approaches. This tension, I argue, is tolerated and even encouraged by society at large.

30. 347 U.S. 483 (1954).

31. 554 U.S. 471 (2008).

32. 481 U.S. 279 (1987).

In conclusion, I argue that a potential solution is for courts to more routinely embrace the legal system's most fundamental tool: the trial itself. I discuss the landmark twelve-day trial for the Proposition 8 case *Hollingsworth v. Perry*,³³ in which plaintiffs successfully challenged the constitutionality of the 2008 California ballot initiative that amended the state constitution to define marriage as being only between a man and a woman. The trial in *Perry* was unique because it turned heavily on disputed so-called "legislative" facts (alternatively known as "social"—as opposed to "adjudicatory"—facts), such as "whether marriage by definition involved the ability to produce children."³⁴ These sorts of questions are regularly assumed unanswerable by courts, and their content might consist entirely of social science evidence. Nonetheless, the *Perry* trial showed that the adversarial process and rigorous cross-examination presents "an excellent opportunity to replace ignorance with knowledge."³⁵ By embracing and more frequently utilizing the trial, courts may construct more certain answers to the complex questions that social science evidence so frequently addresses.

I. "WELCOME TO WAL-MART": AN INTRO TO *DUKES*

Nearly everything about *Wal-Mart Stores, Inc. v. Dukes* was oversized. The case pitted a retail titan, the nation's largest private employer and *world's* biggest corporation,³⁶ against the largest putative class in history: a group of over 1.5 million current and former female Wal-Mart employees, requesting billions in monetary relief for gender discrimination in pay and hiring decisions under Title VII of the Civil Rights Act of 1964.³⁷ Representing all women who had worked in any Wal-Mart store since December 1998, lead plaintiff Betty Dukes requested class certification under Rule 23(b)(2) of the Federal Rules of Civil Procedure,³⁸ a

33. 570 U.S. 693 (2013). The case originated in the District Court, styled as *Perry v. Schwarzenegger*, 704 F. Supp. 2d 921 (N.D. Cal. 2010); when it reached the Supreme Court, it was recaptioned as *Hollingsworth v. Perry*.

34. KENJI YOSHINO, *SPEAK NOW: MARRIAGE EQUALITY ON TRIAL: THE STORY OF HOLLINGSWORTH V. PERRY* 75–76 (2015).

35. *Id.* at 79.

36. *Fortune* 500, FORTUNE, <http://www.fortune.com/fortune500/list> [https://perma.cc/XQN8-YFD9] (last visited Apr. 14, 2020).

37. 42 U.S.C. § 2000e-2 (2018). Title VII forbids discrimination by employers based on characteristics such as sex, race, religion, or ethnicity. *Id.* § 2000e-2(a)(1). This protection is flexible, extending to various forms of discrimination impacting any of the "privileges" or "conditions" of employment, including promotions, evaluations, pay raises, and awards of benefits. *Id.* § 2000e-2(h).

38. FED. R. CIV. P. 23. To establish a putative class's entitlement to certification under Rule 23(b), plaintiffs must first satisfy four prerequisites laid out in Rule 23(a):

(1) the class be so numerous that joinder of all members is impracticable;

move the District Court granted in 2004³⁹ and a divided en banc Ninth Circuit upheld in 2010.⁴⁰

By the time the Supreme Court granted certiorari in late 2010, reactions had become nearly as enlarged as the would-be plaintiff class itself, with many fearing, and countless others heralding, the generational transformation sure to accompany a proplaintiff verdict.⁴¹ Given the hype, one understandably might have forgotten that the Court had agreed to review the propriety of mere class certification—not the actual merits of the case—and since Rule 23 is intended to

-
- (2) there are questions of law or fact common to the class;
 - (3) the claims or defenses of the representative parties are typical of the claims or defenses of the class; and
 - (4) the representative parties will fairly and adequately protect the interests of the class.

These prerequisites are commonly referred to as the “numerosity,” “commonality,” “typicality,” and “adequacy” prongs, respectively.

Id.

39. *Dukes v. Wal-Mart Stores, Inc.*, 222 F.R.D. 137 (N.D. Cal. 2004).

40. *Dukes v. Wal-Mart Stores, Inc.*, 603 F.3d 571 (9th Cir. 2010) (en banc).

41. Compare Melissa Hart, *Learning From Wal-Mart*, 10 EMP. RTS. & EMP. POL’Y J. 355, 356 (2006) (“[W]hat is really significant about the suit is less its remarkable scope and more its implications for the development of discrimination law. As part of a growing trend of gender discrimination class claims, [the case] has the potential to push the boundaries of the law to confront the pervasive, tenacious stereotypes that continue to limit women’s workplace opportunities.”), with Brief of the Chamber of Commerce of the United States of America as *Amicus Curiae* in Support of Petitioner at 5, *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338 (2011) (No. 10-277), 2011 WL 288900, at *5 (arguing that upholding the decision “would erase decades of class-action precedents protecting defendants from unfair trials—and would bury American businesses in abusive class-action lawsuits to the detriment of consumers, the U.S. economy and the judicial system itself”).

The case might also be seen as part of a broader trend of souring public opinion toward the store. In the early- to mid-2000s—about the time when the District Court first certified the *Dukes* class—the notion that Wal-Mart, despite its colossal profits and unceasing expansion, was actually “bad” for American life had become all but fully mainstream. Michelle Gilman, *How the Supreme Court Made Economic Inequality a Whole Lot Worse*, CONVERSATION (Mar. 20, 2016), <https://theconversation.com/how-the-supreme-court-made-economic-inequality-a-whole-lot-worse-56416> [<https://perma.cc/36TM-QN5Q>] (reflecting on Occupy Wall Street and the *Wal-Mart* decision, among others). During this period, Wal-Mart “was battling one image problem after another,” and such prolonged negative PR began significantly impacting the store’s bottom line. See, e.g., Anne D’Innocenzio, *Wal-Mart’s Image Tarnished Once Again*, ARIZ. DAILY STAR (May 13, 2012), http://www.tucson.com/business/national-and-international/walmart-s-image-tarnished-once-again/article_1c122787-1be1-50c8-8f59-446eacd5da2e.html [<https://perma.cc/DLN9-257D>] (noting that between 2005 and 2007, “Wal-Mart acknowledged that the bad publicity [including from *Dukes*] was beginning to hurt its stock,” as its shares fell 20 percent “to an eight-year low of \$. . .”). Considered together, this backdrop might have infused the case with a bit of *David v. Goliath* baggage drawing on some of the same ire motivating, among other examples, the Occupy Wall Street protests, which began less than three months after the Court’s *Wal-Mart* decision.

certify classes “fated to lose as well as classes that are sure to win,”⁴² the most important battles in *Wal-Mart* were, presumably, to remain unfought for still some time.

The plaintiffs’ argument was simple: Wal-Mart’s storewide policy⁴³ of affording its individual regional managers virtually unchecked discretion⁴⁴ over hiring and pay raise decisions, combined with an unusually strong corporate culture, made those choices susceptible to gender bias and stereotyping.⁴⁵ This resulted in both higher pay for men than women—even between employees with identical or similar job titles—and longer wait times for, as well as a lower frequency of, promotions to management for female employees.

Conversely, Wal-Mart’s defense highlighted the uncertainty which would necessarily accompany such a historically large and widely dispersed class. As they put it in their brief for the Court:

The class members—potentially millions of women supervised by tens of thousands of different managers and employed in thousands of different stores throughout the country—assert highly individualized, fact-intensive claims for monetary relief.... The named plaintiffs’ claims cannot conceivably be typical of the claims of the strangers they seek to represent. These intractable problems are compounded by a virtually boundless class definition that produces an across-the-board

42. *Schleicher v. Wendt*, 618 F.3d 679, 686 (7th Cir. 2010).

43. This policy, Plaintiffs argued, amounted to a corporate “pattern or practice” of discrimination. *Wal-Mart*, 564 U.S. at 352. Generally, courts recognize two means by which plaintiffs may show a pattern or practice of Title VII discrimination, “disparate treatment” and “disparate impact.” See *Int’l Brotherhood of Teamsters v. United States*, 431 U.S. 324, 335–36 n.15 (1977). In *Teamsters*, Justice Stewart described disparate treatment as “the most easily understood type of discrimination. The employer simply treats some people less favorably than others because of their race, color, religion, sex, or national origin.” *Id.* Disparate impact, by contrast, involves “employment practices that are facially neutral in their treatment of different groups but that in fact fall more harshly on one group than another and cannot be justified by business necessity.” *Id.* at 336 n.15.

44. In *Watson*, the Court held that “subjective or discretionary employment practices” may serve as the basis for Title VII liability. *Watson v. Fort Worth Bank & Tr.*, 487 U.S. 977, 991 (1988). In *Watson*, the Court found that Fort Worth’s affording its managers “complete, unguided discretion in evaluating applicants for [promotions]” had a racially disparate impact on black applicants. *Id.* at 1009 (Blackmun, J., concurring). Fort Worth had in place no “precise and formal criteria for evaluating candidates for the positions for which Watson unsuccessfully applied. [Instead, the bank] relied . . . on the subjective judgment of supervisors . . .” *Id.* at 982.

45. For more in-depth discussion of the *Dukes* plaintiffs’ arguments, see discussion *infra* Part I. The *Watson* Court noted “the problem of subconscious stereotypes and prejudices” and held that an employer’s “undisciplined system of subjective decisionmaking” can result in “precisely the same effects as a system pervaded by impermissible intentional discrimination.” *Watson*, 487 U.S. at 990–91.

class⁴⁶ pervaded by conflicts among its members. This kaleidoscope⁴⁷ of claims, defenses, issues, locales, events, and individuals makes it impossible for the named plaintiffs to be adequate representatives of the absent class members.⁴⁸

The argument is no doubt an intuitively appealing one. Given what seems, in our abstract imagination of “employment discrimination,” its uniquely personal and local nature, establishing its occurrence on a *nationwide* scale, with any surety, could appear impossible absent perfect information. If this is true, then predicating a multibillion-dollar penalty on anything less than one hundred percent confidence would be unjust. Justice Scalia, writing for the majority, agreed:

Respondents have not identified a common mode of exercising discretion that pervades the entire company—aside from their reliance on [their expert witness, sociologist] Dr. Bielby’s social-framework analysis that we have rejected.⁴⁹ In a company of Wal-Mart’s size and geographical scope, it is quite unbelievable that all managers would exercise their discretion in a common way without some common

46. This language signals to *Falcon* in which the Court disapproved of permissive certification of Title VII class actions based merely on proof of “across-the-board” discriminatory policies. *Gen. Tel. Co. of the Sw. v. Falcon*, 457 U.S. 147, 161 (1982). “[T]he allegation that such discrimination has occurred neither determines whether a class action may be maintained in accordance with Rule 23 nor defines the class that may be certified.” *Id.* at 157. Instead, “a Title VII class action, like any other class action, may only be certified if the trial court is satisfied, after a rigorous analysis, that the prerequisites of Rule 23(a) have been satisfied.” *Id.* at 161.

47. “Kaleidoscope” seems a nod to Chief Judge Alex Kozinski’s short dissent to the Ninth Circuit’s 2010 en banc affirmance, which read:

Maybe there’d be no difference between 500 employees and 500,000 employees if they all had similar jobs, worked at the same half-billion square foot store and were supervised by the same managers. But the half-million members of the majority’s approved class held a multitude of jobs, at different levels of Wal-Mart’s hierarchy, for variable lengths of time, in 3,400 stores, sprinkled across 50 states, with a kaleidoscope of supervisors (male and female), subject to a variety of regional policies that all differed depending on each class member’s job, location and period of employment. Some thrived while others did poorly. They have little in common but their sex and this lawsuit.

Dukes v. Wal-Mart Stores, Inc., 603 F.3d 571, 652 (9th Cir. 2010) (Kozinski, J., dissenting).

48. Brief for Petitioner at 1, *Wal-Mart*, 564 U.S. 338 (No. 10-277), 2011 WL 201045.

49. Generally, “social framework analysis” involves identifying distinctive organizational features, policies, and practices in order to evaluate “how extant social science theory and research applies to the specific circumstances of the . . . setting where discrimination is alleged to have occurred.” William T. Bielby, *Can I Get a Witness? Challenges of Using Expert Testimony on Cognitive Bias in Employment Discrimination Litigation*, 7 EMP. RTS. & EMP. POL’Y J. 377, 390 (2003) (alteration omitted). The expert then “analyzes the specific features of the organization’s policies and practices and evaluates them against what that social science scholarship has shown to be factors that create and sustain bias and those that minimize bias.” *Id.*

direction. Respondents attempt to make that showing by means of statistical and anecdotal evidence, but their evidence falls well short.⁵⁰

At all stages, “commonality” was by far the most important and bitterly contested of the four Rule 23(a) prerequisites.⁵¹ Justice Scalia called commonality “the crux of th[e] case,” as it determines “whether under the particular circumstances maintenance of a class action is economical and whether the named plaintiff’s claim and the class claims are so interrelated that the interests of the class members will be fairly and adequately protected in their absence.”⁵² For Scalia, commonality required plaintiffs to demonstrate that all class members “suffered the same injury,” but even this, by itself, was insufficient: “What matters to class certification . . . is not the raising of common ‘questions’—even in droves—but, rather the capacity of a class-wide proceeding to generate common answers apt to drive the resolution of the litigation.”⁵³

Though never attempted on such a grand scale, the *Dukes* plaintiffs were hardly auditioning some offbeat new theory of Title VII class liability in pointing to statistical and anecdotal evidence. In *Teamsters v. U.S.*,⁵⁴ the Court first held that statistical evidence showing pervasive disparities between minority and white employees, supported by testimony describing “specific instances” of discrimination, was capable of proving employment discrimination.⁵⁵ Looking to

50. *Wal-Mart*, 564 U.S. at 356. As Wal-Mart did in their Brief for Petitioner, *supra* note 48, Scalia endorsed Kozinski’s Ninth Circuit dissent and quoted large portions of it in his opinion. See *id.* at 359–60.

51. FED. R. CIV. P. 23 contains these prerequisites. In the District Court, Wal-Mart did not contest that the plaintiffs satisfied numerosity (1), and Judge Martin J. Jenkins quickly dismissed arguments that certification should fail on either typicality (3) or adequacy (4). *Dukes v. Wal-Mart Stores, Inc.*, 222 F.R.D. 137, 144, 166–69 (N.D. Cal. 2004). This left the parties to focus the most attention on commonality (2), to which Judge Jenkins devoted 22 of the 25 pages in his opinion discussing Rule 23(a). See *id.* at 144–66.

52. *Wal-Mart*, 564 U.S. at 349 n.5 (internal quotation marks omitted) (quoting *Gen. Tel. Co. Sw. v. Falcon*, 457 U.S. 147, 157–58 (1982)).

53. *Id.* at 350 (emphasis omitted) (quoting Richard A. Nagareda, *Class Certification in the Age of Aggregate Proof*, 84 N.Y.U. L. REV. 97, 132 (2009)). Compare this definition with the one proposed by Judge Jenkins: “Rule 23(a)(2) does not require that all questions of law or fact be common. The test is qualitative rather than quantitative—one significant issue common to the class may be sufficient to warrant certification. Indeed, the necessary showing to satisfy commonality is ‘minimal.’” *Dukes*, 222 F.R.D. at 145 (citation omitted). “[P]laintiffs may demonstrate commonality by showing that class members have shared legal issues but divergent facts or that they share a common core of facts but base their claims for relief on different legal theories.” *Id.*

54. *Int’l Brotherhood of Teamsters v. United States (Teamsters)*, 431 U.S. 324 (1977).

55. *Id.* at 338–39 (writing that “[w]e have repeatedly approved the use of statistical proof . . . to establish a prima facie case of racial discrimination in jury selection cases Statistics are equally competent in proving employment discrimination,” and noting that anecdotal accounts can help bring “the cold numbers convincingly to life”). In an amicus opposing Title VII liability based on the statistical proof offered in *Teamsters*, the Equal Employment

join this tradition, the *Dukes* plaintiffs offered proof of a similar sort. To go with an array of anecdotal accounts,⁵⁶ the plaintiffs' expert statistician Dr. Richard Drogin analyzed Wal-Mart regional data and found "statistically significant disparities between men and women . . . in terms of compensation and promotions, that these disparities are wide-spread across regions, and that they can be explained only by gender discrimination."⁵⁷

But given the size and scope of the claim,⁵⁸ something more—and more creative—would be required for a skeptical judge or Justice to connect the transcontinental dots. The plaintiffs, after all, were alleging not one specific affirmative policy which consistently and directly resulted in discrimination by individual store managers. Rather, they argued that the "policy" of establishing *no* policy, combined with Wal-Mart's particularly strong corporate culture, led managers to draw on gendered and sex discriminatory stereotypes to carry out that "no policy" policy. Demonstrating this at the class certification stage,⁵⁹ the theory

Advisory Council warned that accepting such proof would precipitate claims not unlike the *Dukes* plaintiffs', writing that "[t]he issue of the proper role of statistics in a Title VII class action is now before this Court for the first time No issue is more crucial to the rational and effective enforcement of the equal employment opportunity enactments." Brief Amicus Curiae of the Equal Employment Advisory Council at 3, *Teamsters*, 431 U.S. 324 (No. 75-672), 1976 WL 194254, at *3. Nonetheless, the principle that statistical proof plus anecdotal evidence may prove Title VII discrimination remains an influential canon forty-plus years post-*Teamsters*.

56. See *Dukes v. Wal-Mart Stores, Inc.*, 603 F.3d 571, 610 n.34 (9th Cir. 2010) (noting that "[p]laintiffs submitted declarations from each of the class representatives, as well as 114 declarations from putative class members around the country").

57. *Id.* at 604.

58. In an earlier action making up one of several consolidated in *Teamsters*, defendant T.I.M.E.-D.C., a "major interstate common carrier" operated "51 terminals in 26 states . . . and employ[ed] 6,472 people." *United States v. T.I.M.E.-D.C. Inc.*, 517 F.2d 299, 304 (5th Cir. 1975). By contrast, in *Wal-Mart*, the store employed over 230 times as many employees, in over sixty-nine times as many stores, in all fifty states. Dan Ackerman, *Wal-Mart and Sex Discrimination By the Numbers*, *FORBES* (June 22, 2004), https://www.forbes.com/2004/06/23/cx_da_0623topnews.html [<https://perma.cc/8QSX-ZSYU>].

59. Though at the certification stage it might seem premature to take an overly-detailed look at the claim's Title VII merits, the Court in *Falcon* held that such considerations often overlap at this stage: "[S]ometimes it may be necessary for the court to probe behind the pleadings before coming to rest on the certification question." *Falcon*, 457 U.S. at 160. Courts must undertake a "rigorous analysis[] that the prerequisites of Rule 23(a) have been satisfied," and in this process, "the class determination generally involves considerations that are enmeshed in the factual and legal issues comprising the plaintiff's cause of action." *Id.* at 160–61 (internal quotations omitted).

It must be noted, however, that during the precertification stage plaintiffs have a more limited right to discovery than they do postcertification. See FED. R. CIV. P. 23 advisory committee's note to 2003 amendment (stating that "[a]lthough an evaluation of the probable outcome on the merits is not properly part of the certification decision . . . it is appropriate to conduct controlled discovery into the 'merits,' limited to those aspects relevant to making the certification decision on an informed basis"). Given this, an overly extensive foray into the

went, would show that all 1.5 million women similarly satisfied the commonality “crux.” To show this—and, ultimately, to educate the judge and later Justices—the plaintiffs tapped social scientist Dr. William T. Bielby, who examined “statistical and qualitative data” distinct to Wal-Mart at the organizational level, to identify procedural and practical features that established social science research reveals can create or sustain, bias.⁶⁰ In his report, Bielby wrote that he found that Wal-Mart “has and promotes a strong corporate culture—a culture that may include gender stereotyping,” which is “especially likely to influence personnel decisions . . . based on subjective factors, because substantial decisionmaker discretion tends to allow people to seek out and retain stereotyp[e]-confirming information and ignore or minimize information that defies stereotypes.”⁶¹ And because Wal-Mart had the “no-policy” policy of giving its regional managers total discretion over hiring decisions, these managers were more likely to assume that men, for example, as the traditional breadwinners, would be more willing to relocate for, and thus accept, a promotion at a store in a new city.⁶²

Though Scalia was confident in rejecting Bielby’s science—which, it should be noted, commanded the support of the American Sociological Association⁶³—two years later, in *Association for Molecular Pathology v. Myriad Genetics*,⁶⁴ it seemed he had a change of heart. Joining the Court’s judgment—that naturally occurring gene sequences were not patentable by a private company—but withholding his full stamp of approval, Scalia wrote in a short concurrence that:

I join the judgment of the Court, and all of its opinion except Part I–A and some portions of the rest of the opinion going into fine details of molecular biology. I am unable to affirm those details on my own knowledge or even my own belief. It suffices for me to affirm, having

merits—and a reliance on the incomplete record obtained through this “controlled discovery”—may well misrepresent what the claim’s ultimate substance would be postcertification. *Id.*

60. See Brief of Amici Curiae American Sociological Association & the Law and Society Association in Support of Respondents at 10, *Wal-Mart*, 564 U.S. 338 (No. 10-277), 2011 WL 757409, at *9 [hereinafter “ASA Amici”]; Declaration of William T. Bielby, Ph.D. in Support of Plaintiffs’ Motion for Class Certification at 4–5, *Dukes v. Wal-Mart Stores, Inc.*, 222 F.R.D. 137 (N.D. Cal. 2004) (No. 01-2252), 2003 WL 24571701 [hereinafter “Bielby Declaration”]; Exhibit C Rebuttal Report of William T. Bielby, Ph.D., *Dukes*, 222 F.R.D. 137 (No. 01-2252), 2003 WL 24571700; Bielby, *supra* note 49, at 390.

61. *Wal-Mart*, 509 F.3d at 1178.

62. See *Wal-Mart*, 564 U.S. at 370 (Ginsburg, J., dissenting) (“Absent instruction otherwise, there is a risk that managers will act on the familiar assumption that women, because of their services to husband and children, are less mobile than men.”).

63. See generally ASA Amici, *supra* note 60.

64. 569 U.S. 576 (2013).

studied the opinions below and the expert briefs presented here, that the portion of DNA isolated from its natural state sought to be patented is identical to that portion of the DNA in its natural state; and that complementary DNA (cDNA) is a synthetic creation not normally present in nature.⁶⁵

While Scalia critics like Erwin Chemerinsky have accused the late Justice of hiding an alleged conservative activism behind “the cloak of judicial modesty,”⁶⁶ such a tendency hardly explains Scalia sitting sideline in *Myriad*, a 9–0 decision with no clear political winners or losers. Rather, the concurrence stands as a troubling backdrop for Scalia’s treatment of Bielby in *Wal-Mart*, a case with significant and obvious partisan baggage. Since the early 1970s, of course, women’s rights have emerged as a politically polarized issue,⁶⁷ and many of his political opponents found Justice Scalia’s views on the subject in particular to be especially out-of-step.⁶⁸ If we accept the premise, then, that the motive and opportunity was there for Scalia in *Wal-Mart*, a sex-discrimination case, why did he abandon the cloak of “modesty” toward science in *Myriad* for an almost forceful rejection of it in *Wal-Mart*?

65. *Id.* at 596 (Scalia, J., concurring).

66. Erwin Chemerinsky, *Scalia Introduced Conservative Activism to Supreme Court*, ORANGE CTY. REG. (Feb. 21, 2016), <http://www.ocregister.com/2016/02/21/scalia-introduced-conservative-activism-to-supreme-court> [<https://perma.cc/UW4W-GP8M>] (also writing that “Justice Antonin Scalia’s greatest impact likely is his having brought conservative judicial activism to the Supreme Court”).

67. See Christina Wolbrecht, *Explaining Women’s Rights Realignment: Convention Delegates, 1972–1992*, 24 POL. BEHAV. 237 (2002) (“At its 1980 convention, the Republican Party removed the [ERA] from its platform, reversing a pattern of nearly 40 years of official party support. Democrats, on the other hand, not only retained a pro-ERA plank but pledged to provide financial support only to candidates who backed the amendment.”); Val Burris, *Who Opposed the ERA? An Analysis of the Social Bases of Antifeminism*, 64 SOC. SCI. Q. 305, 307 n.2, 311 (1983) (noting a study finding “political conservatism” was a strong predictor of opposition to the ERA, and which for men specifically was “significantly correlated with opposition to the ERA”).

68. See, e.g., Adam Cohen, *Justice Scalia Mouths Off on Sex Discrimination*, TIME (Sept. 22, 2010), <http://content.time.com/time/nation/article/0,8599,2020667,00.html> [<https://perma.cc/B3EY-XRUL>] (discussing Justice Scalia’s reputation for retrogressive women’s rights jurisprudence); Calvin Massey, *The Originalist*, CAL. LAW., Jan. 2011, at 33. In this magazine interview, Justice Scalia contradicts the Court’s holding in *Reed v. Reed*, 404 U.S. 71 (1971), that women have a constitutionally-protected right against sex discrimination, saying:

The only issue is whether [the Constitution] prohibits [discrimination on the basis of sex]. It doesn’t. Nobody ever thought that that’s what it meant. Nobody ever voted for that. If the current society wants to outlaw discrimination by sex, hey we have things called legislatures, and they enact things called laws. You don’t need a constitution to keep things up-to-date. All you need is a legislature and a ballot box.

Id.

An obvious difference between *Myriad* and *Wal-Mart* is the type of science at issue: “hard” in the former, and “soft” in the latter.⁶⁹ As the *Myriad* concurrence makes clear, Scalia was well aware he was not a molecular biologist, but in *Wal-Mart*, he seemed truly unbothered by the fact that he was no sociologist. In fact, this is quite consistent with how social scientists are treated in broad culture. “With so many gifted amateurs working their territory,” some practitioners have reflected, “social scientists have a tougher time asserting the unique nature of their expertise than do astrophysicists, for example. Few of us know much about the dynamics of the cosmos, but we all know plenty about human nature—or at least we think we do.”⁷⁰ The problem, then, was not simply that Scalia did not recognize his own inexpertise, but moreover, in a sort of judicial Dunning-Kruger,⁷¹ so overestimated his grasp of the *Wal-Mart* subject matter that he believed himself an expert on Bielby’s level—or at least knew feigning ignorance would provide a convenient cover.⁷²

So what would we rather have Scalia done in *Myriad*? Should he have been forced to, like the Eleventh Circuit’s Judge Robin S. Rosenbaum, cram for his own “crash course in organic chemistry”⁷³ before weighing in, or had the Court already, sufficiently, comprehended the case’s essential components—even when the fine

69. The word “hard” ordinarily describes a type of scientific study that “[e]mploy[s] quantitative, objective, or rigorous analysis or testing; [such a study is] amenable to precise mathematical treatment or experimental verification or refutation.” *Compare Hard*, OXFORD ENGLISH DICTIONARY 9b, <http://www.oed.com/view/Entry/84122> [<https://perma.cc/4NUF-KXPR>], with *Soft*, OXFORD ENGLISH DICTIONARY 30c, <http://www.oed.com/view/Entry/183898> [<https://perma.cc/S4MG-YJMS>] (“[A] field or method of enquiry which is not amenable to precise mathematical treatment or to experimental verification or refutation[.]”).

70. *In Praise of Soft Science*, 435 NATURE 1003, 1003 (2004).

71. The theory, popularly known as the Dunning-Kruger effect, discusses the “dual burden” of illusory superiority which causes incompetent individuals to overestimate their abilities in intellectual pursuits for the very reason that their “incompetence robs them of the metacognitive ability to realize” that they are, in fact, incompetent. See Justin Kruger & David Dunning, *Unskilled and Unaware of It: How Difficulties in Recognizing One’s Own Incompetence Lead to Inflated Self-Assessments*, 77 J. PERSONALITY & SOC. PSYCHOL. 1121, 1121 (1999). This effect has long roots in philosophic thought: A quote Plato attributes to Socrates in his “Apology of Socrates” states that, “I am wiser than this man; for neither of us really knows anything fine and good, but this man thinks he knows something when he does not, whereas I, as I do not know anything, do not think I do either. I seem, then, in just this little thing to be wiser than this man at any rate, that what I do not know I do not think I know either.” See, e.g., PLATO, PLATO IN TWELVE VOLUMES, VOL. 1, 21d (Harold North Fowler, trans., 1966).

72. See *infra* Subpart II.B for additional discussion.

73. *United States v. Phifer*, 909 F.3d 372, 375 (11th Cir. 2018) (“After careful consideration and a crash course in organic chemistry, we conclude that . . . ‘positional isomer’ does not unambiguously apply to the use of that term as it pertains to butylone and ethylone in this case.”).

details eluded their grasp—to “work out the fundamental truths” at play?⁷⁴ More concerningly, how will judges know for certain when the latter approach will be sufficient in a given case, or whether the former approach is required? And if it is required, will they be capable of executing it?

II. SOCIAL SCIENCE AND THE PROBLEMS OF FOREIGN EXPERTISE AND UNFAMILIAR PROOF

A rather obvious point bears repeating: No matter how wise we may hope for them to be, judges are not scientists, and beyond that, some are even outright science illiterate.⁷⁵ In a more routine or unremarkable case, this might not pose a particularly urgent problem, as unpacking an average Code provision and applying it to the facts of a contracts dispute usually does not require judges be intimately familiar with science hard, soft, or in between. In these cases, judges operate fully within their wheelhouse, simply drawing on their particular, hard-won expertise and skills, gained through their legal education and long careers at the Bar and on the bench.

But in the “unusual” case, where a key evidentiary point requires establishing highly sophisticated or complex pieces of specialized knowledge, the judiciary’s scientific inexperience looms larger.⁷⁶ In these cases, amateur judges must trade robe for lab coat and “science the heck out of” things whether they know how to or not.⁷⁷ Absent truly remarkable luck, of course, they nearly always will not, and this

74. Mark Joseph Stern, *Scalia on DNA Patents: I Don’t Really Understand Science*, SLATE (June 13, 2013, 1:43 PM), <http://www.slate.com/technology/2013/06/myriad-dna-patenting-supreme-court-case-scalia-says-he-doesnt-get-the-science.html> [https://perma.cc/6YAY-TBYS].

75. See, e.g., Seth Stern, *Science-Savvy Judges in Short Supply*, CHRISTIAN SCI. MONITOR (Dec. 21, 2000), <http://www.csmonitor.com/2000/1221/p17s1.html> [https://perma.cc/C6ZK-BQLX] (discussing a survey of 400 state trial judges which found that some 50 percent believe their academic background did not adequately prepare them to handle scientific evidence on the job). Stern notes that Arizona Supreme Court Justice Thomas Zlaket likes to joke he attended law school to for once and all “escape” studying science and math. *Id.*

76. Of course, due to the modern era’s explosion of scientific progress, such cases cannot exactly be classified as wholly “unusual,” and may even be creeping toward the default. See, e.g., SHEILA JASANOFF, *SCIENCE AT THE BAR: LAW, SCIENCE, AND TECHNOLOGY IN AMERICA* ix–x (1995). “[W]e live in a time when the issues raised by science . . . are likely to become even more intertwined with basic decisions about how we live.” *Id.* at ix. “At the heart of the problem,” she writes, “are the rapid technological and scientific advances that make it ever more difficult for those involved in the judicial system and for the citizens who serve on juries to understand the complex information presented by expert witnesses.” *Id.* at x.

77. *Phifer*, 909 F.3d at 374. Judge Rosenbaum—quoting the 2015 space epic *The Martian*—opens her opinion by stating that, “There’s no easy way around it. We’re just going to have to science the heck out of this case. And when we’re done with that, we’re going to have to law the heck out of it.” *Id.*

is why expert witnesses are so crucial to modern litigation.⁷⁸ Notwithstanding above-referenced fears over partisanship, the legal system has for this reason long found ways to tolerate experts as a necessary evil, required to satisfy the system's insatiable need for final—if imperfect—decisions reached quickly and efficiently. Given rapid advancements in the development and accumulation of “new” knowledge and an ever-increasing need for specialization, there no longer exists any individual who could possibly know “everything,”⁷⁹ and as a result, many whole categories of claims would be entirely unprovable, or even nonexistent, absent expert help.⁸⁰

The Rules, too, have recognized this and adapted. The modern procedural framework, codified in Federal Rule of Evidence 702 and stemming from the *Daubert* trilogy of Supreme Court cases,⁸¹ treats judges as “gatekeepers,” requiring them to perform a screening function, whereby juries are insulated from unreliable, unhelpful, or otherwise misleading testimony.⁸² Instead, judges must

78. In a 2016 survey of “expert prevalence,” researchers estimated that expert witnesses appear in just over 85 percent of cases. See Andrew W. Jurs, *Expert Prevalence, Persuasion, and Price: What Trial Participants Really Think About Experts*, 91 IND. L.J. 353, 355 (2016) (noting that this percentage is “identical” to that in two prior research studies).

79. Scientist Thomas Young (1773–1829) was described in 2006 book as “The Last Man Who Knew Everything,” and was said to have “proved Newton wrong, explained how we see, cured the sick, and deciphered the Rosetta Stone, among other feats of genius.” ANDREW ROBINSON, *THE LAST MAN WHO KNEW EVERYTHING* i (2006). We might fancy the pre-information age Young—a polymath, or person of encyclopedic knowledge and wideranging expertise—as one of the final people capable of testifying as an expert in many or even all disciplines, and in his case on matters as diverse as physics, mathematics, physiology, medicine, linguistics, and Egyptology.

80. See JASANOFF, *supra* note 76, at ix (writing that “[t]he questions science puts before us today are well beyond the comprehension of most of us,” and stating that courts routinely resolve “questions that literally are beyond anything that could have been imagined when the foundations of the American legal system were set in place . . . [N]othing tests the courts’ capacity for resolving disputes and setting rules more than the legal issues raised by the scientific revolution.”).

81. The trilogy—made up of *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993), *General Electric Co. v. Joiner*, 522 U.S. 136 (1997), and *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999)—led to a 2000 amendment to Federal Rule of Evidence 702, the Rule governing “Testimony by Expert Witnesses.” As amended, the Rule sets out that a witness “qualified as an expert by knowledge, skill, experience, training, or education” may testify if:

- (a) the expert’s scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
- (b) the testimony is based on sufficient facts or data;
- (c) the testimony is the product of reliable principles and methods; and
- (d) the expert has reliably applied the principles and methods to the facts of the case.

FED. R. EVID. 702 (2000).

82. See 28 U.S.C. app. § 702 at 398. (“In *Daubert* the Court charged trial judges with the responsibility of acting as gatekeepers to exclude unreliable expert testimony, and the Court in

first grapple with that testimony alone, at an earlier procedural juncture, before giving their admissibility *pollice verso*.⁸³ This has led to a precarious regime in which judges consistently wield power incommensurate with ability, as they are hardly better than the average juror at differentiating junk from genius, due to what Dean Jennifer Mnookin might refer to as the judges' epistemic incompetence in the expert's field.⁸⁴ Casting judges as gatekeepers, then, seems a bit like handing them a whistle and asking them to referee a game without reading the rules. In most circumstances, it will be better to have them there than not, as even an amateur umpire should have no problems recognizing the obvious and most egregious fouls. But for the rulings on the margins, in disputes involving more complicated, specialized, or esoteric forms of analysis, this novicehood can stop being helpful and start feeling oppressive.

A. Shortcuts to Knowledge: The Judicial System's Reliance on Secondary Proxies for "Actual" Understanding

Given this inescapable gap in knowledge, judges cannot gatekeep with respect to the substance of expert testimony, as they are, not being subject matter experts, unqualified to verify the validity of that substance. Put another way, judges cannot distinguish good testimony from bad by simply drawing on the "reliable cognitive process" which one must be an expert to possess or access.⁸⁵

Kumho clarified that this gatekeeper function applies to all expert testimony, not just testimony based in science.").

83. *Pollice verso* refers to the turned thumb response, which spectators of Ancient Roman gladiatorial combat made to a "vanquished gladiator begging for his life," to decide whether he would live or die after defeat. Edwin Post, *Pollice Verso*, 13 AM. J. PHILOLOGY 213, 213 (1892). While the stakes are not quite so drastic for our battle-scarred experts, the power imbalance perhaps endures.
84. See Mnookin, *supra* note 19, at 1012. Experts are necessary, Mnookin writes, "precisely because of" what their audiences do not know, as these individuals lack the epistemic competence to intelligently parse the specialized evidence presented: "The second fundamental problem with the adversarial expert is epistemic [competence] . . . [Experts] are supposed to provide information useful to the jury's decision-making that goes beyond what a jury would know without their assistance. But if the jury lacks the knowledge that the expert provides . . . can it rationally evaluate the expertise on offer?" *Id.*
85. Broadly, the notion of "epistemic competence" to which Mnookin refers has roots in an epistemological theory known as "process reliabilism," defined as a "theory of knowledge according to which a belief is justified if it has been reached by a reliable cognitive process, whether or not the person who holds the belief understands the process." *Reliabilism*, OXFORD ENG. DICTIONARY, <http://www.oed.com/view/Entry/273664> [<https://perma.cc/38PR-E2PG>] (last visited Nov. 3, 2018). Under this theory, a mere "belief"—even a true one—transforms into actual knowledge only when it is obtained by such a reliable process; otherwise, these beliefs are not "justified" and are on some level indistinguishable from truths arrived at by accident. See generally Alvin Goldman & Bob Beddor, *Reliabilist Epistemology*, STAN. ENCYC. PHIL. (Dec. 2,

While a true peer from an expert's professional community, with similar preparation and experience, will surely be epistemically competent to examine the data and methodology of an expert report so as to informedly state whether the expert in question's interpretation is sound, an epistemically incompetent judge lacks the luxury. And because most judges cannot, like Scalia in *Myriad*, recuse themselves for lack of personal comprehension, they must instead find ways to create artificially educated guesses. This is done by turning to secondary or "proxy" indicia for reliability,⁸⁶ which include various superficial characteristics, together serving as an inexact barometer for trustworthiness, since these characteristics frequently—though do not always—correlate with valid substance. Accordingly, then-professor Elizabeth Warren referred to such shortcuts for knowledge as "quality signals."⁸⁷ And since all judges must utilize such proxy indicia in some manner, this means that when they are "right" in their gatekeeping, it is by the happy accident that they turned to a particular medley of proxy indicia, which, in a particular circumstance, mapped over the character of an expert possessing the "true" knowledge.

If this point seems at all abstruse—perhaps like epistemological gobbledygook—it is meant no more than as a way of academically expressing a problem Chief Justice Rehnquist noted instantly in his *Daubert* dissent. This new approach, Rehnquist worried:

2015), <http://plato.stanford.edu/archives/win2016/entries/reliabilism> [https://perma.cc/TD8R-8FTP]; see also Andrea Kern, *Knowledge as a Fallible Capacity*, in CONCEPTIONS OF KNOWLEDGE 230 (2011) ("Having an epistemic competence that explains knowledge means that, under the conditions proper to the competence question, one sufficiently frequently produces a true belief on the basis of a true seeming [or opinion].").

Such nonaccidental, true knowledge is obviously superior to "accidentally true" beliefs in a courtroom setting, where experts may be asked to testify about a wide variety of matters, and where certain "knowledge" may not be possible or practicable on every single point, and these "opinions" may be necessary and desirable. Kern goes on to explain that "an epistemic competence cannot consist in the fact that someone who exercises the competence *always* has a true belief. Rather it consists in the fact that, in sufficiently many cases, in which conditions appropriate to the competence in question are present, someone has true beliefs." *Id.* at 231. "More precisely," she writes, "it consists in the fact that, in *sufficiently many cases*, in which conditions appropriate to the competence in question are present, he has true seemings and trusts them, absent signs that they could be false." *Id.*

86. See Mnookin, *supra* note 19, at 1013 (writing that, while "one might not need to be an expert in order to assess expertise . . . the main mechanisms for assessing expertise outside of one's domain of knowledge are, by necessity, secondary indicia, proxies . . ."); see also PAUL M. SNIDERMAN ET AL., *Reasoning Chains*, in REASONING AND CHOICE: EXPLORATIONS IN POLITICAL PSYCHOLOGY 70–92 (1991); *id.* at 70 ("Our account of reasoning and choice is centered on the notion that people compensate for informational shortfalls by taking advantage of judgmental shortcuts, or heuristics.").

87. Elizabeth Warren, *The Market for Data: The Changing Role of Social Sciences in Shaping the Law*, 2002 WIS. L. REV. 1, 6.

[I]s particularly unfortunate [in cases] where the ultimate legal question depends on an appreciation of one or more bodies of knowledge not judicially noticeable

The various briefs filed in this case are markedly different from typical briefs, in that large parts of them do not deal with decided cases or statutory language—the sort of material we customarily interpret. Instead, they deal with definitions of scientific knowledge, scientific method, scientific validity, and peer review—in short, matters far afield from the expertise of judges. This is not to say that such materials are not useful or even necessary . . . but it is to say that the unusual subject matter should cause us to proceed with great caution in deciding more than we have to, because our reach can so easily exceed our grasp.⁸⁸

When the Court's reach exceeds its grasp, it is because the Justices have anchored their decisions in shaky stand-ins for reliability. Under this regime there exists a near-endless number of potential proxy indicia for reliable expertise; the Court in *Daubert* specifically anticipated proxies like “whether the [expert's] theory or technique has been subject to peer review, whether the results have been published, and . . . whether the expert's methods or reasoning enjoy general acceptance⁸⁹ in a relevant scientific community,”⁹⁰ but we should hardly stop there. We might imagine that most judges, consciously or not, will also weigh factors like an expert's age, dress, physical appearance, affect and manner of speech, apparent honesty, perceived prestige or pedigree, and clarity of expression and presentation. The issue, of course, is that the Mad Lib of, say, [a 50-something] with [a Ph.D.] from [an Ivy League university] and [a winning smile] wearing [a suit with elbow patches]—or any number of other permutations of proxy indicia combos—will not necessarily correspond neatly with actual reliability. Given this, experts who do not, or are unable to, perform the “role” of expert⁹¹ in a way the judge expects or wants will be irrationally punished at the gatekeeper stage for

88. *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 598–99 (1993) (Rehnquist, C.J., dissenting).

89. This language has roots in *Frye* which set out the pre-*Daubert* standard for expert admission. *Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923). *Frye* required that testified-to scientific principles enjoy “general acceptance” in their respective fields. *Id.* at 1014. Though *Daubert* overruled the *Frye* standard as the sole criterion for admission, Justice Blackmun stated that “‘general acceptance’ can yet have a bearing on the inquiry” as a secondary indication of reliability. *Daubert*, 509 U.S. at 594.

90. Joseph Sanders, *From Science to Evidence: The Testimony on Causation in the Bendectin Cases*, 46 STAN. L. REV. 1, 65 (1993).

91. See Mnookin, *supra* note 19, at 1013 (in the context of appealing to juries, writing that “[p]arties will, quite rationally, seek out precisely those experts most capable of ‘performing’ the role of expert in just the way that the parties expect that a jury will find credible”).

failing to live up to an expectation that may in fact correlate little, or not at all, with merit.

B. The Impact of Culture as a Proxy in Judicial Decisionmaking

An incredibly powerful—and perhaps the most impactful—secondary proxy for reliability with *no* correlation to merit is a particular proposition’s conformance or nonconformance with a judge’s personal beliefs and opinions, or what Masau Sagiv terms the judge’s “preexisting cultural agenda”⁹² and others have called a social group’s particular “standardized social procedures.”⁹³ One’s cultural agenda could encompass anything from simple “good sense” notions of prudent living to deeply personal, even sacred, matters such as one’s individual political leanings or religious faith. Technically, these considerations are outside the law, given that they do not normally have any independent procedural or substantive significance. But that does not stop them from holding significant sway for judges and jurors: “There is in each of us a stream of tendency,” Benjamin Cardozo, then serving on the New York Court of Appeals, wrote:

[W]hich gives coherence and direction to thought and action. Judges cannot escape that current any more than other mortals. All their lives, forces which they do not recognize and cannot name, have been tugging at them—inherited instincts, traditional beliefs, acquired convictions . . . [which] must determine where choice shall fall.⁹⁴

In some ways, Cardozo’s nameless “forces” are simply another way of conceiving of proxy knowledge itself. That is, judges do not rely on any reliable cognitive processes⁹⁵ in assessing a particular set of facts when culture influences their reaction to it; rather, they are being pulled by invisible and frequently unacknowledged forces that impact them at a more autonomous than conscious level of response. Regardless which specific beliefs they hold, given how deeply such cultural content is baked into human cognitive processes and functioning, it is simply not possible to fully extract it from decisionmaking on demand—and even less possible from the frequently momentous decisionmaking that, for

92. See generally Masua Sagiv, *Cultural Bias in Judicial Decision Making*, 35 B.C. J.L. & SOC. JUST. 229, 231 (2015). Sagiv defines culture as “constitutive of the personality identity of human beings, which is their ‘ability to preserve the attributes that are seen as central by them and the members of their group.’” *Id.* at 231. She writes that this includes “people’s informal knowledge about the world from their social group’s point of view.” *Id.* at 233.

93. David Bidney, *On the Philosophy of Culture in the Social Sciences*, 39 J. PHIL. 449, 449 n.3 (1942) (citation omitted).

94. BENJAMIN N. CARDOZO, *THE NATURE OF THE JUDICIAL PROCESS* 12 (1997).

95. See discussion *supra* note 85 and accompanying text.

instance, Supreme Court Justices must perform. And since “ordinary folk use the orienting force of culture to generate miles of preferences from only inches of fact,”⁹⁶ such decisions may be, simply, wrong.

A judge’s preexisting cultural agenda might undesirably infiltrate his or her legal reasoning in a number of ways—some beyond the scope of this Comment—but most relevant here is the nexus between both confirmation bias⁹⁷ and epistemic or cognitive immunization, which psychologists say explains why, when “people’s confidence in their beliefs is shaken, they become stronger advocates for those beliefs.”⁹⁸ All told, these concepts reflect that people, judges among them, “are well armed with defenses to ward off offensive or inconvenient knowledge.”⁹⁹ But what if this culture clash causes a judge to subconsciously distrust an expert, or seek out ways to minimize or misconstrue that expert’s testimony, if the judge’s own personal beliefs are being contradicted on the stand? Even worse, what if a judge’s personal moral commitments run directly contrary to his or her oath to uphold the U.S Constitution?

Whether they lie at the center or periphery of a case, these emotionally laden topics which are so frequently the focus of social science research—such as research on race, gender, sexuality, politics, or religion—may sneak in and infect judicial determinations, and whether consciously or not, judges may seize on the opportunity to give legal weight to their personal views. This is not to say that no judge or Justice will be able to make principled rulings when their preferred “side” (if one exists) is attacked or contradicted; instead, it is meant to suggest that on a large scale, it is inevitable that, in claims centering around cultural agenda-challenging propositions, at least some, and perhaps many, judges will interpret cultural conflict as a proxy for unreliability and thus discard or minimize relevant evidence and research. For most of us, the notion that cDNA is a synthetic creation, not normally present in nature, does not rock the foundations

-
96. John Gastil et al., *The Cultural Orientation of Mass Political Opinion*, 44 PS: POL. SCI. & POL. 711, 711 (2011) (internal quotation marks omitted) (quoting Aaron B. Wildavsky, *Choosing Preferences by Constructing Institutions: A Cultural Theory of Preference Formation*, 81 AM. POL. SCI. REV. 3, 8 (1987)).
 97. Phoebe C. Ellsworth, *Legal Reasoning and Scientific Reasoning*, 63 ALA. L. REV. 895, 904 (2012) (“Confirmation bias is the tendency to seek, believe, and remember information that agrees with what we already think. There are of course motivational reasons for this bias—we want to believe that we perceive the world as it really is and that people who disagree with us are wrong.”).
 98. Ed Yong, *When in Doubt, Shout—Why Shaking Someone’s Beliefs Turns Them Into Stronger Advocates*, DISCOVER MAG. (Oct. 19, 2010), <http://blogs.discovermagazine.com/notrocketscience/2010/10/19/when-in-doubt-shout-why-shaking-someones-beliefs-turns-them-into-stronger-advocates> [https://perma.cc/XBQ5-27LN] (emphasis omitted).
 99. Stewart Macaulay, *The New Versus the Old Legal Realism: Things Ain’t What They Used to Be*, 2005 WIS. L. REV. 365, 397.

of our belief systems. That women face systemic oppression across the country because those in positions of power do not check their subconscious biases, however, just might.

But why must this judicial behavior implicate social science specifically? Or, rather, what makes the issue especially relevant to the domain of social science when it would seem endemic to the entirety of the law? The answer lies in the notion that social science may itself be inextricably linked with culture. As Sagiv writes, “[c]ultural information is different from other scientific information—it is inseparable from society and politics, and it derives from them.”¹⁰⁰ Moreover, some scholars have suggested that all legal uses of social science, even that “designated as [being] objective, neutral, and scientific uses of objective, neutral, and scientific facts,” might in fact be “inherently political and/or ideological statements, which are shielded and obscured by the scientific mantle in which they are wrapped.”¹⁰¹

These ideas are consistent with the “politics of knowledge” theory, which acknowledges that all scientific research occurs inside of certain social contexts which themselves impact the means by which new or existing ideas are accepted into scientific canons. This exercise is further complicated when such research, occurring inside of a social context, is also about social contexts. An “[a]wareness of the role of the social scientist as a human and political being in shaping research outcomes,” scholars have written, “has brought an unsettling relativity into the calculation.”¹⁰² In 2018, Mitchell Langbert found that the average “Democratic-to-Republican ratio” across a sample of professors with a registered political affiliation at the country’s top-ranked liberal arts colleges was 10.4:1, with social sciences like sociology and psychology having ratios of 43.8:1 and 16.8:1, respectively.¹⁰³ Given this fact, we might wonder whether Chief Justice Roberts used “sociological gobbledygook” to connote not just “unintelligible,” but rather something closer to “liberal trash.”

As a result, for some, the Court’s reticence toward social science may be a feature, not a bug. Because the Justices strongly resist wading into the political

100. Sagiv, *supra* note 92, at 252.

101. Martha L. Fineman & Anne Opie, *The Uses of Social Science Data in Legal Policymaking: Custody Determinations at Divorce*, 1987 WIS. L. REV. 107, 110.

102. Howard Erlanger et al., *Is It Time for a New Legal Realism*, 2005 WIS. L. REV. 335, 342.

103. Mitchell Langbert, *Homogenous: The Political Affiliations of Elite Liberal Arts College Faculty*, 31 ACAD. QUESTIONS 186, 186, 190 (2018). In an article discussing these findings, Cass R. Sunstein notes that “[t]he gap is narrower in science and engineering. In physics, economics and mathematics, the ratio is about 6 to 1. In chemistry, it is 5 to 1, and in engineering, it is just 1.6 to 1.” Cass R. Sunstein, *The Problem With All Those Liberal Professors*, BLOOMBERG (Sept. 17, 2018), <http://www.bloomberg.com/opinion/articles/2018-09-17/colleges-have-way-too-many-liberal-professors> [<https://perma.cc/LP5D-KJZE>].

“thicket” when they can help it,¹⁰⁴ resisting undue influence or overdependence on what may be the inherently political statements of social science could seem consistent with the Court’s wanting to leave such matters to the Legislature, thus protecting separation of powers principles especially important to the Court and its current five-Justice conservative majority. But to adopt such an approach as an across-the-board practice would seem to ignore an impressive track record—at least under the “conventional narrative”¹⁰⁵ of *Brown v. Board of Education*.¹⁰⁶ With apologies to Justice Scalia, the intersection of social sciences and the law is in fact one of the few contexts in which it is quite appropriate to begin “waving the bloody shirt of *Brown* again.”¹⁰⁷ In the decision’s influential footnote 11, Chief Justice Warren rested the Court’s rejection of separate but equal on sociological research, including Kenneth and Mamie Clark’s “doll tests,” an experiment showing the psychic havoc wreaked on black schoolchildren during Jim Crow-era segregation.¹⁰⁸ Warren wrote that the District Court’s claim that “[s]egregation of white and colored children in public schools has a detrimental effect upon the colored children . . . [and this] impact is greater when it has the sanction of the law; for the policy of separating the races is usually interpreted as denoting the inferiority of the negro group” was “amply supported by modern authority.”¹⁰⁹ His footnote then cited to the Clarks’s and six other social science studies

104. For example, in *Colegrove*—where the Court held that the concept of “one person, one vote” did not guarantee districts a right to proportional representation—Justice Frankfurter resisted entering the so-called political thicket. *Colegrove v. Green*, 328 U.S. 549, 556 (1946) (writing that “[t]he remedy for unfairness in districting is to secure State legislatures that will apportion properly, or to invoke the ample powers of Congress”). Just fifteen years later, the Court would overrule *Colegrove* in *Baker v. Carr*, 369 U.S. 186 (1962).

105. See Sanjay Mody, *Brown Footnote Eleven in Historical Context: Social Science and the Supreme Court’s Quest for Legitimacy*, 54 STAN. L. REV. 793, 793 (2002) (discussing “the customary view” that the Court “relied on social science evidence, cited in footnote eleven of the *Brown* opinion, to find that racial segregation caused social and psychological harm to black schoolchildren”).

106. 347 U.S. 483 (1952).

107. “Waving the bloody shirt” makes reference to the Reconstruction-era “standard [political] retort . . . [an] expression of dismissive Southern contempt whenever a Northern politician mentioned any of the thousands upon thousands of murders, whippings, mutilations, and rapes that were perpetrated against freedmen and women and white Republicans in the South.” STEPHEN BUDIANSKY, *THE BLOODY SHIRT: TERROR AFTER APPOMATTOX* 24 (2008). During public appearances throughout his time as a Justice, Scalia was asked so frequently in “public appearances” about the way in which he would have voted in *Brown* “that he [would] say things like ‘Waving the bloody shirt of *Brown* again, eh?’” See Ronald Turner, *A Critique of Justice Antonin Scalia’s Originalist Defense of Brown v. Board of Education*, 62 UCLA L. REV. DISCOURSE 170, 172 (2014) (citation omitted).

108. See Kenneth B. Clark et al., *The Effects of Segregation and the Consequences of Desegregation*, 59 AM. PSYCHOL. 495 (2004) (appendix to appellants’ briefs in the *Brown v. Board of Education*, *Briggs v. Elliott*, and *Davis v. Prince Edward County, Virginia*, cases, coauthored by Kenneth Clark).

109. *Brown*, 347 U.S. at 494.

confirming the fact, buttressing the Court's holding that de jure racial segregation violates the Equal Protection Clause.¹¹⁰

Brown is without question the high-water mark for a collaboration between social science and the law: In a 1955 article, Edmond Cahn even wrote that the decision "added to the dignity and stature of every American."¹¹¹ Warren's inclusion of the studies in his opinion allowed the emerging social awareness of the harms of segregation to be something which the Court could take into account and "permissibly learn" for the first time.¹¹² Relying on and citing to social science, then, gave the Court a means by which it could overcome its long-irrational and indefensible approach, as "[t]he Court that refused to see inequality . . . would be making the only kind of law that can be warranted outrageous in advance—law based on self-induced blindness, on flagrant contradiction of known fact."¹¹³

Despite supporting the holding itself, however, Cahn was critical of the Court's use of social science, writing that he "would not have the constitutional rights of Negroes—or of other Americans—rest on any such flimsy foundation as some of the scientific demonstrations in these records."¹¹⁴ "We don't need evidence for the proposition that segregation is an insult to the Black community," Ronald Dworkin wrote, summarizing Cahn's critique, "[W]e know it; we know it the way we know that a cold causes snuffles."¹¹⁵

Beyond this more fundamental critique, however, Cahn suspected that the Justices' reliance on the studies was a convenient justification meant to support or even compel their desired decision. The evidence, Cahn wrote, "seemed persuasive because it happened to coincide with facts of common knowledge."¹¹⁶ This fear resembles that of appellate attorney John W. Davis, arguing for the defense in *Briggs v. Elliot*,¹¹⁷ one of five school segregation cases eventually consolidated into *Brown*: "[M]uch of that which is handed around under the name of social science is an effort on the part of the scientist to rationalize his own preconceptions. They find usually, in my limited observation, what they go out to

110. *Id.* at 494–95 n.11.

111. Edmond Cahn, *Jurisprudence*, 30 N.Y.U. L. REV. 150, 150 (1955).

112. See Black, *supra* note 27, at 427.

113. *Id.* at 426.

114. Cahn, *supra* note 111, at 157–58. Social science as a discipline, Cahn believed, was "so very young, imprecise, and changeful, [that] their findings have an uncertain expectancy of life. Today's sanguine asseveration may be cancelled by tomorrow's new revelation—or new technical fad." *Id.* at 167.

115. Ronald Dworkin, *Social Sciences and Constitutional Rights—The Consequences of Uncertainty*, 6 J.L. & EDUC. 3, 5 (1977).

116. Cahn, *supra* note 111, at 166.

117. 342 U.S. 350 (1952).

find.”¹¹⁸ In a more contemporary reflection, Lawrence M. Friedman wrote that he “would be very surprised . . . if Warren had read the works cited in footnote 11” at all.¹¹⁹ All told, under this view of *Brown*, the Court relying on footnote 11 was simply a smokescreen from behind which the Justices could grant their preexisting “gut feelings” regarding de jure segregation the force of law. Armchair sociologist Chief Justice Warren did not assess the studies’ substance (if, as Friedman questions, he even read them in the first place); instead, he trusted the proxy of the studies’ conformance with his own preexisting cultural agenda for their reliability. This would mean that social science’s greatest victory in the courtroom was simply to be in service of the same mechanism by which it is so frequently excluded: because it disagrees with the Court majority’s preexisting cultural agenda.

In *Exxon Shipping Co. v. Baker*, in which the Court mulled the propriety of a \$2.5 billion punitive damages award set against the corporation for causing the 1989 Exxon Valdez oil spill, Justice Souter engaged in a particularly striking example of this sin in reverse.¹²⁰ Prior to the case, a number of prominent scholars, including Cass R. Sunstein, Daniel Kahneman, and William R. Freudenburg, had conducted research into the psychology of juries when determining damages, and the consistent “incoherence”¹²¹ of the award amounts. Though these articles seemed to consist of high-quality original research and were published in prestigious legal journals like the *Columbia Law Review*, the *Yale Law Journal*, and the *Journal of Law & Human Behavior*,¹²² several of the studies had originally or in part been funded by Exxon itself.¹²³ While Justice Souter ultimately agreed with the studies’ basic propositions—that there is a “stark unpredictability” in punitive damage awards set by different juries confronted with the same facts—he

118. ANGELO N. ANCHETA, SCIENTIFIC EVIDENCE AND EQUAL PROTECTION OF THE LAW 53 (2006) (noting that Davis “spent several minutes attacking Clark’s doll studies, the testimony of two other experts, and the value of social science evidence in general”).

119. Lawrence M. Friedman, *Brown in Context*, in RACE, LAW, AND CULTURE: REFLECTIONS ON BROWN V. BOARD OF EDUCATION 49, 61–62 (Austin Sarat ed., 1997).

120. See *Exxon Shipping Co. v. Baker*, 554 U.S. 471 (2008).

121. See, e.g., Cass R. Sunstein et al., *Predictably Incoherent Judgments*, 54 STAN. L. REV. 1153, 1173 (2002).

122. See Alan Zarembo, *Funding Studies to Suit Need*, L.A. TIMES (Dec. 3, 2003), <https://www.latimes.com/nation/la-na-valdez3-2003dec03-story.html> [<https://perma.cc/GZP4-XSUJ>].

123. See, e.g., Adam Liptak, *From One Footnote, a Debate Over the Tangles of Law, Science and Money*, N.Y. TIMES (Nov. 24, 2008), <http://www.nytimes.com/2008/11/25/washington/25bar.html> [<https://perma.cc/T53P-W598>] (“Freudenburg . . . took Exxon’s money and conducted preliminary research. Exxon stopped supporting the study when the early findings did not point in a direction helpful to the company. But Exxon did help pay for several studies critical of punitive damages that appeared in places like The Yale Law Journal and The Columbia Law Review.”).

specifically resisted resting the Court's holding on them.¹²⁴ Instead, Souter wrote that he was "aware of a body of literature running parallel to anecdotal reports, examining the predictability of punitive awards by conducting numerous 'mock juries,' where different 'jurors' are confronted with the same hypothetical case. Because this research was funded in part by Exxon, we decline to rely on it."¹²⁵

In many ways, Souter's controversial "footnote 17"¹²⁶ is a veritable grab bag of topics addressed already in this Comment. While his approach seems clearly motivated by a fear of Big Oil infiltrating and pulling the strings of academic discourse in order to influence judicial outcomes, his open privileging of amateur "anecdotal reports" and judges' own informal experiences over legitimate social science scholarship is all too familiar, and this "skepticism clearly had ramifications for the credibility of social science evidence more generally."¹²⁷ Like Warren in *Brown*, we might conceive of this choice as being the result of Souter's preexisting cultural agenda—that is, that the Court should strive to preserve the judiciary from the taint of money. With an allegiance to this goal as his foremost motivator, Souter was forced to remove from conscious consideration any materials offensive to that goal, with soft science serving as the predictable sacrificial lamb. Thankfully, perhaps, because of the fortuitous convergence between anecdotal reports and scholarly consensus, Souter did not have to confront a more troubling question: What if the correct result had in fact turned on his accepting the Exxon-funded studies' conclusions, but he had still ignored them?

The Court has time and again faced this question in capital punishment cases, an area of law about which social scientists have reached a damning consensus the Justices' preexisting cultural agendas keep them from accepting, as, "when the death penalty is involved, lay people reason less rationally, and so do Supreme Court Justices."¹²⁸ "In case after case," Phoebe C. Ellsworth writes in an influential 1988 chapter attacking what she sees as the Court's intellectually dishonest death penalty jurisprudence, "the majority of the Justices have been faced with empirical research that supports

124. *Exxon Shipping Co.*, 554 U.S. at 499.

125. *Id.* at 501 n.17 (citations omitted).

126. See Debra Cassens Weiss, *Footnote 17 in Exxon Decision Gets Dissed and Dissected*, ABA J. (Nov. 25, 2008), http://www.abajournal.com/news/article/footnote_17_in_exxon_decision_gets_dissed_and_dissected [<https://perma.cc/H5CN-59N2>].

127. Moran, *supra* note 19, at 531.

128. Phoebe C. Ellsworth, *Unpleasant Facts: The Supreme Court's Response to Empirical Research on Capital Punishment*, in *CHALLENGING CAPITAL PUNISHMENT: LEGAL AND SOCIAL SCIENCE APPROACHES* 177, 182 (Kenneth C. Haas & James A. Inciardi eds., 1988).

an outcome they do not want.”¹²⁹ As a result, the Court—over multiple decades, and despite repeatedly being presented with research about the death penalty’s nonexistent deterrent effect,¹³⁰ its racially discriminatory application,¹³¹ or the fact that excusing strong death penalty opponents at voir dire creates juries biased towards a guilty verdict¹³²—has consistently found ways to dismiss the data as either inconclusive, irrelevant, or otherwise not worth considering.¹³³

In *McCleskey v. Kemp*, in which social science research on “race-of-victim effects”¹³⁴ was directly at issue, the Court considered a study by David C. Baldus, Charles Pulaski, and George Woodworth¹³⁵ which “had controlled for over 200 other variables and found that none of them—alone or in combination—could explain the pattern of racial discrimination” in death penalty sentencing.¹³⁶ Ellsworth writes that, in his majority opinion, Justice Powell:

[M]ade a faint-hearted attempt to question the conclusiveness of the research by arguing that the correlational data presented could not actually *prove* that race was the cause of the sentencing disparities History, common sense, and the data all overwhelmingly supported the simple, obvious explanation. The racial effects were due to race.

...

Once again, it [wa]s clear that the majority Justices . . . value the death penalty so highly that they are willing to ignore the inescapable implications of the social science research and to undermine fundamental constitutional guarantees.¹³⁷

McCleskey shows that—just as in *Exxon Shipping Co.*, *Wal-Mart*, or *Gill*—the Court simply does not value social science as highly as it does other forms of

129. *Id.* at 178.

130. *Id.* at 178–82.

131. *Id.* at 182–89.

132. *Id.* at 189–204.

133. *Id.* at 178.

134. *McCleskey v. Kemp*, 481 U.S. 279, 355 (1987). For instance, Marvin E. Wolfgang and Marc Riedel found that if a “defendant is black and the victim is white, the defendant is about eighteen times more likely to receive the death penalty than when the defendant is in any other racial combination of defendant and victim.” Marvin E. Wolfgang & Marc Riedel, *Race, Judicial Discretion, and the Death Penalty*, 407 ANNALS AM. ACAD. POL. & SOC. SCI. 119, 132 (1973).

135. David C. Baldus, Charles Pulaski & George Woodworth, *Comparative Review of Death Sentences: An Empirical Study of the Georgia Experience*, 74 J. CRIM. L. & CRIMINOLOGY 661 (1983).

136. Ellsworth, *supra* note 128, at 188.

137. *Id.*

knowledge, and certainly not enough to dislodge a majority of Justices from their preexisting positions should cultural conflict arise. Instead—as in the Friedman view of *Brown*—it is, at best, window dressing for a predestined holding with which the evidence happens to accord, or, at worst, a pest that the Justices must swat away, lest the conclusions that stem from it frustrate or challenge a preferred outcome. In either instance, the social science is sidelined, fated to become a prompt for a judicial creative writing exercise: *Given this science, Judge, how will you still manage to say whatever you'd been planning to already?*

C. The Judiciary's Normative Expectations for Good Experts and Expertise

Beyond the substance of research, another highly influential factor contributing to social science's repeated exclusion from the courtroom is not just its content, but form. Given features inherent to both the discipline itself, and the ways in which its specialists pursue and gain knowledge, social scientists will disproportionately fail to conform to the judiciary's normative expectations for what (1) good experts and (2) good expert evidence should look like. Time and again, as a result, this unfamiliarity serves as a strong proxy for unreliability.

1. The Archetypal Expert

Not unlike the tort fiction of a never-negligent reasonable person,¹³⁸ we might also imagine a sort of “archetypal expert,” a perfect figurehead embodying all of the judiciary's hopes for desirable testimony, and against whom judges will implicitly compare all comers at the *Daubert* gatekeeper stage. This individual performs the “role” of expert superbly. She is well-qualified and credentialed. She is brilliant and whip-smart, speaking intelligently but not inaccessibly. She is experienced, but no expert-for-hire. She is a hedgehog,¹³⁹ fully immersed in the relevant field, and not bogged down by extraneous knowledge at its edges which might unnecessarily complicate her perfect and 100 percent certain conclusions. She speaks specifically, without qualification, and without hesitancy, satisfying the law's need for finality and immediate resolution by pointing to objectivity and

138. RESTATEMENT (SECOND) OF TORTS § 283 cmt. c (AM. LAW INST. 1965).

139. See generally Maxine D. Goodman, *A Hedgehog on the Witness Stand—What's the Big Idea?: The Challenges of Using Daubert to Assess Social Science and Nonscientific Testimony*, 59 AM. U. L. REV. 635, 637, 638, 673 (2010) (conceiving of experts as either hedgehogs, who burrow into “one big idea” and “impel[] the judge toward certainty,” or foxes, who apply a “more open-minded cognitive approach [and] who tend to see alternative explanations and embrace ambiguity...”).

consensus. She does not confront judges with the impossible contours of a prolonged and difficult debate about which reasonable minds currently or might someday disagree, and this fact will keep the judge safe from being made a fool years or decades down the road. Above all, she is too good to be true. No such individual could possibly exist. But if an expert is able to check off enough boxes (or, to use our parlance, embody enough of these secondary proxies), her testimony will be overwhelmingly privileged at the gatekeeper stage.

In many respects, the *Dukes* plaintiffs' expert sociologist William Bielby embodied the archetype to a T, possessing a large number of what are typically deemed highly-relevant proxy indicia for a reliable expert: He had impeccable academic credentials and more than eighty career publications, as well as numerous professional awards, honors, and commendations earned over decades of work.¹⁴⁰ At the time he authored his original report in February 2003, he was serving as President of the American Sociological Association,¹⁴¹ and that same year, he published a scholarly article on virtually the same subject he testified to in *Wal-Mart*,¹⁴² suggesting he was not merely speculating on matters outside his professional specialty to earn a quick buck. Beyond these "positives," however, were qualities a skeptical judge or Justice might reflexively have viewed as proxies for unreliability. These could vary from the shallow—even if amusing—qualities, such as Bielby's time dabbling as an amateur musician, subverting the clichéd view of stuffy academia and once earning him the moniker "the Rock 'n Roll Sociologist"¹⁴³, to the pejorative, such as Bielby's

140. See William T. Bielby, *Curriculum Vitae*, 1–2 (Feb. 2017), http://cpi.stanford.edu/affiliates/cv/bielby_williamCV.pdf [<https://perma.cc/C3P3-6ALP>]. Bielby earned a master's degree in Social Sciences from the University of Illinois at Urbana-Champaign (1972), his Ph.D. in Sociology from the University of Wisconsin (1976), and obtained a prestigious fellowship in the Center for Advanced Study in the Behavioral Sciences at Stanford University (1983–84). *Id.* at 1–2. In his academic career, Bielby had professorships at the University of California, Santa Barbara (1977–2005, in the Department of Sociology, with years 1992–98 spent as Department Chair; and 1993–2005, as an Affiliated Professor in the Department of Statistics and Applied Probability), the University of Pennsylvania (2005–07), UCLA (1985), Northwestern University (2004), and the University of Arizona (2010–16), among others. *Id.* at 1. Bielby specialized in and taught graduate and undergraduate courses about organizational behavior, labor market, quantitative methods, social inequality, and research methods for the social sciences. See Bielby Declaration ¶ 3, *Dukes v. Wal-Mart Stores, Inc.*, 222 F.R.D. 137 (N.D. Cal. 2004) (No. 01-2252), 2003 WL 24571701. As for scholarship, Bielby has published essentially continuously from 1970 through the present, see Bielby, *Curriculum Vitae*, *supra* note 138, at 3–10.

141. See Bielby, *Curriculum Vitae*, *supra* note 140, at 2. Based on this and his résumé generally, we might conclude—admittedly, based on an additional layer of secondary proxies—that Bielby was not only accepted by his peer community, but was a leading voice within it.

142. See generally Bielby, *supra* note 49.

143. Roger Friedland, *Rock 'n Roll Sociologist: William T. Bielby*, FOOTNOTES, Sept./Oct. 2002, <http://www.asanet.org/sites/default/files/savvy/footnotes/septoct02/indexthree.html>

history of testifying in a few prior sex discrimination cases, which for some might raise “hired gun” concerns.¹⁴⁴

Considered alone, however, these more superficial proxies seem insufficient to explain Scalia’s decision. What is missing is a consideration of not just Bielby the Individual, but of his actual testimony, and—something which should give us great pause—its substance. That is, judges will want not just for the archetypal expert to appear as such, but for her to come wielding the archetypal proof, the type that, as we know, the judiciary will lack the epistemic competence to independently construct.

2. The Perfect Proof

The Rules purport to create a regime in which the “perfect proof” does not vary from one substantive context to the next: That is, a class claim certified under Rule 23 has the same “commonality” or “typicality” standard whether it asserts Title VII discrimination, 42 U.S.C. § 1983 violations, or securities fraud. Regardless of underlying substance, the Rules require plaintiffs meet an identical quantum of proof in order to satisfy relevant benchmarks, and courts, procedurally speaking, should not scrutinize a civil rights violation any differently from a wrong alleged under the Wilderness Act. This principle is known as transsubstantivity.¹⁴⁵ Though the concept is highly influential in civil litigation and in both the development and administration of the Federal Rules, scholars like

[<https://perma.cc/37AE-CG8Y>] (describing Bielby’s longtime fantasy of reuniting his high school rock band “The Newports,” a dream he made a reality in a special show at the American Sociological Association’s 2002 Annual Meeting, a performance which reportedly “really rocked”).

144. See Bielby Declaration, *supra* note 140, ¶ 2. By 2003, Bielby had testified as an expert witness in both California Superior Court and federal District Court in cases involving workplace discrimination like in *Wal-Mart*. E.g., *Stender v. Lucky Stores, Inc.*, 803 F. Supp. 259, 301 (N.D. Cal. 1992) (testifying in a case where he had examined “employment policies regarding job assignment and promotion, and [used] his expertise in the areas of the organization of work and gender in work to explain the kinds of gender based job assignment and promotion patterns that he observed”).

Additionally, Bielby had served as an expert “in several other cases involving issues of gender discrimination in large, multiestablishment national and regional retail firms, including class action cases involving Lucky Stores, Publix, Sherwin-Williams, and Home Depot.” See Bielby Declaration, *supra* note 140, ¶ 2. Richard Nagareda has casually referred to Bielby “as a repeat-player expert.” Richard A. Nagareda, *Class Certification in the Age of Aggregate Proof*, 84 N.Y.U. L. REV. 97, 154 n.218 (2009). Given that Scalia quoted directly from Nagareda’s article in his opinion, we might wonder how significant an impact the categorization made. See *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 350 (2011).

145. See David Marcus, *Trans-Substantivity and the Processes of American Law*, 2013 BYU L. REV. 1191, 1194 (writing that transsubstantivity “is one of the most fundamental principles of doctrinal design for modern civil procedure”).

J. Maria Glover have questioned whether the Roberts Court's bipolar class action jurisprudence requires us to view at least some of the Rules as in fact operating "non-transsubstantively."¹⁴⁶ In this view, the required quantum hinges not on the applicable Rule, but on the type of claim that Rule is facilitating, and the application requires the Court to supply meaning to, and make substantive judgments as to how, the laws in question should operate. The exercise for plaintiffs, then, becomes about satisfying the Justices' often never-before-articulated expectations for conduit theories of liability.¹⁴⁷ In doing so, the safest course will be to appeal to Justices' presumed preferences for proxies under the particular law, such as those seen as implicating the Justices' larger judicial philosophies¹⁴⁸ or cultural beliefs¹⁴⁹—"messier" proxies, correlating weakly, or again perhaps even not at all, with validity.

In a case with unfamiliar or irregular substance—as in a historically large one like *Wal-Mart*, or a long-nonjusticiable one like *Gill*—decisionmaking will rely quite crucially on these messy proxies, and in fact, the irregularity itself may be determinative. One explanation for this is the simple, powerful familiarity heuristic, which dictates that "familiar equals better equals safer."¹⁵⁰ This cognitive bias leads not just to the conclusion that the familiar is safer, but that the unfamiliar is dangerous. Charlatanry. Even gobbledygook. And because unfamiliar liability

146. See generally J. Maria Glover, *The Supreme Court's "Non-Transsubstantive" Class Action*, 165 U. PA. L. REV. 1625 (2017).

147. Glover writes that, in *Wal-Mart*, the Roberts Court's newly articulated Rule 23(a) standard is not what sunk the plaintiffs. *Id.* Rather, the Court interpreted the plaintiffs' conduit theory of liability as being invalid "as a matter or substantive Title VII liability policy," and "whether plaintiffs had satisfied the new commonality standard would have simply required an evaluation of whether the class had presented sufficient evidence, at the certification stage, to support that theory. Instead, plaintiffs could not satisfy the commonality requirement . . . because the Court . . . refused to interpret Title VII as encompassing plaintiffs' conduit theory." *Id.* at 1637, 1640.

148. See generally Glover, *supra* note 146, at 1626–27. Glover contrasts the outcome in *Wal-Mart* with the Court's opposite holding in *Tyson Foods v. Bouaphakeo*, 136 S. Ct. 1036 (2016), in which the Justices "vindicated the use of statistical proof to satisfy Rule 23 requirements." Glover, *supra* note 146, at 1626. *Tyson Foods* dealt not with Title VII, but FLSA, and Glover writes that "the majority's predominance analysis in *Tyson* necessarily rests upon interpreting FLSA as not requiring a perfect 'fit' between a damages model and the underlying theory of FLSA liability, at least for certification purposes." *Id.* at 1643. "In *Tyson*," she writes, "*y* is an acceptable approximation of *x* under the FLSA, and a common model that calculates *y* thus satisfies [FLSA's Rule 23 analogue]. The procedural consequences flow from the interstitial lawmaking vis-à-vis the FLSA," rather than from any procedural holdings in other cases. *Id.*

149. See discussion *supra* Subpart II.B (discussing the influence of culture in judicial decisions).

150. See, e.g., Wray Herbert, *Heuristics Revealed*, OBSERVER MAG., Oct. 2010, <http://www.psychologicalscience.org/observer/heuristics-revealed> [<https://perma.cc/BYLL8-XMDK>] ("The familiarity heuristic is one of the most robust heuristics known It is a potent mental tool that we draw on every day for hundreds of decisions, and basically what it says is this: if something comes quickly to mind, trust it.").

theories send up gigantic proxy red flags of invalidity, the judiciary will pore over newfangled proof for signs of either the familiar or foreign. In this process, questions such as “*Is this case simply ‘too big,’ based on my past experience as a judge, to be a valid Title VII action?*”¹⁵¹ or “*Does sex discrimination, in my personal opinion, actually take place in this manner or to this extent?*”¹⁵² will take on significant importance, despite the fact that “size of class” or “personal experience or beliefs of the judge” have no bearing on reality. This is the true curse of the adventuresome plaintiff—having to navigate not-necessarily deserved judicial skepticism for the mere fact that their conduit theory of liability looks “different” from theories that have been long accepted.

And what if this irregular or unfamiliar conduit theory exists only because of, or can be animated only through, scientific research? In this case, the decision will morph into another—but this time more vicious—Mad Lib: What is [*this Justice’s*] subjective belief as to what [*a Title VII claim*] brought under [*Rule 23*] usually requires, and does the plaintiffs’ unfamiliar [*method of scientific proof*] conform closely enough with this usual requirement? Judicial approval, then, may turn on the judge accepting particular “types” of proof—in *Wal-Mart* and *Gill*, social science research—that he or she will primarily examine for signs of unfamiliarity and dangerousness. This is where a bias against social science becomes oppressive, as it could shut out claims that can only be supplied content through unfamiliar or otherwise alien soft science help. Consider once again the sphere of voting rights litigation. Election law scholar Richard H. Pildes has written that:

Law and social science are perhaps nowhere more mutually dependent than in the voting-rights field. Empirical studies in this area are not merely of evidential significance; nor is cutting-edge social-scientific research simply a relevant resource to which litigants and courts might turn as one potential tool for interpreting and applying the law. Instead, the critical elements of the cause of action that the [VRA] . . . creates are

151. See, e.g., William H. J. Hubbard, *Optimal Class Size, Dukes, and the Funny Thing About Shady Grove*, 62 DEPAUL L. REV. 693 (2013).

152. See, e.g., Linda Hamilton Krieger & Susan T. Fiske, *Behavioral Realism in Employment Discrimination Law: Implicit Bias and Disparate Treatment*, 94 CALIF. L. REV. 997, 1004 (2006). The authors argue that the law’s focus on traditional formulations of intent is misguided because the law’s inexpert conception of how and why discrimination occurs is flawed: what judges call “intentional discrimination” can “just as easily result from the uncontrolled application of implicit, unconscious, or automatic stereotypes and other subtle ingroup preferences as from the operation of conscious discriminatory designs.” *Id.* In response, research like Dr. Bielby’s means to account for this uncertainty, and attempts to locate corporate organizational features which may explain how unconscious stereotyping causes the statistical disparities in the data. See text accompanying *supra* note 60.

defined in terms of legal concepts that necessarily must be given content through the kind of data that social-scientific analysis makes available.¹⁵³

The VRA, by necessity, does not focus on injury at the “individual” level, and instead searches for harms experienced by entire groups,¹⁵⁴ quite like the larger discipline of social science itself: “Most social sciences,” Ellsworth has written, “[is] not concerned with specific cases, except when they suggest an exception to a general rule or assumption.”¹⁵⁵ So while courts may want an opinion about how *this* defendant in *this particular* instance acted, an expert social scientist will feel more at home reasoning from group data to describe what is *likely* to have happened to a person *similar to* the defendant in a situation *analogous to* the defendant’s.¹⁵⁶

In VRA litigation, the judiciary’s preferred method is not an option. Given this fact, if the legal system does not sanction—or at least tolerate—a means, scientific or otherwise, by which VRA plaintiffs may creatively but credibly prove aggregate harms, then the Act would be rendered legally inoperative, and the judiciary would be acting in direct contravention of Congressional will.¹⁵⁷

153. Richard H. Pildes, *Is Voting-Rights Law Now at War with Itself? Social Science and Voting Rights in the 2000s*, 80 N.C. L. REV. 1517, 1518 (2002).

154. See *id.* (writing that the VRA “addresses the aggregate behavior, not of individual voters or state actors acting in isolation, but of groups of white and black (and other minority) voters whose combined patterns of voting behavior are the law’s focal point”). Viewing vote dilution through a group lens like this is noteworthy, as in non-VRA vote dilution cases the right is usually viewed by courts as being “individual and personal in nature.” See, e.g., *Gill v. Whitford*, 138 S. Ct. 1916, 1935 (2018) (Kagan, J., concurring).

155. Ellsworth, *supra* note 97, at 897.

156. See Carl E. Fisher et al., *Toward a Jurisprudence of Psychiatric Evidence: Examining the Challenges of Reasoning from Group Data in Psychiatry to Individual Decisions in the Law*, 69 U. MIAMI L. REV. 685, 685 (2015) (discussing how typical best practices in clinical psychiatry settings “do not necessarily comport with the ways in which aggregate data is applied to an individual case in the courtroom”). Similarly, in Constance R. Lindman, *Sources of Judicial Distrust of Social Science Evidence: A Comparison of Social Science and Jurisprudence*, 64 IND. L.J. 755, 759 (1989), the author distinguishes between the law’s favored “idiographic knowledge,” defined as knowledge focused “on the individual case and the facts idiosyncratic to that case,” with social science’s “nomothetic information” which “is concerned with general principles, relationships and patterns which transcend the individual instance.” Difficulties emerge, then, because “relying on nomothetic information creates uncertainty in an idiographic inquiry[,] . . . [and] [i]diographic knowledge about the behavior of individuals cannot be reliably inferred from nomothetic knowledge about the behavior of whole classes of people.” *Id.*

157. Through the years, the VRA has generally received strong legislative support. Most recently, in 2006, Congress approved a twenty-five-year extension of the Voting Rights Act by a 390–33 margin in the House and a unanimous 98–0 vote in the Senate. In 1992, the House passed another extension 237–125 and the Senate 75–20; in 1982, 389–24 (House) and 85–8 (Senate); in 1975, 346–56 and 77–12; and in 1970, 237–132 and 64–12. See, e.g., KEVIN J. COLEMAN,

Consider Section 2 of the VRA, for instance, which mandates that citizens of minority groups be afforded the same “opportunity” as the majority body to “participate in the political process and elect representatives of their choice.”¹⁵⁸ While this language may seem uncomplicated on its face, the calculations required by the mandate are anything but. In light of, among other things, the secret ballot confidentiality of almost all U.S. elections, determining whether a district experiences something like “racially polarized voting”¹⁵⁹ because its citizens vote in “bloc” or politically cohesive¹⁶⁰ racial coalitions can require some fairly advanced statistical science, not dissimilar from *Gill*’s EG, such as a technique known as “ecological inference.”¹⁶¹ In the VRA context, courts regularly trust and rely on these social science metrics to overcome the fact that all individual voters cannot be polled about their ballots, which would otherwise defeat their ability to prove they have been “harmed.” Given this, without the assistance of soft science, VRA claims would be entirely theoretical.¹⁶²

In *Gill*—same as in *Wal-Mart*—the plaintiffs’ cause of action required social science research for a reason virtually identical to the text of Rule 23: The injured group in question was so numerous that “joinder” of all members would be impracticable. Given that such claims may realistically only acquire meaning through an interaction with social science, denying plaintiffs use of this proof removes not only the sole means by which they may prove the claims, but in some contexts, the only means by which they may fully or meaningfully realize

CONGRESSIONAL RES. SERV., THE VOTING RIGHTS ACT OF 1965: BACKGROUND AND OVERVIEW 1, 22 (2015), <https://fas.org/sgp/crs/misc/R43626.pdf> [<https://perma.cc/F7SS-LMLN>]; J. Morgan Kousser, *The Strange, Ironic Career of Section 5 of the Voting Rights Act, 1965–2007*, 86 TEX. L. REV. 667, 686–87, 706, 761 (2008).

158. 52 U.S.C. § 10301(b) (2018).

159. See *Thornburg v. Gingles*, 478 U.S. 30, 30, 50–53, 53 n.20 (1986) (laying out the eponymous *Gingles* factors which are required to bring a Section 2 claim, and sanctioning the use of social science methods known as ecological regression and bivariate analysis as means of proving racially polarized voting).

160. *Id.*

161. See generally D. James Greiner, *Ecological Inference in Voting Rights Act Disputes: Where Are We Now, and Where Do We Want to Be?*, 47 JURIMETRICS 115, 123 (2007) (describing a variety of modern ecological inference methods used to “determine whether voting in a jurisdiction is racially polarized”).

162. In fact, such claims are at least somewhat frequently successful. In an article published in the aftermath of the *Shelby County* decision, see discussion *supra* notes 2, 10, Stephanopoulos estimated there had been at least “18 wins since 1982” in Section 2 cases. Nicholas Stephanopoulos, *The Future of the Voting Rights Act*, SLATE (Oct. 23, 2013, 4:37 PM), <http://www.slate.com/news-and-politics/2013/10/section-2-of-the-voting-rights-act-is-more-effective-than-expected-new-research-shows.html> [<https://perma.cc/B8TT-K5Y5>] (estimating this represents “a success rate of about 50 percent”).

Constitutional rights.¹⁶³ This notion of gaining content through a collaboration with social science is critical, because while Pildes may be correct in saying no other field displays *so* extreme a degree of mutual dependence between law and social science research as voting rights does, the general pattern reoccurs in countless substantive contexts. Unlike with the VRA, however, in these other fields, the Justices refuse to grit teeth and bear less than perfection.

This is the result of what Rachel Moran refers to as the judiciary's preference for the "common sense approach" of a "private model" of the law, "governed by relatively manageable concepts like intent and fault, rather than by inquiries into complex, cumulative harms."¹⁶⁴ This approach privileges analyzing harm at the individual—rather than the group—level, an approach jibing both with Justice Scalia's own personal judicial philosophy,¹⁶⁵ clear in *Wal-Mart*, and the broader doctrine of Article III standing, which ultimately sank the *Gill* plaintiffs' claim.¹⁶⁶

In *Gill*, the Court denied the plaintiffs standing for their inability to demonstrate an "invasion of a legally protected interest" that is "concrete and particularized" . . . [and] "affect[s] [them] in a personal and individual way"—that is, their failure to show the statewide districting plans caused *them specifically* to suffer an "injury-in-fact," one of three Constitutionally-required standing requirements.¹⁶⁷ In the voting rights setting, it seems unfortunate to restrict (non-VRA) litigants access to federal court review by construing the injury-in-fact requirement this way. After all, voting "harms" caused by gerrymandering are demonstrable *only* in aggregate, given that it is impossible to determine whether a single vote was gerrymandered without examining *all* of the votes at once, and critically, in relation to one another. "The concept of 'representation' necessarily applies to groups," Justice Lewis Powell wrote in the partisan gerrymandering case

163. See Brandon L. Garrett, *Aggregation and Constitutional Rights*, 88 NOTRE DAME L. REV. 593, 641–42 (2012) ("Assuming that a right is defined in a way that it applies outside particular individualized circumstances, aggregate litigation may provide a better forum to develop the meaning of that right.").

164. Moran, *supra* note 19, at 526.

165. "The province of the Court is, solely, to decide on the rights of *individuals* . . ." *Marbury v. Madison*, 5 U.S. 137, 170 (1803) (emphasis added). Throughout his career, Justice Scalia was quite fond of this particular line, quoting it often, including in his opinion for the plurality in the seminal standing decision in *Lujan*, and multiple times in an article published three years before joining the Court. See *Lujan v. Defenders of Wildlife*, 504 U.S. 555, 576 (1992); Antonin Scalia, *The Doctrine of Standing as an Essential Element of the Separation of Powers*, 17 SUFFOLK U. L. REV. 881, 883 (1983).

166. *Gill v. Whitford*, 138 S. Ct. 1916, 1933–34 (2018).

167. *Id.* at 1929 (quoting *Lujan*, 504 U.S. at 560).

Davis v. Bandemer, given that “groups of voters elect representatives, [and] individual voters do not.”¹⁶⁸

Simply put, gerrymandering is a *group* harm inflicted on all voters in a jurisdiction; metrics like EG are designed to capture and illustrate the occurrence of this harm. To find, as the Court did in *Gill*, that the plaintiffs lacked standing because EG does “not address the effect that a gerrymander has on the votes of *particular* citizens,”¹⁶⁹ then, is to focus on matters beside the point, and to ignore how capably very similar methods prove very similar harms in actions brought pursuant to Section 2 of the VRA.

This tension between the legal standard applied and the harm suffered is, perhaps surprisingly, quite in keeping with the Court’s regular practice. Scholars like Brandon L. Garrett have written that the Court over time has interpreted many “sub-constitutional” rights as requiring individualized inquiries or particular causation, an approach which frequently sinks claims seeking remedies for group harms.¹⁷⁰ In many ways, *Wal-Mart* owes its outcome to the Court’s preference for individualization and for the “common sense” approach embodied in its (and Scalia’s own preferred) standing jurisprudence. Though *Wal-Mart* is not a standing case, the *Dukes* plaintiffs lost for their inability to show discrete instances of individually-suffered harms, which seems a nearly identical point. At root, in *Wal-Mart*, Scalia craved a simple “yes” or “no” answer from Bielby, confidently speaking as granularly as down to each individual hiring decision. Did *these* individual store managers injure *these* specific employees? Are we sure? And if we are sure, can we say for certain that the injury was caused by *discrimination based on sex*, and not something else? Scalia writes:

Relying on ‘social framework’ analysis, Bielby . . . could not, however, determine with any specificity how regularly stereotypes play a meaningful role in employment decisions at Wal-Mart. At his deposition . . . Dr. Bielby conceded that he could not calculate whether

168. *Davis v. Bandemer*, 478 U.S. 109, 167 (1986) (Powell, J., concurring in part and dissenting in part). In *Bandemer*, the Court found that partisan gerrymandering claims are justiciable under the Equal Protection Clause, but was not able to agree on and adopt a “judicially manageable” standard for resolving the claims, the issue which continues still to this day. See also *Common Cause v. Rucho*, 279 F. Supp. 3d 587 (M.D.N.C. 2018) (holding that partisan gerrymandering claims present a nonjusticiable political question because the Court has yet to agree upon a judicially discoverable and manageable standard for adjudication).

169. *Gill*, 138 S. Ct. at 1933 (emphasis added).

170. See Garrett, *supra* note 163, at 632–33 (discussing *City of Los Angeles v. Lyons*, 461 U.S. 95 (1983), in which the Court held that plaintiff Lyons lacked standing to pursue injunctive relief against the LAPD for its use of a dangerous chokehold technique since he could not guarantee that he specifically would be subjected to the practice again in the future). Garrett writes that the decision has been “much criticized for its interpretation of Article III standing.” *Id.*

0.5 percent or 95 percent of the employment decisions at Wal-Mart might be determined by stereotyped thinking If Bielby admittedly has no answer to that question, we can safely disregard what he has to say. It is worlds away from significant proof that Wal-Mart operated under a general policy of discrimination.¹⁷¹

Saying Bielby “conceded” this distorts the issue, as calculating whether 0.5 or 95 percent of decisions were impacted by stereotyped thinking was not his goal. Instead, Bielby’s report sought to identify organizational structures which, among other things, explained the striking contents of fellow plaintiff expert Dr. Drogin’s statistical proof.¹⁷² Scalia’s rejection of Bielby’s testimony for the reasons stated seem as unimportant as Chief Justice Roberts’ reasons for denying standing in *Gill*: After all, why should the Court focus on a metric’s lack of certainty as to individual employee hiring decisions when the *Dukes* plaintiffs’ theory of liability—systemic disparate treatment under Title VII—seeks “to paint a complete picture of the structural nature” of discrimination within an organization?¹⁷³

Indeed, though a generalized approach’s nonspecificity is, in the Scalia view, a *shortcoming* of social science, in the class action setting it just as easily could have been recognized and celebrated as a profound *virtue*—much like it is in VRA Section 2 litigation. Consider, for instance, Judith Resnick’s use of the famed “lemon” problem, on loan from the field of economics, as a way of understanding the *Dukes* plaintiffs’ proof. “The argument,” she writes, “is that, for certain claims, outcomes across a comprehensive set can identify, with greater accuracy, the meritorious claims than can individual decisions, which produce noise and distortions obscuring obvious patterns.”¹⁷⁴ In a Title VII systemic disparate treatment case, any method by which judges may see this forest for the trees should be highly desirable, not a reason for heightened skepticism. Justice Scalia holding Bielby accountable for what he incorrectly identified as being the weaknesses of his method, then, is to punish a trumpet for its failure to be a piano, and to deny

171. Wal-Mart Stores, Inc. v. Dukes, 564 U.S. 338, 354–55 (2011) (internal quotation marks omitted).

172. Scalia mentions Dr. Drogin’s proof only perfunctorily in the actual opinion; for a simple comparison, Scalia’s opinion mentions Bielby by name thirteen times, while Drogin’s appears only three times, all in one short paragraph. *Id.* at 356. See also text accompanying *supra* note 57 for additional discussion regarding Dr. Drogin’s proof.

173. Ramona L. Paetzold & W. Steven Rholes, Wal-Mart v. Dukes: *Justice Scalia and Systemic Disparate Treatment Theory*, 21 EMP. RTS. & EMP. POL’Y J. 115, 117 (2017).

174. Judith Resnik, *Fairness in Numbers: A Comment on AT&T v. Concepcion, Wal-Mart v. Dukes, and Turner v. Rogers*, 125 HARV. L. REV. 78, 146 (2011) (discussing George A. Akerlof, *The Market for “Lemons”: Quality Uncertainty and the Market Mechanism*, 84 Q.J. ECON. 488 (1970)). Resnick describes the problem as related to “the analytic distortions produced by owning one product that was a ‘lemon’ and by the asymmetries of ‘marketplace information.’” *Id.* at 146.

Bielby's expertise its specific value by first mischaracterizing and then minimizing it. This approach, scholars like Paetzold and Rholes have argued, can only be explained by either ignorance of—or outright malice toward—the method, or toward social science. As they put it in a forcefully-argued 2017 article, Scalia's opinion:

[E]vinced . . . contempt for a model of discrimination that attempts to focus beyond the individual level of harm. This hostility appears to exist at a theoretical level, but implicit in this is hostility toward social science, statistics, and methods of proof that have long been found to be acceptable by the Supreme Court.¹⁷⁵

Overall, the Court's characterizations of both Bielby's report and Stephanopoulos and McGhee's EG is both (1) consistent with its typical treatment of social science proof showing harms suffered by groups; and (2) a symptom of the judiciary's broader inability to evaluate the substance of expert testimony, which results in overreliance on the "familiar" and "traditional." In cases like *Wal-Mart* and *Gill*—for which the harms alleged *require* analysis tailored to the group—this regime results in the undeserved exclusion or discounting of relevant testimony, all to preserve the Court's treasured archetypal "perfect proof." Social science can competently provide a means by which the judiciary could perform this group-level analysis—if only the Court would let it.

III. TRANSLATING BETWEEN THE LANGUAGES OF SCIENCE AND LAW

An additional layer of difficulty perpetuating the deep-rooted judicial devaluing of social science relates to problems in translating between the contrasting languages of law and science. Consider again, briefly, Scalia's rejection of Bielby's research for a simple example of the translation problem in action. Despite Scalia's repeated uses of the term, Bielby was not, in fact, seeking to perform a "social framework analysis." Social framework analysis is "not a sociological method, but rather a legal term for some kinds of research."¹⁷⁶ In this respect, it is much like the court-created construct of "insanity," a diagnosis

175. Paetzold & Rholes, *supra* note 173, at 150–51 ("So where did [Scalia] go wrong? Everywhere.").

176. ASA Amici, *supra* note 60, at 3 ("Social framework analysis' is not a sociological method, but rather a legal term for some kinds of research. What constitutes high quality 'social framework analysis' continues to be vigorously debated among scholars. As such, the Court should assess the underlying social science methods, as practiced by social science researchers and vetted in the peer-reviewed journals of those fields, instead of the 'social framework analysis' construct when deciding whether social scientific work is valid."). The concept of social framework analysis as a legal construct made its debut in Laurens Walker & John Monahan, *Social Frameworks: A New Use of Social Science in Law*, 73 VA. L. REV. 559 (1987).

which does not actually exist with any “one-to-one” correspondence in psychiatry’s regular nosology, but to which the law frequently asks that experts testify.¹⁷⁷ Though this could seem a minor point, it can have broader, deleterious effects on judges’ ultimate evaluations of methodology.

This occasions a closer look at psychiatry more generally, as it is a soft science field with presumably much to offer the law, but which frequently frustrates its highly specific needs because of its specialized “language” and disciplinary norms.¹⁷⁸ Similar to Bielby or Stephanopoulos and McGhee, who pull from far-flung group data to offer conclusions likely to apply to individual units, psychiatric research “inevitably” employs and extrapolates from aggregate data to offer generalized conclusions applicable “across persons, places, or things,” a challenge Fisher, Faigman, and Appelbaum term “the G2i problem,” or more simply, the difficulty with applying general data to individual cases.¹⁷⁹ Broadly, of course, all scientists gain insight into observed phenomena by tallying, from sufficiently large samples, various data relating to a specific phenomenon’s occurrence, including the many causal, complicating, or other predictive “manifestations” which accompany its presence or incidence.¹⁸⁰ In medical research, for instance, when a particular phenomenon, or syndrome, is well understood generally in the sense that professionals consistently associate specific manifestations, or symptoms, *with* the syndrome, then practitioners may begin diagnosing individual patients with the syndrome when they too possess those symptoms.¹⁸¹ Essentially, the act

177. Fisher et al., *supra* note 156, at 695 (“Insanity is defined in many jurisdictions as a cognitive impairment at the time of the offense that rendered the defendant unaware of the wrongfulness of his conduct. But in psychiatry there is no diagnostic category that corresponds fully to this definition of insanity.”).

178. As psychiatry of course makes up part of the “M” in STEM, psychiatrists certainly have a credible claim to being “hard” scientists. However, due to the twin difficulties of psychiatric study and treatment—namely, that we cannot simply autopsy a living subject to determine the innerworkings of his mental illness, and even if we could, we would not know nearly enough what to look for—“[t]he perception of many, including health care professionals, is that psychiatry is a ‘soft’ science and treatment of mental illness is not focused or specific and certainly not effective.” John S. McIntyre, *Usefulness and Limitations of Treatment Guidelines in Psychiatry*, 1 *WORLD PSYCHIATRY* 186, 187 (2002). Because of this, and the fact that psychiatry shares countless methodological and practical traits with psychology—a virtually undisputed soft science—in this Comment, I treat psychiatry as a soft science.

179. See generally Fisher et al., *supra* note 156, at 685.

180. *Id.* at 687, 691 (giving the example of meteorologists, who will “study weather systems to identify variables that will permit accurate forecasting of rain on a particular day”).

181. Further, diagnosis is not simply a descriptive process useful only for categorization, but also one offering guidance for future action. Fisher et al. describe diagnoses as “anchor[ing] the clinical judgment of the expert, helping to restrain unfounded speculation that transcends the generally accepted characteristics of the disorder in question.” *Id.* at 691. They go on to write, “[a] diagnostic and functional assessment is essentially an inference that connects an individual

employs the same processes underpinning analogical legal reasoning, which “involves the comparison of closely related cases to the current case and the placement of the current case in the category of those it most closely resembles, or, occasionally, the creation of a new distinction that differentiates this case from all previous cases.”¹⁸²

With certain syndromes, diagnosis may require identifying or locating just a single characteristic always overlapping with the syndrome: For instance, if a doctor suspects a patient with swollen lymph nodes, sore throat, and fever has strep throat, the doctor will test for the presence of a specific biologic characteristic, or “biomarker”—the existence of group-A β -hemolytic streptococcus bacteria in a throat culture. In a more complicated but still well-established diagnosis, such as multiple sclerosis (MS), the doctor will first determine whether there exists a suite of relevant symptomatology (from a laundry list of potential manifestations) so as to justify ordering specific tests for two or more biomarkers, including the presence of oligoclonal bands in the patient’s cerebrospinal fluid (detected through a spinal tap) and white matter lesions in the brain or cervical spine (by MRI).¹⁸³ Though not *all* people with cerebrospinal fluid oligoclonal bands have MS (they may have Lyme disease, or meningitis), and neither will all of those with brain lesions (lupus, neuromyelitis optica), when clinicians observe *both* biomarkers together and then rule out, by blood test, conditions causing similar symptoms, they will be fairly confident in making a diagnosis.¹⁸⁴

While this process may seem like its own tightrope walk—and, in fact, MS misdiagnosis is relatively common¹⁸⁵—it is dramatically more refined than many typical psychiatric diagnoses, such as bipolar disorder, posttraumatic stress

case to diagnostic labels that have been developed on the basis of group data and a collective clinical understanding.” *Id.* at 692–93.

182. Ellsworth, *supra* note 97, at 900.

183. See generally *Multiple Sclerosis*, MAYO CLINIC (Apr. 19, 2019), <https://www.mayoclinic.org/diseases-conditions/multiple-sclerosis/diagnosis-treatment/drc-20350274> [<https://perma.cc/M8Q7-Y5NP>].

184. *E.g., Other Conditions to Rule Out*, NAT’L MULTIPLE SCLEROSIS SOC’Y, <http://www.nationalmssociety.org/Symptoms-Diagnosis/Other-Conditions-to-Rule-Out> [<https://perma.cc/GVY6-VL6S>] (last visited Nov. 8, 2018) (listing more than twenty other conditions which should be rejected before settling on a diagnosis of MS); Neil Scolding, *The Differential Diagnosis of Multiple Sclerosis*, 71 J. NEUROLOGY, NEUROSURGERY & PSYCHIATRY ii9, ii9 (2001) (noting the range of unrelated disorders that mimic MS symptoms, as well as certain syndromes which frequently accompany the disease).

185. Andrew J. Solomon et al., *The Contemporary Spectrum of Multiple Sclerosis Misdiagnosis*, 87 NEUROLOGY 1393, 1393 (2016) (noting that “misdiagnosis of MS is frequently acknowledged” and “remains common”).

disorder, or anxiety.¹⁸⁶ Though we are very confident these conditions exist, just like we have no doubt that partisan gerrymandering, or workplace sex discrimination actually occur, we have a much more difficult time diagnosing the conditions in specific cases with certainty because *zero* biomarkers exist, and psychiatrists must instead engage in the inexact exercise of interpreting highly-variable symptoms, behaviors, and many other complicating factors which we are not even close to understanding fully.¹⁸⁷ Though the MS diagnosis also requires significant interpretive work and creativity in that doctors must pull from a number of disparate sources to confidently render diagnoses, a psychiatric diagnosis requires a great deal more. Such differences—in both the amount of interpretation required and resultant postconclusion “certainty”—are exactly why neurologists are generally labeled “hard” scientists, and psychiatrists “soft.” The processes are virtually identical—phenomena are identified generally, and then experts must decide whether individual cases fit the definition based on all available evidence—but soft scientists simply have more factors to account for, or factors which they are only capable of accounting for in more limited or otherwise flawed ways.¹⁸⁸

This divide typifies and perhaps further explains the disciplines’ respective positions on the so-called “Hierarchy of the Sciences,” a rank order generally listing natural sciences like physics or chemistry at the “top” and social sciences like sociology or psychology at the bottom.¹⁸⁹ Generally, a field’s rank correlates to its

186. See, e.g., Gabrielle Moss, *5 Mental Health Disorders That Are Often Misdiagnosed*, BUSTLE (June 9, 2016), <http://www.bustle.com/articles/165839-5-mental-health-disorders-that-are-often-misdiagnosed> [<https://perma.cc/8N6D-L6ZJ>] (listing borderline personality disorder and ADHD as additional frequently misdiagnosed disorders).

187. While research is surging and profound leaps forward in understanding have taken place in the last half-century, the amount we actually know about the brain remains shockingly little. See, e.g., James Gorman, *Learning How Little We Know About the Brain*, N.Y. TIMES (Nov. 10, 2014), <http://www.nytimes.com/2014/11/11/science/learning-how-little-we-know-about-the-brain.html> [<https://perma.cc/W4C2-49M9>] (noting that “[s]cientists have puzzled out profoundly important insights about how the brain works, like the way the mammalian brain navigates and remembers places . . . [y]et the growing body of data . . . represents a paradox of progress, with the advances also highlighting great gaps in understanding”).

188. See Fisher et al., *supra* note 156, at 693 (“Psychiatry, of course, is itself an applied science. Hence, as with all scientific disciplines, it begins by identifying certain general phenomena and then, in application, determining whether a particular case is an instance of that phenomenon.”).

189. See, e.g., Stephen Cole, *The Hierarchy of the Sciences?*, 89 AM. J. SOC. 111, 111 (1983). The idea of such a hierarchy has existed for close to 200 years. As originally formulated by philosopher Auguste Comte in his pivotal text “The Course in Positive Philosophy,” a potential fully-developed field of sociology (which he termed “social physics”) was the paragon scientific achievement, and further down on the hierarchy was biology (organic physics), physics and chemistry (terrestrial physics), and astronomy (celestial physics). Though Comte ranked sociology first on his list and as the ultimate *goal* of all research, he also viewed it as the most

performance relative to others in a variety of conceptual metrics, with those performing the “best” ranking as most hard, and the worst, soft—thus making the Hierarchy something of a scientometric Mohs’ scale.¹⁹⁰ These measures include tests of concepts such as the amount of scholarly consensus in the discipline,¹⁹¹ the “replicability” of its published studies,¹⁹² the field’s relative level of predictive and explanatory success,¹⁹³ and its overall number of “positive results.”¹⁹⁴ On the

complex, least understood, and field most dependent on continued development of other sciences. See 1 AUGUSTE COMTE, *COURS DE PHILOSOPHIE POSITIVE* (1st ed. 1830).

190. Referring to the mineralogy “hardness test” named for geologist Friedrich Mohs, which ranks diamonds, one of the very hardest minerals, as a “10” and talc, one of the softest, a “1.” *Mohs Hardness*, ENCYC. BRITANNICA (Apr. 29, 2017), <https://www.britannica.com/science/Mohs-hardness> [<https://perma.cc/KP4T-YTLV>].

191. Some scholars believe hard sciences generally have a higher level of scholarly consensus as to the “significance of new knowledge and the continuing relevance of old.” See Daniele Fanelli & Wolfgang Glänzel, *Bibliometric Evidence for a Hierarchy of the Sciences*, 8 PLOS ONE 1, 1 (2013). In the “ideal” science, which hard sciences are believed to more closely approximate, “scholars share a common background of established theories, facts and methods”; but “[m]oving towards ‘softer’ fields, this consensus becomes less likely to be reached, the common background shrinks and fractures, and . . . data become less able to ‘speak for themselves.’” *Id.*

192. Another troublesome metric for soft science is its experts’ comparatively lesser ability to recreate published studies, which some in the early 2010s termed a “Replicability Crisis.” See generally Harold Pashler & Eric-Jan Wagenmakers, *Editors’ Introduction to the Special Section on Replicability in Psychological Science: A Crisis of Confidence?*, 7 PERSP. ON PSYCH. SCI. 528 (2012). This is problematic, obviously, because if an original study’s results cannot be duplicated by peers—or, occasionally, even the original expert him or herself—then there is a good chance its conclusion is in at least some way unreliable.

193. See Paul Humphreys, *A Conjecture Concerning the Ranking of the Sciences*, 9 TOPOI 157, 157 (1990). Humphreys theorizes that hard sciences generally perform better on measures of what he terms “predictive success” and “explanatory success.” Regarding predictive success, he writes, “Few, even those well-disposed towards the social sciences, would dispute that chemistry is predictively more successful than, say, macroeconomics,” indicating that experts have a significantly higher mastery over the former than the latter, and we can thus more consistently count on their findings, and less on the macroeconomist’s. *Id.* Similarly, Humphreys offers that “the essential feature of [explanatory success] is the requirement that adequate explanations list contributing and counteracting causes, where to be a cause is to produce the effect invariantly.” *Id.* at 159.

As in Fanelli & Glänzel’s study, see Fanelli & Glänzel, *supra* note 191, at 1, where those fields performing studies on “non-biological material” tend to rank higher on the Hierarchy than do those relying on inexact measurements of various aspects of human behavior and cognition, Humphreys views the social sciences as less able to account for all contributing and counteracting causes leading to a result’s guaranteed occurrence. Humphreys, *supra*, at 157; see also *infra* note 194.

194. Some studies have found that a field’s measure of “positive results,” determined by the amount of academic papers in the discipline reporting confirmation (rather than invalidation) of the paper’s tested hypothesis, correlates with its rank on the Hierarchy. See, e.g., Daniele Fanelli, *“Positive” Results Increase Down the Hierarchy of Sciences*, 5 PLOS ONE 1, 3 (2010). Certain social sciences have a high prevalence of positive results: In Fanelli’s study, psychology and psychiatry had the highest prevalence of positive results (91.5 percent), while “Space Science” the lowest (70.2 percent), which might raise fears that the methodologies employed in the former were faultily constructed—perhaps in such a way that the positive result was inevitable,

whole, the metrics privilege the empirical over what might seem the subjective, and the quantitative for the qualitative. In sum, what they all describe is some version of the proposition that *soft sciences are less likely to reach certain conclusions than hard sciences are*. For judges, this fact is no doubt a strong indicator of unreliability notwithstanding what may be an individual piece of social science's true merit. In an especially cruel application of the G2i problem, judges will reason from available group data—that is, their awareness that soft sciences, disciplinarily speaking, less frequently provide certain conclusions than do hard sciences—to make predictions about individual pieces of evidence. All things being equal, this will heavily prejudice judges against soft science.

Disciplines speaking simply to cold hard numbers, on the other hand, seem to simply report incontrovertible fact, which must necessarily make them more “reliable.” Their results simply tell us what was already the case, and our job is just to uncover that truth by applying the correct gauntlet of formulas. Hard sciences, though themselves very complicated, produce results which are anything but. Binary. Y/N. Biomarkers, essentially. Soft sciences, on the other hand, are stuck in an unending state of “maybe,” or “yes, but only in the event that—” or “it depends.” They do not sound like our archetypal expert. The hard science approach, in fact, sounds a great deal closer to Justice Scalia's favored interpretive tool, legal formalism, which:

[S]eems to start from the assumption that the law is a closed, logical system. Judges do not make law: they merely declare the law which, in some Platonic sense, already exists. The judicial function has nothing to do with the adaptation of rules of law to changing conditions; it is restricted to the discovery of what the true rules of law are and indeed always have been. Past error can be exposed and in that way minor corrections can be made, but the truth, once arrived at, is immutable and eternal.¹⁹⁵

Formalism, Moran writes, “maximizes the tensions between law and science as ways of knowing,” in that formalists “look to their judicial predecessors, not contemporary social scientists, to determine what the law should look like. . . . Law is immutable; science is tentative; law is certain; science is speculative.”¹⁹⁶ Is the

even though in an ideally-constructed study it might have been negative—and their conclusions therefore meaningless. *Id.*

Overall, Fanelli's study finds that positive results are 2.3 times more frequent in social sciences than in physical sciences, and 3.4 times higher in “studies applying behavioural and social methodologies on people” compared to “physical and chemical studies on non-biological material.” *Id.* at 1.

195. GRANT GILMORE, *THE AGES OF AMERICAN LAW* 62 (1977).

196. Moran, *supra* note 19, at 537.

true rivalry, then, not a battle between philosopher and physicist, but between formalist and other, more flexible modes of legal interpretation? This might be so, but in some respects, it may not be particularly illuminating to separate chicken from egg. The end result is identical: That, in the courtroom, outside knowledge able to more credibly vouch for its own “certainty,” as hard science is, coexists more comfortably with a judicial philosophy rooted in natural, rather than positive, law. Social science, which frequently challenges the soundness of long-existing frameworks, is by its nature in conflict with a jurisprudence hinging on the system being unchanging and permanent.

And though social science’s perpetual lack of certainty is obviously anathema to the law’s aspirations, it is strikingly similar to its operation.¹⁹⁷ Much like the law, social science analyzes the behavior of highly unpredictable human actors by synthesizing diverse and constantly-evolving bodies of knowledge and research. Both seek to make inherently abstract concepts concrete in order to explain and possibly remedy them. And at bottom, the endgame for each is to offer intelligent guesses to problems whose full solutions are only ascertainable through omniscience. In deciding between murder and manslaughter, for instance, judges and juries might parse from contextual clues whether a defendant possessed the “depraved and malignant heart” of malice aforethought, even though seeking full certainty on the point would be an absurd exercise. In the criminal trial, the court makes do with imperfection; when a testifying social scientist falls short of the impossible standard, she may be written off. In one setting, uncertainty is fatal; in the other, it may well undergird the decision to execute.

Broadly, a massive loss for plaintiffs which regularly is a direct byproduct of the formalist approach is the inability to point to *results* capturing obvious structural harms as proof. Of course, a results-only (or at least more results-sensitive) regime is not simply a fantastical wish with no place at all in the law. Consider again the VRA’s Section 2, which explicitly is such a test. The Section forbids “any voting standard, practice, or procedure that *results in* the denial or abridgement of the right of any citizen to vote on account of race, color, or

197. “[A]nd if the law is a science at all, surely it must be a social science.” Ellsworth, *supra* note 97, at 897. Ellsworth notes that judges, like social scientists, will be “faced with an individual case, a case that is unique but that resembles various other cases that she and others have seen and sometimes described,” and will then apply “the case method” to chart the proper course. *Id.* at 898; see also Lindman, *supra* note 156, at 758 (“If social scientists are scrupulous about revealing inadequacies and uncertainties, it may seem ironic that jurists, who specialize in dealing with evidence containing varying degrees of probativeness, view social science evidence so skeptically.” (internal quotation marks omitted)).

membership in a language minority group.”¹⁹⁸ If those results show that a voting system has impacted a minority group’s ability to vote, the plaintiffs are entitled to remedy regardless of any finding of discriminatory intent. Outside of this and other narrow exceptions, however, the law overwhelmingly privileges intent over impact alone, an approach which has only increased under the Rehnquist and Roberts Courts.¹⁹⁹ As a result, in the usual case, social science findings about aggregate impact are rendered irrelevant since they are incapable of speaking to individual intent. This again shields the Court from broad consensus and keeps it more firmly trapped in the trees rather than the forest.

Such a tendency towards formalism leads to a privileging of traditional intent, “grounded in a belief in personal responsibility and free will,” over the deterministic, results-oriented assumptions of social science, which “impl[y] that other people, perhaps most people, would have behaved the same way in the same circumstances, [and] suggesting that the circumstances largely determined the behavior.”²⁰⁰ By and large, the law depends on judges being able to locate individual fault occurring as a result of deliberate choice, and the existence of “free will” itself may be “necessary to carrying out the judicial function.”²⁰¹ Social science, by introducing into the equation the uncomfortable fact that a group’s behavior may have been preordained or in some other way restricted by its circumstances, calls into question the soundness of key foundational aspects of the whole legal system.

The Court’s unease with the deterministic view is clear in *Wal-Mart*, where Bielby’s research into *unconscious* corporate biasing was premised on the presupposition that managers operating under a particular set of policies might be unable to “choose” right (not stereotyping) from wrong (relying on stereotypes) because their circumstances (working in a company with certain policies in place), in broad strokes, dictate their behavior. Regardless, Scalia treated the research’s failure to demonstrate individual, traditional intent fault within Drogin’s “uncontested descriptive statistics” as dispositive.²⁰² The statistics were not acceptable because they were “just” results, and the interpretive framework the

198. *Section 2 of the Voting Rights Act*, U.S. DEP’T JUST. (Sept. 14, 2018), <https://www.justice.gov/crt/section-2-voting-rights-act> [https://perma.cc/RP2S-8ES9] (emphasis added).

199. Ellsworth, *supra* note 97, at 916 (discussing Moran, *supra* note 19, at 541–43, and noting that “Moran argues that this emphasis . . . has actually increased under the Rehnquist and Roberts Courts”).

200. *Id.* (emphasis omitted).

201. *Id.* at 918.

202. These statistics indicated, for instance, that “[w]omen fill 70 percent of the hourly jobs in the retailer’s stores but make up only 33 percent of management employees.” *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 370 (2011) (Ginsburg, J., dissenting) (internal quotations omitted).

Dukes plaintiffs used to explain the results frustrated Scalia's favored conception of the law.

Despite the legal system's affinity for this binary certainty of "either/or" categorical thinking—that "a defendant is either liable or not liable, sane or insane, a danger to society or not"—occasionally, some valuable information might lie between 0 and 1.²⁰³ To use our previous example, if an expert must testify to whether or not a defendant was insane at the time he committed a crime, perhaps the psychiatrist's own professional "language" would characterize his mental state more descriptively than the "yes" or "no" the law's expediciencies demand. In spite of what might be an incredibly complex symptomatology, there is little room for a nuanced answer when the question asked is simply: "Should the defendant be incarcerated, or not?"

Returning once more to the Court's line of death penalty cases, Chief Justice Rehnquist's opinion in *Lockhart v. McCree*²⁰⁴ is a master class in bad faith manipulation of social science. By 1986, scholarly consensus held—as it continues to—that if all individuals making up a jury favor the death penalty generally, then those individuals are highly biased toward finding a defendant guilty in a capital case.²⁰⁵ Despite this, as Ellsworth notes, Rehnquist took a series of fifteen empirical studies—conducted over a period of thirty years, across geographic locations and demographic groups, and employing a variety of methods—and picked them apart "one by one, finding a flaw in each and discarding it from the set, until only one was left."²⁰⁶ "Surely a per se constitutional rule as far-reaching as the one McCree proposes," Rehnquist then stated, "should not be based on the results of [a] lone study."²⁰⁷

Under this "all or nothing" regime, a study being inadmissible means that it cannot influence decision and contribute to a larger pattern, even as an obvious mosaic emerges from the collage. Since the notion of "large consensus on almost all, but not every, key point" is not a legally cognizable one, when no individual study is perfectly satisfactory, the judiciary has no way of perceiving the widespread, even overwhelming, agreement which may exist. And perhaps mirroring the very content in the studies now being discounted—that people's preexisting belief about the death penalty prejudices their judgments about

203. Ellsworth, *supra* note 97, at 913. "Scientists, especially social scientists," Ellsworth writes, "are more likely to think in terms of continuous variables; there is always a grey area between the sane and the insane, the dangerous and the safe, and the deliberate or unintentional behavior." *Id.*

204. 476 U.S. 162 (1986).

205. See, e.g., Ellsworth, *supra* note 128, at 189–204.

206. Ellsworth, *supra* note 97, at 914.

207. *Lockhart*, 476 U.S. at 172–73.

specific applications of it—Justice Rehnquist was able to cast the fifteen studies’ chorus of consensus as being a matter still genuinely in debate. Regarding a prior death penalty case with a remarkably similar outcome, *Gregg v. Georgia*,²⁰⁸ Ellsworth writes that “[b]ehind this is a tacit requirement of complete consensus within the social science community before its research on an issue even need be *considered*, a standard far beyond what is required for any other sort of evidence.”²⁰⁹ It is “distressing,” she writes, to think that the most influential thing a social scientist can do is give the Court an excuse to ignore a body of research.”²¹⁰

If this at all sounds familiar, it might be because it echoes a point Justice Sotomayor made at the *Gill* oral arguments: that “every single social science metric points in the same direction.”²¹¹ Moreover, as Justice Kagan pointed out, the various metrics’ conclusions were by no means in flux: “[W]hat I’m suggesting is that this is not kind of hypothetical, airy-fairy, we guess, and then we guess again. I mean, this is pretty scientific by this point.”²¹² Nonetheless, Chief Justice Roberts was able to discount it all and call broad certainty into question by casting the method employed to reach it in the case at bar as “sociological gobbledygook.” Time and again, this pattern recurs and, over and over, the approach leads to counterintuitive, even illogical decisions that seem blind to a reality clear in the data.

And as for an expert who, like Dr. Bielby, is asked to testify to a court-manufactured legal framework rather than using the established methodologies of their discipline, there exists an additional, perhaps even more vexing problem: the Court’s ability to unilaterally “alter” the norms of that discipline either by genuinely not understanding it or by strategically misconstruing it to justify a holding desirable to that Justice. In moving the goalposts by demanding, for instance, between 0.5 and 95 percent certainty, the Court not only rejects one specific piece of social science research, but essentially the field of sociology itself. Paetzold and Rholes contend that this error could only have occurred as a result of ignorance of or hostility toward social sciences,²¹³ and in fact, it certainly is difficult to imagine the Justices so flagrantly resetting the norms of a hard science discipline. As *Myriad* makes clear, Scalia would rather recuse himself from decision than engage with the specialized knowledge of molecular biology, and so it is difficult to imagine him trying to “get away” with not studying for a test on the subject if a decision truly hinged on it.

208. 428 U.S. 153 (1976).

209. See Ellsworth, *supra* note 128, at 182.

210. *Id.*

211. Transcript of Oral Argument, *supra* note 8, at 16.

212. *Id.* at 40.

213. See *supra* note 175 and accompanying text.

Perhaps the reason why is simple: Society at large allows him to get away with either misunderstanding or misrepresenting social science because of entrenched biases toward the disciplines. “Half of Americans,” researchers have found, think young people don’t pursue STEM because it is simply “too hard.”²¹⁴ Behind this proposition, of course, hides the nasty corollary that other subjects, including the soft sciences, are in fact “easy,” and if they are, then why must we bother deferring to their experts at all? It is through precisely this cultural posture that potentially relevant and valuable social science so often enters judicial proceedings dead on arrival, left to be misconstrued, distrusted, or ignored. But what harm is there really, after all, when all that’s lost is a bit of sociological gobbledygook?

CONCLUSION: CONSTRUCTING “CERTAINTY” THROUGH TRIAL

An above-raised question still unanswered is how, exactly, would we prefer our judges and Justices to engage with disputed subject matter falling outside their epistemic competence? Do we truly want to enroll them all in their own “crash courses” in organic chemistry, molecular biology, sociology, or psychology so that they may cram valiantly and become as close to Junior Experts as possible before rendering decisions? Certainly not, or at least not always. It seems unlikely that judges could ever sufficiently bring themselves up to speed on any and all issues, given crushing docket demands and the incredibly complex nature of certain proofs. But neither should we throw up our hands and end there. Even where the disciplinary norms and languages of social science and law clash, or where crucial facts of a dispute seem impossible to ascertain with certainty, perhaps the most foundational aspect of the legal system provides a powerful solution: going to trial.

In his book *Speak Now: Marriage Equality on Trial*, Kenji Yoshino details the historic twelve-day trial in *Hollingsworth v. Perry*, conducted in the Northern District of California in 2010.²¹⁵ In *Perry*, the plaintiffs successfully challenged the constitutionality of California’s Proposition 8, an amendment to the California state constitution defining marriage as between one man and one woman, passed by voters in a hotly-contested 2008 ballot initiative. At the time of Prop 8’s enactment, there was perhaps no topic more perverted by cultural pre- and misconceptions, or preexisting cultural agendas, than gay marriage. The issue has

214. Brian Kennedy et al., *Half of Americans Think Young People Don’t Pursue STEM Because It Is Too Hard*, PEW RES. CTR. (Jan. 17, 2018), <http://www.pewresearch.org/fact-tank/2018/01/17/half-of-americans-think-young-people-dont-pursue-stem-because-it-is-too-hard> [https://perma.cc/8MKG-K9US].

215. See generally YOSHINO, *supra* note 34; see also *Perry v. Schwarzenegger*, 704 F. Supp. 2d 921 (N.D. Cal. 2010).

incited strong feelings along political, religious, and gender lines since time immemorial, and a variety of unsupported—sometimes outlandish—talking points slowly had wormed their way into the national discourse and there remained unchallengeable, legally, for literal decades. Prior to trial, the Prop 8 proponents had argued that these “disputed sociological issues could not be resolved through the judicial process”²¹⁶ and urged Judge Vaughn Walker that the only way to resolve questions of so-called “legislative” facts was to simply “go back to [his] chambers . . . and pull down any dictionary, pull down any book that discusses marriage”²¹⁷

The proponents’ wariness toward trial as a way of resolving legislative/social facts, Yoshino argues, certainly would not have “come into play in the arena of the natural sciences, where people feel that they can reach some objective answer untainted by ideology.”²¹⁸ This should sound quite familiar. And given that marriage equality turns in large part *not* on the empiric safety of the hard sciences, but on the best guesses of the more dangerous soft, directing Judge Walker to “books,” to contemplate in the quiet of his own chambers, might seem the closest approximation of untainted objectivity possible under the system’s constraints. What it actually sounds like, though, is our misguided impulse to enroll judges in a “crash course” in marriage equality, so they may strive for an epistemic competence surely unattainable under the compressed timeframe of a summary judgment motion, and in which their sources of information may more easily be restrained and manipulated by either side of a dispute.²¹⁹

Most crucially, however, is that a book cannot be cross-examined. As the trial in *Perry* showed, the power of confrontation creates a mechanism by which ignorance may speedily be replaced with knowledge,²²⁰ and in which all sides of even the most multifaceted issues can be presented, probed, and weighed, in a fashion falling quite familiarly within judges’ own expertise. In *Perry*, Judge Walker heard from highly-qualified scholars of American history, political science, Asian American social and political movements, social history, public policy, developmental psychology, economics, and government, among others. Along with these experts’ case-in-chief testimony, however, Walker also heard intense cross examination, which even the most seasoned experts in the group said was unmatched in intensity by anything since their graduate school oral

216. YOSHINO, *supra* note 34, at 9.

217. *Id.* at 75.

218. *Id.* at 76.

219. See Lindman, *supra* note 156, at 757 (“Because judges are exempt from scrupulous jury trial rules, their search for legislative facts is left unrestricted.” (citation omitted)).

220. YOSHINO, *supra* note 34, at 78–9.

examinations.²²¹ Incredibly, in this setting, it becomes easier—even simple—to spot genius from within the junk. “Claims that might seem unassailable in a thirty-second spot on television,” Yoshino writes, “could be dismantled through hours of methodological questioning in the dock.”²²² Unlike the Rehnquist approach in *Lockhart*, where small quibbles with individual pieces of evidence pulled them each out of consideration, unable to contribute to big-picture patterns of certainty, in a well-run trial, the meritorious aspects of particular disputes naturally coalesce into an undeniable picture.

“When Judge Walker called for a trial,” Yoshino writes, “he transformed the litigation by turning on a threshing machine. Factual allegations would not be taken at face value but instead shunted into a device that would separate wheat from chaff.”²²³ Bare untruths, such as the “canard[]” that homosexual abuse of children is higher than heterosexual abuse—disproven in the 1970s, but nonetheless hanging around the outer edges of the marriage equality debate for decades thereafter—simply fall apart under good adversarial questioning.²²⁴ “When they come into court and they have to support those opinions and they have to defend those opinions under oath and cross-examination,” plaintiffs’ attorney Jon Boies said after the trial, “those opinions just melt away.”²²⁵ As Yoshino concludes:

The reason the trial was lopsided and the outcome comprehensively in favor of the plaintiffs was not due to the limitations of the trial’s fact-finding procedures . . . or judicial bias. It was because the opponents finally had nowhere to turn, and their arguments were revealed for what they are: wholly unpersuasive And so I say again—for the next great legal controversy that turns on key legislative facts: *Let there be a trial.*²²⁶

Ultimately, it may be true that Justice Ginsburg, dissenting in *Wal-Mart*, was no different from Justice Scalia, and simply offered her own epistemically incompetent read of the *Dukes* plaintiffs’ evidence, motivated by her own faulty preexisting cultural agenda. If this is true, then it is perhaps unavoidable in any case. But in denying class certification, Scalia weaponized uncertainty as a justification for leaving the merits permanently unresolved, while Ginsburg instead suggested we certify the class and “let there be a trial.” Embracing trial more frequently will save judges and Justices from their worst tendencies, and keep

221. *Id.* at 112.

222. *Id.* at 158.

223. *Id.* at 155.

224. *Id.* at 162.

225. *Id.* at 8.

226. *Id.* at 280.

them from substituting their own skepticism for the knowledge potentially ascertainable through the normal operation of the adversarial process. Wigmore famously wrote that cross-examination is “beyond any doubt the greatest legal engine ever invented for the discovery of truth.”²²⁷ It is time to utilize this engine. By doing so, judges will be able to construct certain answers to the unresolved, but vital, questions pursued by social scientists.

227. 5 JOHN HENRY WIGMORE, EVIDENCE IN TRIALS AT COMMON LAW § 1367 at 32 (Chadbourn rev. 1974).