

Texas Law Review

PARTY RULEMAKING:
MAKING PROCEDURAL RULES THROUGH PARTY CHOICE
Robert G. Bone

INJUNCTIONS AS MORE (OR LESS) THAN “OFF SWITCHES”:
PATENT-INFRINGEMENT INJUNCTIONS’ SCOPE
John M. Golden

BOOK REVIEW—STILL CONVICTING THE INNOCENT
Jennifer E. Laurin

BOOK REVIEW—CONSTRUCTING A CANON OF LAW-RELATED POETRY
Alexandra J. Roberts

MONETARY DAMAGES AND THE (B)(2) CLASS ACTION:
A CLOSER LOOK AT *WAL-MART V. DUKES*

WOULD YOU WANT WILLIAM HUNG AS YOUR TRIER OF FACT?
THE CASE FOR A SPECIALIZED MUSICOLOGY TRIBUNAL

Texas Law Review

A national journal published seven times a year

Recent and Forthcoming Articles of Interest

visit www.texasrev.com for more on recent articles

TALK DERBY TO ME: INTELLECTUAL PROPERTY NORMS GOVERNING ROLLER DERBY PSEUDONYMS

David Fagundes
April 2012

SYMPOSIUM—RESHAPING CAPITAL MARKETS AND INSTITUTIONS: TWENTY YEARS ON

June 2012

Individual issue rate: \$15.00 per copy
Subscriptions: \$47.00 (seven issues)

Order from:

School of Law Publications
University of Texas at Austin
727 East Dean Keeton Street
Austin, Texas USA 78705
(512) 232-1149

<http://www.utexas.edu/law/publications>

Texas Law Review *See Also*

Responses to articles and notes found in this and other issues are
available at www.texasrev.com/seealso

SWITCHING TO PROPHYLACTIC INJUNCTIONS

Tracy A. Thomas

Receive notifications of all *See Also* content—sign up at www.texasrev.com.

TEXAS LAW REVIEW ASSOCIATION

OFFICERS

NINA CORTELL
President-Elect

HON. DIANE P. WOOD
President

TANIA M. CULBERTSON
Executive Director

JAMES A. HEMPHILL
Treasurer

Immediate Past President

BOARD OF DIRECTORS

ALISTAIR B. DAWSON
KARL G. DIAL
GARY L. EWELL
STEPHEN FINK
DIANA M. HUDSON
DEANNA E. KING

JEFFREY C. KUBIN
D. MCNEEL LANE
LEWIS T. LECLAIR
JOHN B. MCKNIGHT
MICHAEL H. NEWMAN
ERIC J.R. NICHOLS

ELLEN PRYOR
CHRIS REYNOLDS
DAVID M. RODI
REAGAN W. SIMPSON
STEPHEN L. TATUM
MARK L.D. WAWRO

SCOTT J. ATLAS, *ex officio Director*
MICHAEL T. RAUPP, *ex officio Director*

Texas Law Review (ISSN 0040-4411) is published seven times a year—November, December, February, March, April, May, and June. The annual subscription price is \$47.00 except as follows: Texas residents pay \$50.88 and foreign subscribers pay \$55.00. All publication rights are owned by the Texas Law Review Association. *Texas Law Review* is published under license by The University of Texas at Austin School of Law, P.O. Box 8670, Austin, Texas 78713. Periodicals Postage Paid at Austin, Texas, and at additional mailing offices.

POSTMASTER: Send address changes to The University of Texas at Austin School of Law, P.O. Box 8670, Austin, Texas 78713.

Complete sets and single issues are available from WILLIAM S. HEIN & CO., INC., 1285 Main St., Buffalo, NY 14209-1987. Phone: 1-800-828-7571.

Single issues in the current volume may be purchased from the *Texas Law Review* Publications Office for \$15.00 per copy plus shipping. Texas residents, please add applicable sales tax.

The *Texas Law Review* is pleased to consider unsolicited manuscripts for publication but regrets that it cannot return them. Please submit a single-spaced manuscript, printed on one side only, with footnotes rather than endnotes. Citations should conform with *The Greenbook: Texas Rules of Form* (12th ed. 2010) and *The Bluebook: A Uniform System of Citation* (19th ed. 2010). Except when content suggests otherwise, the *Texas Law Review* follows the guidelines set forth in the *Texas Law Review Manual on Usage & Style* (12th ed. 2011), *The Chicago Manual of Style* (16th ed. 2010), and Bryan A. Garner, *A Dictionary of Modern Legal Usage* (2d ed. 1995).

Except as otherwise noted, the *Texas Law Review* is pleased to grant permission for copies of articles, notes, and book reviews to be made for classroom use, provided that (1) a proper notice of copyright is affixed to each copy, (2) the author and source are identified, (3) copies are distributed at or below cost, and (4) the Texas Law Review Association is notified of the use.

© Copyright 2012, Texas Law Review Association

Editorial Offices: *Texas Law Review*
727 East Dean Keeton Street, Austin, Texas 78705
(512) 232-1280 Fax (512) 471-3282
tlr@law.utexas.edu
<http://www.texasrev.com>

THE UNIVERSITY OF TEXAS SCHOOL OF LAW

ADMINISTRATIVE OFFICERS

STEFANIE A. LINDQUIST, B.A., J.D., Ph.D.; *Interim Dean, A.W. Walker Centennial Chair in Law.*
ALEXANDRA W. ALBRIGHT, B.A., J.D.; *Associate Dean for Academic Affairs, Senior Lecturer.*
ROBERT M. CHESNEY, B.S., J.D.; *Associate Dean for Academic Affairs, Charles I. Francis Professor in Law.*
WILLIAM E. FORBATH, A.B., B.A., Ph.D., J.D.; *Associate Dean for Research, Lloyd M. Bentsen Chair in Law.*
EDEN E. HARRINGTON, B.A., J.D.; *Associate Dean for Clinical Education & Program Initiatives, Dir. of William Wayne Justice Ctr. for Public Interest Law, Clinical Professor.*
KIMBERLY L. BIAR, B.B.A.; *Assistant Dean for Financial Affairs, Certified Public Accountant.*
CARLA COOPER, B.A., M.A., Ph.D.; *Assistant Dean for Alumni Relations and Development.*
MICHAEL J. ESPOSITO, B.A., J.D., M.B.A.; *Assistant Dean for Continuing Legal Education.*
KIRSTON FORTUNE, B.F.A.; *Assistant Dean for Communications.*
MICHAEL HARVEY, B.A., B.S.; *Assistant Dean for Technology.*
MONICA K. INGRAM, B.A., J.D.; *Assistant Dean for Admissions and Financial Aid.*
DAVID A. MONTOYA, B.A., J.D.; *Assistant Dean for Career Services.*
REYMUENDO RAMOS, B.A.; *Assistant Dean for Student Affairs.*

FACULTY EMERITI

HANS W. BAADE, A.B., J.D., LL.B., LL.M.; *Hugh Lamar Stone Chair Emeritus in Civil Law.*
RICHARD V. BARNDT, B.S.L., LL.B.; *Professor Emeritus.*
WILLIAM W. GIBSON, JR., B.A., LL.B.; *Sylvan Lang Professor Emeritus in Law of Trusts.*
ROBERT W. HAMILTON, A.B., J.D.; *Minerva House Drysdale Regents Chair Emeritus.*
DOUGLAS LAYCOCK, B.A., J.D.; *Alice McKean Young Regents Chair Emeritus.*
J. L. LEBOWITZ, A.B., J.D., LL.M.; *Joseph C. Hutcheson Professor Emeritus.*
JOHN T. RATLIFF, JR., B.A., LL.B.; *Ben Gardner Sewell Professor Emeritus in Civil Trial Advocacy.*
MICHAEL M. SHARLOT, B.A., LL.B.; *Wright C. Morrow Professor Emeritus in Law.*
JOHN F. SUTTON, JR., J.D.; *A.W. Walker Centennial Chair Emeritus.*
JAMES M. TREECE, B.A., J.D., M.A.; *Charles I. Francis Professor Emeritus in Law.*
RUSSELL J. WEINTRAUB, B.A., J.D.; *Ben H. & Kitty King Powell Chair Emeritus in Business & Commercial Law.*

PROFESSORS

JEFFREY B. ABRAMSON, B.A., J.D., Ph.D.; *Professor of Law and Government.*
DAVID E. ADELMAN, B.A., Ph.D., J.D.; *Harry Reasoner Regents Chair in Law.*
DAVID A. ANDERSON, A.B., J.D.; *Fred & Emily Marshall Wulff Centennial Chair in Law.*
MARK L. ASCHER, B.A., M.A., J.D., LL.M.; *Joseph D. Jamail Centennial Chair in Law.*
RONEN AVRAHAM, M.B.A., LL.B., LL.M., S.J.D.; *Thomas Shelton Maxey Professor in Law.*
LYNN A. BAKER, B.A., B.A., J.D.; *Frederick M. Baron Chair in Law, Co-Director of Center on Lawyers, Civil Justice, and the Media.*
MITCHELL N. BERMAN, A.B., M.A., J.D.; *Richard Dale Endowed Chair in Law.*
BARBARA A. BINTLIFF, M.A., J.D.; *Joseph C. Hutcheson Professor in Law.*
LYNN E. BLAIS, A.B., J.D.; *Leroy G. Denman, Jr. Regents Professor in Real Property Law.*
ROBERT G. BONE, B.A., J.D.; *G. Rollie White Teaching Excellence Chair in Law.*
OREN BRACHA, LL.B., S.J.D.; *Howrey LLP and Arnold, White, & Durkee Centennial Professor.*
J. BUDZISZEWSKI, B.A., M.A., Ph.D.; *Professor.*
NORMA V. CANTU, B.A., J.D.; *Professor of Law and Education.*
LOFTUS C. CARSON, II, B.S., M. Pub. Affrs., M.B.A., J.D.; *Ronald D. Krist Professor.*
MICHAEL J. CHURGIN, A.B., J.D.; *Raybourne Thompson Centennial Professor.*
JANE M. COHEN, B.A., J.D.; *Edward Clark Centennial Professor.*
FRANK B. CROSS, B.A., J.D.; *Herbert D. Kelleher Centennial Professor of Business Law.*
WILLIAM H. CUNNINGHAM, B.A., M.B.A., Ph.D.; *Professor.*
JENS C. DAMMANN, J.D., LL.M., Dr. Jur., J.S.D.; *William Stamps Farish Professor in Law.*
JOHN DEIGH, B.A., M.A., Ph.D.; *Professor of Law and Philosophy.*
MECHELE DICKERSON, B.A., J.D.; *Arthur L. Moller Chair in Bankruptcy Law and Practice.*
GEORGE E. DIX, B.A., J.D.; *George R. Killam, Jr. Chair of Criminal Law.*
JOHN S. DZIENKOWSKI, B.B.A., J.D.; *Dean John F. Sutton, Jr. Chair in Lawyering and the Legal Process.*
KAREN L. ENGLE, B.A., J.D.; *Cecil D. Redford Prof. in Law, Director of Bernard & Audre Rapoport Center for Human Rights & Justice.*
KENNETH FLAMM, Ph.D.; *Professor.*
JULIUS G. GETMAN, B.A., LL.B., LL.M.; *Earl E. Sheffield Regents Chair.*
JOHN M. GOLDEN, A.B., J.D., Ph.D.; *Loomer Family Professor in Law.*
STEVEN GOODE, B.A., J.D.; *W. James Kronzer Chair in Trial and Appellate Advocacy, University Distinguished Teaching Professor.*
LINO A. GRAGLIA, B.A., LL.B.; *A. Dalton Cross Professor.*
CHARLES G. GROAT, B.A., M.S., Ph.D.; *Professor.*
PATRICIA I. HANSEN, A.B., M.P.A., J.D.; *J. Waddy Bullion Professor.*
HENRY T.C. HU, B.S., M.A., J.D.; *Allan Shivers Chair in the Law of Banking and Finance.*
BOBBY R. INMAN, B.A.; *Professor.*
DEREK P. JINKS, B.A., M.A., J.D.; *The Marrs McLean Professor in Law.*
STANLEY M. JOHANSON, B.S., LL.B., LL.M.; *James A. Elkins Centennial Chair in Law, University Distinguished Teaching Professor.*
CALVIN H. JOHNSON, B.A., J.D.; *Andrews & Kurth Centennial Professor.*
EMILY E. KADENS, B.A., M.A., Dipl., M.A., Ph.D., J.D.; *Baker and Bots Professor in Law.*
SUSAN R. KLEIN, B.A., J.D.; *Alice McKean Young Regents Chair in Law.*
SANFORD V. LEVINSON, A.B., Ph.D., J.D.; *W. St. John Garwood & W. St. John Garwood, Jr. Centennial Chair in Law, Professor of Gov't.*

VIJAY MAHAJAN, M.S.Ch.E., Ph.D.; *Professor.*
 BASIL S. MARKESINTS, LL.B., LL.D., DCL, Ph.D.; *Jamail Regents Chair.*
 INGA MARKOVITS, LL.M.; *"The Friends of Joe Jamail" Regents Chair.*
 RICHARD S. MARKOVITS, B.A., LL.B., Ph.D.; *John B. Connally Chair.*
 THOMAS O. MCGARITY, B.A., J.D.; *Joe R. & Teresa Lozano Long Endowed Chair in Administrative Law.*
 STEVEN A. MOORE, B.A., Ph.D.; *Professor.*
 LINDA S. MULLENIX, B.A., M. Phil., J.D., Ph.D.; *Morris & Rita Atlas Chair in Advocacy.*
 STEVEN P. NICHOLS, B.S.M.E., M.S.M.E., J.D., Ph.D.; *Professor.*
 ROBERT J. PERONI, B.S.C., J.D., LL.M.; *The Fondren Foundation Centennial Chair for Faculty Excellence.*
 H. W. PERRY, JR., B.A., M.A., Ph.D.; *Associate Professor of Law and Government.*
 LUCAS A. POWE, JR., B.A., J.D.; *Anne Green Regents Chair in Law, Professor of Government.*
 WILLIAM C. POWERS, JR., B.A., J.D.; *President of The University of Texas at Austin, Hines H. Baker & Thelma Kelley Baker Chair, University Distinguished Teaching Professor.*
 DAVID M. RABBAN, B.A., J.D.; *Dahr Jamail, Randall Hage Jamail, & Robert Lee Jamail Reg. Chair, Univ. Distinguished Teaching Prof.*
 ALAN S. RAU, B.A., LL.B.; *Mark G. & Judy G. Yudof Chair in Law.*
 DAVID W. ROBERTSON, B.A., LL.B., LL.M., J.S.D.; *W. Page Keeton Chair in Tort Law, University Distinguished Teaching Professor.*
 JOHN A. ROBERTSON, A.B., J.D.; *Vinson & Elkins Chair.*
 WILLIAM M. SAGE, A.B., M.D., J.D.; *Vice Provost for Health Affairs, James R. Dougherty Chair for Faculty Excellence.*
 LAWRENCE G. SAGER, B.A., LL.B.; *John Jeffers Research Chair in Law, Alice Jane Drysdale Sheffield Regents Chair.*
 JOHN J. SAMPSON, B.B.A., LL.B.; *William Benjamin Wynne Professor.*
 CHARLES M. SILVER, B.A., M.A., J.D.; *Roy W. & Eugenia C. MacDonald Endowed Chair in Civil Procedure, Professor in Government, Co-Director of Center on Lawyers, Civil Justice, & the Media.*
 ERNEST E. SMITH, B.A., LL.B.; *Rex G. Baker Centennial Chair in Natural Resources Law.*
 JAMES C. SPINDLER, B.A., M.A., J.D., Ph.D.; *The Sylvan Lang Professor.*
 MATTHEW L. SPITZER, B.A., Ph.D., J.D.; *Hayden W. Head Regents Chair for Faculty Excellence.*
 JANE STAPLETON, B.S., Ph.D., LL.B., D.C.L., D. Phil.; *Ernest E. Smith Professor.*
 JORDAN M. STEIKER, B.A., J.D.; *Judge Robert M. Parker Endowed Chair in Law.*
 MICHAEL F. STURLEY, B.A., J.D.; *Fannie Coptin Regents Chair.*
 GERALD TORRES, A.B., J.D., LL.M.; *Bryant Smith Chair in Law.*
 GREGORY J. VINCENT, B.A., J.D., Ed.D.; *Professor.*
 WENDY E. WAGNER, B.A., M.E.S., J.D.; *Joe A. Worsham Centennial Professor.*
 LOUISE WEINBERG, A.B., J.D., LL.M.; *William B. Bates Chair for the Administration of Justice.*
 OLIN G. WELLBORN, A.B., J.D.; *William C. Liedtke, Sr. Professor.*
 JAY L. WESTBROOK, B.A., J.D.; *Benno C. Schmidt Chair of Business Law.*
 ABRAHAM L. WICKELGREN, A.B., Ph.D., J.D.; *Bernard J. Ward Professor in Law.*
 ZIPPORAH B. WISEMAN, B.A., M.A., LL.B.; *Thos. H. Law Centennial Professor.*
 PATRICK WOOLLEY, A.B., J.D.; *Beck, Redden & Secrest Professor in Law.*

ASSISTANT PROFESSORS

MARILYN ARMOUR, B.A., M.S.W., Ph.D.	MIRA GANOR, B.A., M.B.A., LL.B., LL.M., J.S.D.
DANIEL M. BRINKS, A.B., J.D., Ph.D.	JENNIFER E. LAURIN, B.A., J.D.
JUSTIN DRIVER, B.A., M.A., M.A., J.D.	ANGELA K. LITWIN, B.A., J.D.
ZACHARY S. ELKINS, B.A., M.A., Ph.D.	MARY ROSE, A.B., M.A., Ph.D.
JOSEPH R. FISHKIN, B.A., M.Phil., D.Phil., J.D.	SEAN H. WILLIAMS, B.A., J.D.
CARY C. FRANKLIN, B.A., M.S.T., D.Phil., J.D.	

SENIOR LECTURERS, WRITING LECTURERS, AND CLINICAL PROFESSORS

WILLIAM P. ALLISON, B.A., J.D.; <i>Clinical Prof., Dir. of Criminal Defense Clinic.</i>	HARRISON KELLER, B.A., M.A., Ph.D.; <i>Vice Provost for Higher Education Policy, Senior Lecturer.</i>
MARJORIE I. BACHMAN, B.S., J.D.; <i>Clinical Instructor</i>	JEANA A. LUNGWITZ, B.A., J.D.; <i>Clinical Prof., Dir. of Domestic Violence Clinic.</i>
PHILIP C. BOBBITT, A.B., J.D., Ph.D.; <i>Distinguished Sr. Lecturer.</i>	TRACY W. McCORMACK, B.A., J.D.; <i>Lect.; Dir. Trial Advoc. Prg.</i>
KAMELA S. BRIDGES, B.A., B.J., J.D.; <i>Lecturer.</i>	ROBIN B. MEYER, B.A., M.A., J.D.; <i>Lecturer.</i>
CYNTHIA L. BRYANT, B.A., J.D.; <i>Clinical Prof., Dir. of Mediation Clinic.</i>	JANE A. O'CONNELL, B.A., M.S., J.D.; <i>Lecturer, Assoc. Dir. for Patron Services, Instruction, & Research.</i>
PAUL J. BURKA, B.A., LL.B.; <i>Senior Lecturer.</i>	RANJANA NATARAJAN, B.A., J.D.; <i>Clinical Prof., Dir. of National Security Clinic.</i>
JOHN C. BUTLER, B.B.A., Ph.D.; <i>Clinical Assoc. Prof.</i>	ROBERT C. OWEN, A.B., M.A., J.D.; <i>Clinical Prof.</i>
MARY R. CROUTER, A.B., J.D.; <i>Lecturer, Asst. Dir. of William Wayne Justice Center for Public Interest Law.</i>	SEAN J. PETRIE, B.A., J.D.; <i>Lecturer.</i>
TIFFANY J. DOWLING, B.A., J.D.; <i>Clinical Instructor, Dir. of Actual Innocence Clinic.</i>	WAYNE SCHIESS, B.A., J.D.; <i>Sr. Lecturer, Dir. of Legal Writing.</i>
LORI K. DUKE, B.A., J.D.; <i>Clinical Prof.</i>	STACY ROGERS SHARP, B.S., J.D.; <i>Lecturer.</i>
ARIEL E. DULITZKY, J.D., LL.M.; <i>Clinical Prof., Dir. of Human Rights Clinic.</i>	PAMELA J. SIGMAN, B.A., J.D.; <i>Adjunct Prof., Dir. Juv. Just. Clinic.</i>
ELANA S. EINHORN, B.A., J.D.; <i>Lecturer.</i>	DAVID S. SOKOLOV, B.A., M.A., J.D., M.B.A.; <i>Distinguished Sr. Lecturer, Dir. of Student Life.</i>
TINA V. FERNANDEZ, A.B., J.D.; <i>Lecturer.</i>	LESLIE L. STRAUCH, B.A., J.D.; <i>Clinical Prof. Child. Rights Clinic.</i>
LYNDA E. FROST, B.A., M.Ed., J.D., Ph.D.; <i>Clinical Assoc. Prof.</i>	GRETCHEN S. SWEEN, B.A., M.A., Ph.D., J.D.; <i>Lecturer.</i>
DENISE L. GILMAN, B.A., J.D.; <i>Clinical Prof., Dir. of Immigr. Clinic.</i>	MELINDA E. TAYLOR, B.A., J.D.; <i>Sr. Lecturer, Exec. Dir. of Ctr. for Global Energy, Int'l Arbitration, & Environmental Law.</i>
KELLY L. HARAGAN, B.A., J.D.; <i>Lecturer, Dir. of Environmental Law Clinic.</i>	HEATHER K. WAY, B.A., B.J., J.D.; <i>Lecturer, Dir. of Commun. Dev. Clinic.</i>
BARBARA HINES, B.A., J.D.; <i>Clinical Prof., Dir. of Immigr. Clinic.</i>	ELIZABETH M. YOUNGDALE, B.A., M.L.I.S., J.D.; <i>Lecturer.</i>

ADJUNCT PROFESSORS AND OTHER LECTURERS

ROBERT J. ADAMS, JR., B.S., M.B.A., Ph.D.
 WILLIAM R. ALLENSWORTH, B.A., J.D.
 CRAIG D. BALL, B.A., J.D.
 SHARON C. BAXTER, B.S., J.D.
 KARL O. BAYER, B.A., M.S., J.D.
 WILLIAM H. BEARDALL, JR., B.A., J.D.
 JERRY A. BELL, B.A., J.D.
 ALLISON H. BENESCH, B.A., M.S.W., J.D.
 CRAIG R. BENNETT, B.S., J.D.
 JAMES B. BENNETT, B.B.A., J.D.
 MELISSA J. BERNSTEIN, B.A., M.L.S., J.D.
 ROBERT S. BICKERSTAFF, B.A., J.D.
 MURFF F. BLEDSOE, B.A., J.D.
 WILLIAM P. BOWERS, B.B.A., J.D., LL.M.
 HUGH L. BRADY, B.A., J.D.
 STACY L. BRAININ, B.A., J.D.
 ANTHONY W. BROWN, B.A., J.D.
 JAMES E. BROWN, B.A., LL.B.
 TOMMY L. BROYLES, B.A., J.D.
 W.A. BURTON, JR., B.A., M.A., LL.B.
 JOHN C. BUTLER, B.B.A., Ph.D.
 AGNES E. CASAS, B.A., J.D.
 RUBEN V. CASTANEDA, B.A., J.D.
 EDWARD A. CAVAZOS, B.A., J.D.
 ELIZABETH S. CHESTNEY, B.A., J.D.
 JEFF CIVINS, A.B., M.S., J.D.
 GARY COBB, B.A., J.D.
 ELIZABETH COHEN, B.A., M.S.W., J.D.
 JAMES W. COLLINS, B.S., J.D.
 PATRICIA J. CUMMINGS, B.A., J.D.
 DICK DEGUERIN, B.A., LL.B.
 ADAM R. DELL, B.A., J.D.
 STEVEN K. DEWOLF, B.A., J.D., LL.M.
 CASEY D. DUNCAN, B.A., M.L.I.S., J.D.
 PHILIP DURST, B.A., M.A., J.D.
 BILLIE J. ELLIS, JR., B.A., M.B.A., J.D.
 JAY D. ELLWANGER, B.A., J.D.
 TORIA J. FINCH, B.S., J.D.
 JOHN C. FLEMING, B.A., J.D.
 KYLE K. FOX, B.A., J.D.
 DAVID C. FREDERICK, B.A., Ph.D., J.D.
 GREGORY D. FREED, B.A., J.D.
 FRED J. FUCHS, B.A., J.D.
 CHARLES E. GHOLZ, B.S., Ph.D.
 MICHAEL J. GOLDEN, A.B., J.D.
 DAVID HALPERN, B.A., J.D.
 ELIZABETH HALUSKA-RAUSCH, B.A., M.A., M.S., Ph.D.
 JETT L. HANNA, B.B.A., J.D.
 KELLY L. HARAGAN, B.A., J.D.
 CLINT A. HARBOUR, B.A., J.D., LL.M.
 ROBERT L. HARGETT, B.B.A., J.D.
 JAMES C. HARRINGTON, B.A., M.A., J.D.
 CHRISTOPHER S. HARRISON, Ph.D., J.D.
 JOHN R. HAYS, JR., B.A., J.D.
 P. MICHAEL HEBERT, A.B., J.D.
 STEVEN L. HIGHLANDER, B.A., Ph.D., J.D.
 SUSAN J. HIGHTOWER, B.A., M.A., J.D.
 KENNETH E. HOUP, JR., J.D.
 RANDY R. HOWRY, B.J., J.D.
 MONTY G. HUMBLE, B.A., J.D.
 PATRICK O. KEEL, B.A., J.D.
 DOUGLAS L. KEENE, B.A., M.Ed., Ph.D.
 SCOTT A. KELLER, B.A., J.D.
 CHARI L. KELLY, B.A., J.D.
 ROBERT N. KEPPLER, B.A., J.D.
 MARK L. KINCAID, B.B.A., J.D.
 KURT H. KUHN, B.A., J.D.
 AMI L. LARSON, B.A., J.D.
 KEVIN R. LASHUS, B.A., J.D.
 JODI R. LAZAR, B.A., J.D.
 KEVIN L. LEAHY, B.A., J.D.
 MAURIE A. LEVIN, B.A., J.D.
 JIM MARCUS, B.A., J.D.
 PETER D. MARKETOS, B.A., J.D.
 HARRY S. MARTIN, A.B., M.L.S., J.D.
 FRANCES L. MARTINEZ, B.A., J.D.
 LAURA A. MARTINEZ, B.A., J.D.
 RAY MARTINEZ, III, B.A., J.D.
 LISA M. MCCLAIN, B.A., J.D., LL.M.
 BARRY F. MCNEIL, B.A., J.D.
 ANGELA T. MELINARAAB, B.F.A., J.D.
 MARGARET M. MENICUCCI, B.A., J.D.
 JO A. MERICA, B.A., J.D.
 RANELLE M. MERONEY, B.A., J.D.
 ELIZABETH N. MILLER, B.A., J.D.
 DARYL L. MOORE, B.A., M.L.A., J.D.
 EDWIN G. MORRIS, B.S., J.D.
 SARAH J. MUNSON, B.A., J.D.
 MANUEL H. NEWBURGER, B.A., J.D.
 MILAM F. NEWBY, B.A., J.D.
 DAVID G. NIX, B.S.E., LL.M., J.D.
 PATRICK L. O'DANIEL, B.B.A., J.D.
 M. A. PAYAN, B.A., J.D.
 ELIZA T. PLATTS-MILLS, B.A., J.D.
 LAURA L. PRATHER, B.B.A., J.D.
 JONATHAN PRATTER, B.A., M.L.S., J.D.
 VELVA L. PRICE, B.A., J.D.
 BRIAN C. RIDER, B.A., J.D.
 ROBERT M. ROACH, JR., B.A., J.D.
 BRIAN J. ROARK, B.A., J.D.
 BETTY E. RODRIGUEZ, B.S.W., J.D.
 JAMES D. ROWE, B.A., J.D.
 MATTHEW C. RYAN, B.A., J.D.
 KAREN R. SAGE, B.A., J.D.
 MARK A. SANTOS, B.A., J.D.
 CHRISTOPHE H. SAPSTEAD, B.A., J.D.
 SUSAN SCHULTZ, B.S., J.D.
 AMY J. SCHUMACHER, B.A., J.D.
 SUZANNE SCHWARTZ, B.J., J.D.
 RICHARD J. SEGURA, JR., B.A., J.D.
 DAVID A. SHEPPARD, B.A., J.D.
 JUDGE ERIC M. SHEPPERD, B.A., J.D.
 RONALD J. SIEVERT, B.A., J.D.
 AMBROSIO A. SILVA, B.S., J.D.
 STUART R. SINGER, A.B., J.D.
 HON. BEA A. SMITH, B.A., M.A., J.D.
 LYDIA N. SOLIZ, B.B.A., J.D.
 KATHERINE J. SOLON, B.A., M.S., J.D.
 JAMES M. SPELLINGS, JR., B.S., J.D.
 DAVID B. SPENCE, B.A., J.D., M.A., Ph.D.
 KACIE L. STARR, B.A., J.D.
 WILLIAM F. STUTTS, B.A., J.D.
 MATTHEW J. SULLIVAN, B.S., J.D.
 JEREMY S. SYLESTINE, B.A., J.D.
 BRADLEY P. TEMPLE, B.A., J.D.
 SHERINE E. THOMAS, B.A., J.D.
 TERRY O. TOTTENHAM, B.S., LL.M., J.D.
 MICHAEL S. TRUESDALE, B.A., M.A., J.D.
 JEFFREY K. TULIS, B.A., M.A., Ph.D.
 ROBERT W. TURNER, B.A., LL.B.
 TIMOTHY J. TYLER, B.A., J.D.
 SUSAN S. VANCE, B.B.A., J.D.
 LANA K. VARNEY, B.J., J.D.
 SRIRAM VISHWANATH, B.S., M.S., Ph.D.
 DEBORAH M. WAGNER, B.A., M.A., J.D.
 CLARK C. WATTS, B.A., M.D., M.A., M.S., J.D.
 JANE M. WEBRE, B.A., J.D.
 WARE V. WENDELL, A.B., J.D.
 RODERICK E. WETSEL, B.A., J.D.
 DARAJ. WHITEHEAD, B.A., M.S.
 RANDALL B. WILHITE, B.B.A., J.D.
 DAVID G. WILLE, B.S.E.E., M.S.E.E., J.D.

MARK B. WILSON, B.A., M.A., J.D.
CHRISTINA T. WISDOM, B.A., J.D.
JUDGE PAUL L. WOMACK, B.S., J.D.
LUCILLE D. WOOD, B.A., J.D.

DENNEY L. WRIGHT, B.B.A., J.D., LL.M.
LARRY F. YORK, B.B.A., LL.B.
DANIEL J. YOUNG, B.A., J.D.

VISITING PROFESSORS

OWEN L. ANDERSON, B.A., J.D.
ANTONIO H. BENJAMIN, LL.B., LL.M.
PETER F. CANE, B.A., LL.B., D.C.L.

VICTOR FERRERES, J.D., LL.M., J.S.D.
LARRY LAUDAN, B.A., M.A., Ph.D.
NADAV SHOKED, LL.B., S.J.D.

Texas Law Review

Volume 90

Number 6

May 2012

MICHAEL T. RAUPP
Editor in Chief

DANIEL C. CLEMONS
Managing Editor

DREW D. PENNEBAKER
Chief Articles Editor

TANIA M. CULBERTSON
Administrative Editor

SIDNEY K. SMITH
Chief Notes Editor

NICHOLAS D. STEPP
Book Review Editor

STEPHEN A. FRASER
Online Content Editor

ATHENA J. PONCE
Research Editor

DAVID E. ARMENDARIZ
PAIGE M. JENNINGS
GABRIEL H. MARKOFF
KATHLEEN L. NANNEY
Articles Editors

NEIL K. GEHLAWAT
KRISTIN M. MALONE
KARSON K. THOMPSON
Notes Editors

JOSHUA H. PACKMAN
GEORGE M. PADIS
JEFFREY T. QUILICI
ERIC T. WERLINGER
Articles Editors

F. GIBBONS ADDISON
BETH N. BEAURY
JUSTIN L. BERNSTEIN
ROBERT W. BROWN
ROBERT T. CARMAN

L. JOSEPH DENBINA
ADRIAN L. DUNBAR
RYAN M. GOLDSTEIN
ALDEN G. HARRIS

ANDREW T. INGRAM
DAVID J. KOHTZ
LINDSEY A. MILLS
ADAM J. MOSS
KEITH P. SYSKA

Associate Editors

Members

MICHAEL ABRAMS
BRITTANY R. ARTIMEZ
ALESE L. BAGDOL
BRIAN J. BAH
JULIA C. BARRETT
MOLLY M. BARRON
BRADEN A. BEARD
ERIC S. BERELOVICH
CECILIA BERNSTEIN
CHRISTOPHER M. BEZEG
DAWSON A. BROTEMARKLE
KRISTIN C. BURNETT
CHARLES D. CASSIDY
BENJAMIN D. CLARK
CHASE E. COOLEY
WILLIAM P. COURTNEY
LESLIE S. DYKE
MAXIM FARBEROV
SHEILA FORJUOH
CHARLES E. FOWLER
AMELIA A. FRIEDMAN

SCOTT J. FULFORD
ALLISON L. FULLER
ERIN L. GAINES
MONICA E. GAUDIOSO
PARTH S. GEJI
DANIEL D. GRAVER
MATTHEW A. GREENBERG
SEAN M. HILL
ALEXANDER G. HUGHES
MONICA R. HUGHES
MATTHEW P. KAIPUST
LISA D. KINZER
JONATHAN LEVY
TYSON M. LIES
NATHANIEL H. LIPANOVICH
ROSS M. MACDONALD
RALPH C. MAYRELL
WILLIAM J. MCKINNON
KYLE E. MITCHELL
BENJAMIN S. MORGAN
MATTHEW M. OLSON
CORBIN D. PAGE

CHRISTOPHER S. PATTERSON
ADAM R. PERKINS
KATHRYN G. RAWLINGS
JEREMY M. REICHMAN
BRETT S. ROSENTHAL
LAUREN K. ROSS
RAFE A. SCHAEFER
MICHAEL N. SELKIRK
STEPHEN STECKER
KASYN M. STEVENSON
MARTHA L. TODD
WILLIAM C. VAUGHN
COLIN M. WATTERSON
JAMES T. WEISS
SARAH K. WEYNAND
COLLIN R. WHITE
LECH WILKIEWICZ
WRIGHT T. WOMMACK
CATHARINE S. WRIGHT
D. PATRICK A. YARBOROUGH
JOYCE Y. YOUNG

PAUL N. GOLDMAN
Business Manager

MITCHELL N. BERMAN
JOHN S. DZIENKOWSKI
Faculty Advisors

TERI GAUS
Editorial Assistant

Texas Law Review

Volume 90, Number 6, May 2012

ARTICLES

- Party Rulemaking: Making Procedural
Rules Through Party Choice
Robert G. Bone 1329
- Injunctions as More (or Less) than “Off Switches”:
Patent-Infringement Injunctions’ Scope
John M. Golden 1399

BOOK REVIEWS

- Still Convicting the Innocent
Jennifer E. Laurin 1473
- reviewing* Brandon L. Garrett’s
CONVICTING THE INNOCENT: WHERE
CRIMINAL PROSECUTIONS GO WRONG
- Constructing a Canon of Law-Related Poetry
Alexandra J. Roberts 1507
- reviewing*
POETRY OF THE LAW:
FROM CHAUCER TO THE PRESENT

NOTES

- Monetary Damages and the (b)(2) Class Action:
A Closer Look at *Wal-Mart v. Dukes*
Neil K. Gehlawat 1535
- Would You Want William Hung as Your Trier of Fact?
The Case for a Specialized Musicology Tribunal
Eric M. Leventhal 1557

Articles

Party Rulemaking: Making Procedural Rules Through Party Choice

Robert G. Bone*

I.	Introduction.....	1330
II.	Framing the Problem	1337
A.	The Problem Clarified and in Context	1337
B.	Current Law on Party Rulemaking	1342
1.	<i>Forum Choice</i>	1343
2.	<i>Pretrial</i>	1345
3.	<i>Trial</i>	1348
4.	<i>Remedies</i>	1350
5.	<i>Appeals</i>	1351
6.	<i>Summary</i>	1351
III.	The Normative Dimension: Conventional Arguments	1352
A.	Considerations Favoring Party Rulemaking	1354
1.	<i>The Flawed Argument from Arbitration</i>	1354
2.	<i>The Benefits of Party Rulemaking</i>	1355
B.	Considerations Opposing Party Rulemaking	1360
1.	<i>The Problem of Consent</i>	1360
2.	<i>The Baseline Argument</i>	1369
3.	<i>The Costs of Party Rulemaking</i>	1372
C.	Balancing the Competing Considerations.....	1380
IV.	The Legitimacy Critique.....	1384
A.	Identifying Adjudication’s Core	1385
B.	Party Choice and Principled Reason.....	1391
C.	Justifying the Limits	1394
V.	Conclusion	1397

* G. Rollie White Professor of Law, The University of Texas School of Law. I am grateful to Larry Sager, Victor Ferreres, Alan Rau, Joey Fishkin, and Larry Ribstein for extremely helpful conversations, guidance, and comments and to participants in faculty colloquia at the University of Illinois College of Law and The University of Texas School of Law. I would also like to thank Ashwin Rao and Neil Gehlawat for their valuable research assistance.

I. Introduction

The procedures used to adjudicate civil lawsuits in federal court have many different sources. Some procedures are constitutionally mandated. The Due Process Clause of the United States Constitution, for example, imposes minimum notice and hearing requirements,¹ and the Seventh Amendment guarantees a jury trial in many cases.² Other procedures are prescribed by statute. The Private Securities Litigation Reform Act, for example, creates strict pleading requirements for securities fraud suits³ and penalizes frivolous filings.⁴ Most procedures are based on the Federal Rules of Civil Procedure (FRCP), the work of a committee-based rulemaking process authorized by statute.⁵ In addition, each federal district has the power to make local rules for cases litigated in its own district courts,⁶ and district judges have broad discretion to create procedures tailored to a specific case through pretrial management and trial supervision.⁷ And federal judges also have some power to make procedural rules through the common law process.⁸

This Article focuses on yet another source of procedure: party choice. One aspect of party-made procedure is fundamental to litigation. In the American adversary system, litigants enjoy broad freedom to make their own litigation choices. They do so within the constraints and guidelines of general procedural rules and case-specific judicial-management decisions, but these constraints leave wide latitude for parties to make procedural moves strategically.⁹

1. U.S. CONST. amend. V.

2. *Id.* amend. VII.

3. Private Securities Litigation Reform Act of 1995, Pub. L. No. 104-67, § 101(b), 109 Stat. 737, 746-47 (codified as amended at 15 U.S.C. § 78u-4(b) (2006)).

4. *Id.* § 101(c), 109 Stat. 737, 747-48 (codified as amended at 15 U.S.C. § 78u-4(c) (2006)).

5. Rules Enabling Act § 401(a), 28 U.S.C. §§ 2072-2073 (2006).

6. FED. R. CIV. P. 83. Also, federal judges have the power to adopt standing orders for cases that they hear. *Id.*

7. See Robert G. Bone, *Who Decides? A Critical Look at Procedural Discretion*, 28 CARDOZO L. REV. 1961, 1965-70 (2007) [hereinafter Bone, *Procedural Discretion*] (explaining that judges enjoy both “explicit” case-management discretion delegated by the Federal Rules and “interpretive” discretion where those Rules are vague).

8. Judges also have latitude to craft general procedural rules through interpretation of vague constitutional, statutory, and FRCP language. Technically, these interpretations are part of the law being interpreted. Sometimes, however, it is difficult to distinguish between rule interpretation and common law creation, especially when a judge interprets an extremely vague Federal Rule of Civil Procedure, such as Rule 56 before the 2010 amendments. Compare, e.g., *Celotex Corp. v. Catrett*, 477 U.S. 317, 322-25 (1986) (purporting to construe the provision of FRCP 56(c) requiring that the litigants’ filings “show that there is no genuine issue as to any material fact” to support a rule that the movant need only “point[] out” a lack of evidence supporting the nonmovant’s case), with *id.* at 337-39 (Stevens, J., dissenting) (criticizing the Court’s “abstract exercise in Rule construction”).

9. For example, a plaintiff has broad freedom to choose how much factual detail to include in her complaint, how much discovery to take, and which parties and claims to join to her suit—all within the constraints of the relevant procedural rules and subject to the defendant’s likely response.

This point is obvious perhaps, but it is also important. The fact that parties exercise substantial control over how their lawsuits are adjudicated raises the question of whether they should be able to control procedure in other ways as well. There are three additional possibilities: First, parties might commit in advance to the same actions they could choose strategically during litigation. This is possible to some extent today, and subpart II(B) describes typical instances where it occurs, such as when parties select the forum, contractually waive jury trial, or limit discovery in advance. In effect, the parties create a simple rule for the later suit that requires the action the parties have agreed to take.

Second, parties can make procedure by precommitting to general procedural rules for their lawsuit. They do this now when they choose rules to govern an arbitration proceeding, and there is no conceptual problem with extending the same privilege to adjudication. If this were done, existing procedural rules would serve as defaults and apply only when the parties failed to choose something different. In fact, the current system permits some party rulemaking along these lines. As subpart II(B) explains, parties can contract in advance for a shorter statute of limitations and for exceptions to the American rule on attorneys' fees. However, this type of party rulemaking appears to be extremely limited. This prompts the question of how much more freedom parties should have. Should they be able to contract for a different pleading standard or summary judgment test, or even a particular jury composition or method of judicial decision?¹⁰

The third way parties might make procedure is to choose rules not just in advance of a suit, but also at the time of implementation. This third option is more difficult to envision because it is quite foreign to litigation as we know it. But it is conceptually feasible. Suppose a plaintiff, at the time she files her complaint, could choose among different combinations of pleading standards, discovery limits, and summary judgment procedures.¹¹ If she chose a strict pleading standard, for example, she would be allowed to opt for broad discovery, and any summary judgment motions would be postponed until later in the suit. If she chose a liberal pleading standard, on the other

See FED. R. CIV. P. 8–9 (addressing pleading); FED. R. CIV. P. 26–37 (addressing discovery); FED. R. CIV. P. 13–14, 18–20, 24 (addressing joinder).

10. An extreme example would be an agreement for the case to be decided by a judicial coin flip. Perhaps more realistically, parties could agree that the judge decide on the basis of moral principles, even if those principles conflict with the law. Or they might agree that the judge use a particular approach to statutory interpretation or narrow her attention to a delimited set of precedents. Cf. Frank Partnoy, *Synthetic Common Law*, 53 U. KAN. L. REV. 281, 283 (2005) (proposing a system of private judges who decide cases according to a menu of precedents selected or designed by the parties).

11. We might also involve the defendant in the choice by giving him a veto or allowing him to modify the plaintiff's choice. One compelling reason to give these options to the parties would be to harness the superior information parties have about their own lawsuits. However, it would be important to design the procedural combinations and the choice conditions in a way that incentivized parties to make the socially optimal choice for their case.

hand, discovery would be limited and summary judgment available at an earlier stage.¹² The idea would be to allow the plaintiff to make her own trade-off between case-screening methods and information-access opportunities. Notice that, unlike the second possibility, there is no precommitment in these scenarios. The choice results from noncooperative strategic interaction during litigation.

In previous writing, I have discussed the relative merits of three different methods of procedural rulemaking: the committee-based rulemaking process currently in place, legislative rulemaking, and common law.¹³ This Article examines a fourth possibility: procedural rulemaking by parties. It focuses on the first two modes of party rulemaking described above: committing in advance to a specific action, and committing in advance to a general rule different from the official rule that would otherwise apply. The third mode, choosing general rules noncooperatively at the time of implementation, is too complicated to explore here and deserves separate treatment.

The proper scope of party rulemaking is an important topic today. Ever since the Supreme Court opened the door to liberal enforcement of forum-selection clauses in *The Bremen v. Zapata Off-Shore Co.*¹⁴ and extended its holding to consumer contracts in *Carnival Cruise Lines, Inc. v. Shute*,¹⁵ the question of party freedom to contract for procedure has received a great deal of scholarly attention.¹⁶ In recent years, the stakes have risen. Companies now include arbitration clauses in consumer and employment contracts to avoid undesirable procedures in court. For example, in *AT&T Mobility LLC*

12. This assumes, of course, that plaintiffs use the threat of broad discovery to leverage settlements in frivolous and weak suits. See, e.g., *Bell Atlantic Corp. v. Twombly*, 550 U.S. 554, 557–59 (2007) (expressing concern about high discovery costs pressuring unjustified settlements). If the judge could screen more vigorously for frivolous suits at the pleading stage, there would be less need for a summary judgment screen and less concern about limiting discovery.

13. Robert G. Bone, *The Process of Making Process: Court Rulemaking, Democratic Legitimacy, and Procedural Efficacy*, 87 GEO. L.J. 887 (1999) [hereinafter Bone, *The Process of Making Process*].

14. 407 U.S. 1, 12–14 (1972).

15. 499 U.S. 585, 593–95 (1991).

16. See generally, e.g., Linda S. Mullenix, *Another Choice of Forum, Another Choice of Law: Consensual Adjudicatory Procedure in Federal Court*, 57 FORDHAM L. REV. 291 (1988) (analyzing and critiquing “consensual adjudicatory procedure”). Professors Paul Carrington and Paul Haagen wrote one of the early, in-depth critical examinations of the Supreme Court’s generous approach to forum-selection agreements, both for courts and for arbitration. Paul D. Carrington & Paul H. Haagen, *Contract and Jurisdiction*, 1996 SUP. CT. REV. 331. These early articles inspired an extensive follow-on literature, which was fueled as well by the Supreme Court’s increasingly solicitous attitude toward arbitration agreements during the 1990s and 2000s. See, e.g., Jean R. Sternlight, *Rethinking the Constitutionality of the Supreme Court’s Preference for Binding Arbitration: A Fresh Assessment of Jury Trial, Separation of Powers, and Due Process Concerns*, 72 TUL. L. REV. 1 (1997) (challenging some widespread uses of arbitration as possibly unconstitutional); see also *infra* notes 21–22 (collecting other sources). For background on the Supreme Court’s expansion of arbitration, see THOMAS E. CARBONNEAU, *THE LAW AND PRACTICE OF ARBITRATION*, at xi–xxix (3d ed. 2009).

v. Concepcion,¹⁷ the United States Supreme Court last term reversed a Ninth Circuit decision that refused to enforce a class-action-waiver clause as part of a consumer arbitration agreement.¹⁸ The arbitration and class-action-waiver clauses were included in a service contract between a cell phone company and its subscribers, and the company used the provisions to avoid consumer class actions in arbitration.¹⁹ This is not an example of party rulemaking within adjudication; the company altered applicable procedures by switching to arbitration. But it does raise the question of how far parties should be allowed to shape their own procedures. Indeed, this question applies equally to arbitration and adjudication. For example, had the Supreme Court in *Concepcion* affirmed the Ninth Circuit's holding, firms might well have shunned consumer arbitration altogether and switched back to adjudication.²⁰ In that case, they would almost certainly have tried to use contractual waivers to avoid class actions in court, so the courts would have had to deal with the enforceability of these waivers in the adjudication setting. Thus, the general question is whether there should be limits on the power of parties to contractually modify procedural rules and whether those limits should differ between adjudication and arbitration.

The debate over party rulemaking is sharply divided. Some scholars favor broad freedom to customize procedure both before and after a dispute arises.²¹ Others urge much stricter limits.²² Both sides agree that relative

17. 131 S. Ct. 1740 (2011), *rev'g* Laster v. AT&T Mobility LLC, 584 F.3d 849 (9th Cir. 2009).

18. *Id.* at 1745, 1753.

19. *Id.* at 1744.

20. See Transcript of Oral Argument at 42, *Concepcion*, 131 S. Ct. 1740 (No. 09-893) (statement of Kagan, J.) (noting that refusals to enforce class action waivers are likely to encourage companies to exit arbitration); *cf.* Samuel Issacharoff & Erin F. Delaney, *Credit Card Accountability*, 73 U. CHI. L. REV. 157, 179 (2006) (explaining that credit card companies have proven to be “even less enthusiastic about classwide arbitration than about class action litigation,” illustrating the point with an example of one company whose contractual language revealed that if it “[could]n’t compel individual arbitration, it [did]n’t want to be in arbitration at all”).

21. See, e.g., Daphna Kapeliuk & Alon Klement, *Contracting Around Twombly*, 60 DEPAUL L. REV. 1, 1–2 (2010) (arguing that if contracting parties are allowed to modify pleading standards, “they will be able to solve problems of inadequate screening and to realize both pre- and post-dispute opportunities that would be unworkable otherwise,” resulting in improved “justice and efficiency”); Michael L. Moffitt, *Customized Litigation: The Case for Making Civil Procedure Negotiable*, 75 GEO. WASH. L. REV. 461, 462 (2007) (suggesting that existing procedural rules should be treated as “default rules, rather than as nonnegotiable parameters”); Henry S. Noyes, *If You (Re)Build It, They Will Come: Contracts to Remake the Rules of Litigation in Arbitration’s Image*, 30 HARV. J.L. & PUB. POL’Y 579, 581 (2007) (proposing a system whereby instead of opting out of the public system of adjudication in favor of private arbitration, parties “opt-in” and choose the public courts as the forum for dispute resolution, yet waive, modify, and displace the ‘normal’ litigation rules”); Daphna Kapeliuk & Alon Klement, *Contractualizing Procedure 3* (Dec. 31, 2008) (unpublished manuscript) [hereinafter Kapeliuk & Klement, *Contractualizing Procedure*], available at <http://ssrn.com/abstract=1323056> (“Unlike the limited benefits that litigants can gain from procedural arrangements made after the dispute arises, namely reduction of litigation costs and risks, pre-dispute modifications of procedural rules can achieve significant benefits not only after the dispute, but also before it.”); see also Robert E. Scott & George G. Triantis, *Anticipating Litigation in Contract Design*, 115 YALE L.J. 814, 856–60 (2006) (explaining how parties can benefit by designing litigation procedures through contract).

bargaining power and externalities are important factors, but they disagree about how the costs and benefits should be weighed. Proponents of a broad scope for party rulemaking argue that the efficiency and autonomy benefits outweigh the costs.²³ Those who advocate strict limits place greater weight on the potential unfairness to weaker parties and focus more strongly on risks to adjudicative legitimacy and on other institutional costs.²⁴

The problem with these arguments is that they are insufficiently theorized. Both sides assume that all the benefits and costs can be catalogued, roughly measured, and weighed properly. But this assumption is flawed. Benefits and costs are extremely difficult to evaluate in the intensely strategic environment of litigation. For example, it might seem that contractual limits on discovery will reduce the social costs of litigation, but in fact, they might increase costs if the lower discovery burden reduces the gains from settlement and makes trial more attractive.²⁵

22. See, e.g., Judith Resnik, *Procedure as Contract*, 80 NOTRE DAME L. REV. 593, 622–26, 666 (2005) (emphasizing the importance of adjudicating conflicts with public rules); David H. Taylor & Sara M. Cliffe, *Civil Procedure by Contract: A Convolved Confluence of Private Contract and Public Procedure in Need of Congressional Control*, 35 U. RICH. L. REV. 1085, 1159–61 (2002) (arguing for congressional limitations on private procedural agreements); Elizabeth Thornburg, *Designer Trials*, 2006 J. DISP. RESOL. 181, 208–10 (recounting some of the benefits to having judges control procedure and cautioning that certain equity-based factors should weigh heavily in a judge’s decision about whether to enforce “a contract for a customized trial”); see also Kevin E. Davis & Helen Hershkoff, *Contracting for Procedure*, 53 WM. & MARY L. REV. 507 (2011) (taking a critical but more balanced view); Jaime Dodge, *The Limits of Procedural Private Ordering*, 97 VA. L. REV. 723 (2011) (same). See generally Christopher R. Drahozal & Peter B. Rutledge, *Contract and Procedure*, 94 MARQ. L. REV. 1103 (2011) (reviewing the debate). A leading procedure treatise is quite hostile to party rulemaking. See 8 CHARLES ALAN WRIGHT, ARTHUR R. MILLER & RICHARD L. MARCUS, FEDERAL PRACTICE AND PROCEDURE § 2005, at 52 (3d ed. 2010) (criticizing enforcement of an agreement that would disregard procedural rules designed to serve the public interest); 14D CHARLES ALAN WRIGHT, ARTHUR R. MILLER & EDWARD H. COOPER, FEDERAL PRACTICE AND PROCEDURE § 3803.1, at 76 (3d ed. 2007) (championing the benefits of an approach to forum-selection clauses that limits the parties’ freedom to refashion the FRCP).

23. See, e.g., Kapeliuk & Klement, *Contractualizing Procedure*, *supra* note 21, at 2 (endorsing and expanding on the insight of Professors Scott and Triantis that “contracting parties can structure procedural rules in ways that would increase their joint surplus from the contract,” and agreeing that “by varying the degree of precision of contract provisions and terminology, contracting parties can shift costs between the time of contracting and the time of dispute, in ways that would prove beneficial for both” (citing Scott & Triantis, *supra* note 21, at 856–60)).

24. See, e.g., Resnik, *supra* note 22, at 623–24 (noting that the traditional conception of adjudication values “public and disciplined factfinding,” “norm enforcement,” and transparent process as central to judicial legitimacy, and that the new Contract Model of Procedure threatens those values); Taylor & Cliffe, *supra* note 22, at 1100, 1103–04 (noting that private procedural agreements operate “outside of the publicly crafted rules” with the result that “wisdom gained from public debate is subverted” and also observing that private procedure can give one side unfair strategic advantages); Thornburg, *supra* note 22, at 210 (stressing, among other things, “the development of legal rules” and “the educational and symbolic functions of the courts” and concluding that “[t]he normative and political values supporting public adjudication cannot be bargained away without the involvement of anyone representing the public interest” (footnote omitted)).

25. See *infra* note 115 and accompanying text.

There is an even more serious problem. I argue in Part III below that the most salient objection to party rulemaking has to do with its potential impact on adjudicative legitimacy; yet commentators tend to treat the legitimacy objection superficially, as a concern about public perceptions or adverse reputational effects for the judiciary.²⁶ As Part IV explains, legitimacy is about much more than public perception. If giving parties the power to design court procedure threatens the normative legitimacy of adjudication, it must be because party rulemaking offends some feature of adjudication that is essential to its legitimacy. It follows then that one must be prepared to explain which features are essential, and this requires a deep theoretical inquiry into the nature of adjudication. Critics of party rulemaking do not even attempt that task.

In fact, there are widely accepted features of the current system of civil litigation that are difficult to square with an aversion to broad party choice. For one thing, parties can design their own procedure through the simple expedient of agreeing to arbitration. This at least raises the question of why they should not be allowed to do the same thing in adjudication. Moreover, parties have considerable freedom to choose which substantive law to apply to their dispute.²⁷ Why then should they not have broad control over procedure?²⁸ The prevalence of settlement also presents serious problems for critics of party rulemaking. Settlement allows parties to alter the default system of procedure radically by replacing adversarial adjudication with a process of negotiation or mediation. Most people believe that settlements achieve good results under a wide range of conditions and without undermining the legitimacy or integrity of adjudication.²⁹ How can these beliefs

26. See *infra* subsection III(B)(3)(d). For example, Professors Daphna Kapeliuk and Alon Klement, in what is perhaps the most rigorous economic analysis of party rulemaking in the existing literature, assume that any threat to judicial legitimacy can be cashed out in the form of reputation costs and that judges can trade these costs off against the benefits of enforcing *ex ante* contracts in individual cases. See Kapeliuk & Klement, *Contractualizing Procedure*, *supra* note 21, at 44 (explaining how party rulemaking can have an adverse reputational effect on the court). Professors Davis and Hershkoff take a somewhat broader view, Davis & Hershkoff, *supra* note 22, at 541–63, but they do not analyze adjudicative legitimacy rigorously in the way I do here.

27. See ERIN A. O'HARA & LARRY E. RIBSTEIN, *THE LAW MARKET* 56–60 (2009) (describing the scope of power parties have to choose substantive law).

28. See generally Gary Lawson, *Stipulating the Law*, 109 MICH. L. REV. 1191 (2011) (noting that courts have allowed parties to stipulate to the answer to a specific legal question that was then used as a basis for deciding the case, and evaluating this practice).

29. See, e.g., Janet Cooper Alexander, *Do the Merits Matter? A Study of Settlements in Securities Class Actions*, 43 STAN. L. REV. 497, 498 (1991) (“Most cases are resolved through settlement. Indeed, federal policy (and probably that of most states) favors settlement over trial, to such an extent that it is a familiar axiom that a bad settlement is almost always better than a good trial.” (internal quotation marks omitted)); Samuel R. Gross & Kent D. Syverud, *Getting to No: A Study of Settlement Negotiations and the Selection of Cases for Trial*, 90 MICH. L. REV. 319, 320 (1991) (“With some notable exceptions, lawyers, judges, and commentators agree that pretrial settlement is almost always cheaper, faster, and better than trial.”). But see Marc Galanter, *The Quality of Settlements*, 1988 J. DISP. RESOL. 55, 56–59, 82 (chronicling an “anthology of assertions

be squared with an aversion to agreements that shape procedure in other ways?

Still, there is something to the concerns of the party-rulemaking critics. Imagine that two parties enter into a contract for a pleading standard stricter than plausibility as well as a stricter version of Rule 11 sanctions. Imagine that they also agree to bar joinder, class actions, and pleading amendments, and adopt a very narrow summary judgment rule, all in order to cut costs and expedite the pretrial stage. Suppose too that their contract limits discovery, provides for only truncated evidentiary presentations, and bars appeals, all in order to achieve a final result more quickly and efficiently in the event settlement fails. And suppose that it bars motions for judgment as a matter of law (directed verdict or JNOV) and provides that the case should be decided by the judge with the assistance of a panel of experts in the field—like the merchant juries that Karl Llewellyn endorsed for commercial cases³⁰—with an instruction that the judge and the experts apply only certain specified legal rules or simply follow their own sense of fairness and practicality.

My guess is that many, maybe all, proponents of broad party rulemaking would find this degree of procedural manipulation disturbing for adjudication, even though something like it can occur in arbitration.³¹ Moreover, they would still find it disturbing even if there were no apparent reason to think that the modifications would increase the public costs of litigation or negatively affect other litigants. It is hard to dispel the nagging sense that something is wrong, that parties should not be allowed to change adjudication to this extent. The question is why.

This Article sets out to answer that question. Part II frames the problem more precisely. It describes in broader context and defines with greater care the distinct ways in which parties can choose procedure. With this background in place, Part III critically examines the conventional arguments for and against party rulemaking. It evaluates those arguments from utilitarian and rights-based perspectives, finds them unconvincing except in certain special cases, and identifies three limited scenarios where concerns about the costs of party rulemaking justify judicial restraint in enforcing agreements. Part III concludes that if party rulemaking is to be limited or barred in a wider range of cases, it must be because giving parties control over procedure risks jeopardizing the normative legitimacy of adjudication.

Part IV then attempts to flesh out a coherent and convincing version of the legitimacy critique. First, it identifies core elements of adjudication that are critical to its institutional legitimacy. Those elements have to do with

about the estimable qualities of settlements drawn from judges and legal scholars” but concluding that “[s]ettlements are not intrinsically good or bad, anymore than adjudication is good or bad”).

30. Zipporah Batshaw Wiseman, *The Limits of Vision: Karl Llewellyn and the Merchant Rules*, 100 HARV. L. REV. 465, 512–13 (1987).

31. There could be Article III problems with the parties' chosen decision protocol, especially if the contract gives the experts any decision-making power. But the negative reaction I imagine in the text is not likely to depend on a proper interpretation of Article III.

adjudication's commitment to a distinctive method of reasoning. Part IV then argues that the most troubling examples of party rulemaking are those that tinker with procedures that frame, guide, or incentivize this reasoning process. This analysis identifies some party-made rules as problematic on legitimacy grounds—namely, those that seriously interfere with the adjudicative reasoning process—but it also leaves considerable room for party rulemaking. Part IV closes by examining reasons why parties should not be allowed to alter procedures that are central to the reasoning process even when those parties are willing to assume the risks. Part V concludes.

II. Framing the Problem

The following discussion frames the problem of party rulemaking by placing it in a broader setting and summarizing the current law.

A. *The Problem Clarified and in Context*

To see how party choice relates to other methods of making procedure, it is helpful to organize the various sources of procedural law in the form of a nested hierarchy starting with the Constitution at the top.³² Each level of the hierarchy must be consistent with all the levels higher up. For example, federal procedural statutes lie just below the Constitution and must conform to constitutional requirements. The Federal Rules of Civil Procedure are at the next level; they are supposed to conform to the Constitution and federal statutes.³³ One step further down lie the local district rules, which are supposed to be consistent with the Federal Rules, federal statutes, and the Constitution.³⁴ Common law procedure lies below that, followed by a trial judge's discretionary case-management orders.³⁵

Ordinarily we think of party choice as lying at the very bottom of this hierarchy. Parties fashion procedure only to the extent that the rules and

32. There are a few exceptions that do not fit neatly into the hierarchy, but the most important sources are included.

33. An exception is the Supersession Clause, which provides that the Federal Rules of Civil Procedure supersede any inconsistent statute adopted before their enactment. Rules Enabling Act § 401(a), 28 U.S.C. § 2072(b) (2006). This clause was added to the original Federal Rules of Civil Procedure “to obviate a need for explicit repeal of any provisions of the Judicial Code which might later be found to be in conflict with the new rules.” Paul D. Carrington, “*Substance*” and “*Procedure*” in the Rules Enabling Act, 1989 DUKE L.J. 281, 322. The Supersession Clause, however, has rarely been applied. *Id.*

34. Local districts were criticized in the 1980s for adopting local rules that went well beyond and sometimes conflicted with the Federal Rules of Civil Procedure. See, e.g., Stephen N. Subrin, *Federal Rules, Local Rules, and State Rules: Uniformity, Divergence, and Emerging Procedural Patterns*, 137 U. PA. L. REV. 1999, 2019–20 (1989) (describing multiple studies done in the 1980s by various groups that demonstrated “[t]he proliferation of local rules”).

35. Trial judges also have some inherent power, but this power ordinarily must be exercised in a manner consistent with rules higher up in the hierarchy. See *Chambers v. NASCO, Inc.*, 501 U.S. 32, 44–47 (1991) (outlining the various “facets to a federal court’s inherent power,” while noting that “the exercise of the inherent power of lower federal courts can be limited by statute and rule”).

norms higher up allow it. This view is correct for party choice of actions before or during litigation, but it is not correct for party choice of general procedural rules. Empowering rule choice inverts the usual order, for it allows parties to alter rules higher up in the hierarchy. This is one of the reasons that party choice of general rules seems problematic.³⁶

The previous point assumes that the distinction between a general rule and an action is intuitively obvious. The distinction is clear enough in the ordinary case when parties are already engaged in strategic litigation. The general procedural rules define the set of permissible actions, and lawyers for the parties choose actions from the permissible set.³⁷ The lawyers make their choices strategically, by which I mean that each chooses her action anticipating and responding to the choices that others make (or will make), all the while aware that others will anticipate and respond to what she does. Thus, assuming rationality, each lawyer's choice is constrained to some extent by what she believes other lawyers will do.

The distinction between rules and actions also applies when parties make choices in advance of the time when the rule would be implemented or the action taken. These choices require cooperation and are usually implemented through some kind of agreement. While the usual case involves an agreement entered into prior to a dispute, there is nothing stopping parties from agreeing after a dispute arises or while a lawsuit is ongoing. As we shall see in subpart II(B) below, most of the cases involve contracts in which the parties commit to particular actions that the general procedural rules would permit them to take during litigation. For example, parties can waive jury trial in the course of litigation, and they can also agree to waive jury trial by contract in advance. Similarly, they can stipulate to limits on discovery during litigation, and they can also agree to discovery limits in advance. In a sense, this type of agreement fashions a rule for the future suit, but the rule is a very simple one, merely a directive requiring implementation of the specific action. Most importantly for our purposes, the simple rule does not alter the general rules that would otherwise apply; instead, it just moves to an earlier point in time a choice that the general rules allow parties to make later on.

36. This Article assumes specific enforcement of agreements to choose procedural rules or actions in advance. It is only through specific enforcement that parties can directly shape procedure for their cases.

37. An analogy to game theory might be helpful. See generally ERIC RASMUSEN, GAMES AND INFORMATION: AN INTRODUCTION TO GAME THEORY 12–14 (3d ed. 2001) (explaining that in a game, the rules of the game identify the players, the choices available at each node, the information structure, and the payoffs). Procedural rules create the rules of the litigation game. They identify those who can be parties to litigation (the players of the game), the stages at which choices can be made, and the actions available to each party at each stage. Litigation payoffs are defined by the substantive and procedural law (procedural law insofar as it affects the magnitude of litigation costs), and the information structure is a result of all the numerous factors that influence access to information, including the procedural rules.

Parties can choose not only actions but also general rules. One way they do so today is by choosing the forum. If a suit can be filed in Court *A* or Court *B*, and *A* would apply different procedural rules than *B*, the plaintiff can choose the rules applicable to the case by choosing the appropriate court. This can happen in federal courts when local rules or interpretations of the Federal Rules of Civil Procedure vary across different districts, but it is much more likely to happen in state courts. Similarly, if the plaintiff files in state court and the case can be removed to federal court—and the rules differ between state and federal court—the defendant can choose the rules by removing the case. Choice of rules through choice of forum can be accomplished not only through unilateral strategic choice, as in these examples, but also through mutual agreement on choice of forum in advance of suit. Like choice of action in the previous paragraph, however, choosing rules by choosing a forum does not place party choice in conflict with what the chosen court would otherwise have applied. It merely selects among different sets of official rules.

Choice of rules need not be so limited. In theory at least, parties might agree on their own procedural rules rather than select among different official rule systems. The rules they choose could be different than—indeed, inconsistent with—the official rules that would otherwise apply. This is the type of party rulemaking that I focus on in this Article.³⁸ As we shall see in subpart II(B) below, there is relatively little support for this form of rulemaking in the current case law. Moreover, it raises serious concerns, especially when the chosen rule differs markedly from the default rule that would otherwise apply.

We can organize the different types of party-made procedure into a two-by-two matrix with the method of choice on one axis and the object of choice on the other:

Table 1. Party-Made Procedure

	Choice of Actions	Choice of Rules
No Agreement	Type I	Type IV
Agreement	Type II	Type III

Type I is the usual sort of party-made procedure associated with ordinary strategic interaction in the course of litigation. As mentioned above, Type IV procedure is rather foreign to our current litigation system and

38. This type of rulemaking also includes choice of actions when the chosen action conflicts with a general rule that would otherwise apply at the later litigation stage. In this situation, the parties alter the general rule indirectly by creating an exception to it. In other words, they create a new rule consisting of the existing rule plus the exception.

difficult to envision—aside from the simple case of a party choosing procedural rules by selecting the forum. This leaves Type II and Type III. Both are methods of making rules. Type II is the most common form of party rulemaking. Parties commit in advance to a specific action that one or both of them could have taken during litigation. Type III involves contractual choice of general rules rather than choice of specific actions. It includes agreements that choose general rules different from the official rules that would otherwise apply.³⁹ While we will consider Type II rulemaking, our principal interest is in Type III, especially Type III rulemaking that alters otherwise applicable rules.

Before proceeding, a word of caution is in order. The difference between Type II and Type III rulemaking is not perfectly sharp. The distinction is analytically useful; it helps to highlight the difference in normative stakes. But as the following subpart illustrates, some examples of party rulemaking are not easy to classify in one category or the other.

So far we have been assuming a bilateral agreement between two parties. However, the trial judge can be involved in the agreement as well. Trial judges have broad case-management powers and often “negotiate” with the parties to establish discovery limits, trial deadlines, and the like. While the arrangement is not exactly contractual—the judge has the final word and in theory, at least, can impose procedures unilaterally—the result can resemble a three-party contract when the judge has something to lose by thwarting the parties’ preferences (such as overseeing a more protracted litigation). Moreover, even when parties agree in advance, the trial judge often has power to check the agreement *ex post*.⁴⁰ This can set up a bargaining game at the time of implementation, and the possibility of later renegotiation with the judge is likely to affect the parties’ contracting incentives *ex ante*.

One final point warrants clarification. Type II and Type III rulemaking can take place at any time before an action or rule is implemented. This includes during the course of litigation as well as before a lawsuit arises. Some commentators assume that cooperation is nearly impossible during litigation, but they tend to exaggerate the difference between *ex ante* and *ex post*.⁴¹

To be sure, agreement can be more difficult to reach after a dispute materializes. In the *ex ante* world, both parties are uncertain about the kind of suit they will face and what position (plaintiff or defendant) they will occupy. As a result, they take expectations over all future states of the world, and if the benefits and burdens cancel out in expectation, the parties have

39. It also includes situations where the parties choose the rules indirectly by choosing the forum.

40. See *infra* subpart II(B).

41. See, e.g., Kapeliuk & Klement, Contractualizing Procedure, *supra* note 21, at 13–14, 16–17 (observing the divergence of parties’ predispute and postdispute interests and noting that “after the parties have embarked on the litigation war-game, cooperation is difficult to achieve”).

incentives to make a procedural choice that reduces joint litigation costs.⁴² For example, two contracting parties might agree on limited discovery if they both believe that they are just as likely to be a defendant who is benefitted by limited discovery as they are to be a plaintiff who is burdened by it. After a dispute materializes, however, the defendant knows he will benefit from limited discovery and the plaintiff knows she will be burdened. So it is more difficult to find a bargaining range.

However, the informational differences between *ex ante* and *ex post* are not as stark as some commentators assume. The *ex ante* world is not a Rawlsian veil of ignorance, and the *ex post* world is not a state of complete information.⁴³ Parties in the *ex ante* position are often able to predict salient characteristics of future litigation. Consider a contract between a seller and consumer for the sale of goods. The seller knows with reasonable confidence that he is likely to be the defendant and the consumer is likely to be the plaintiff in any future litigation over the quality of the goods. The same is true for contracts between firms. For example, contracts between building owners and building contractors frequently give rise to lawsuits in which the owner sues the contractor for building defects or the contractor sues the owner for nonpayment.

Moreover, uncertainty exists in the *ex post* world as well as the *ex ante*. For example, parties early in a lawsuit are likely to be uncertain about the fruits of discovery or the prospect of success on a later summary judgment motion. They must make litigation decisions by taking expectations over all the future possibilities, and these expectations permit some canceling of benefits and burdens. Given uncertainty about summary judgment, for example, parties might be able to agree to a modified summary judgment rule or even no summary judgment at all. Doing so could make both parties better off in expectation by reducing litigation costs and securing a speedier trial.⁴⁴

It is true that there is more room for making side payments *ex ante*.⁴⁵ But side payments are also possible during litigation. For example, one party might agree to pay the other a portion of the additional costs incurred by a new procedure or agree to adjust the settlement or trial outcome. Also,

42. See Bruce L. Hay, *Procedural Justice—Ex Ante vs. Ex Post*, 44 UCLA L. REV. 1803, 1828–39 (1997) (describing the difference between *ex ante* and *ex post* perspectives when information differs).

43. See Robert G. Bone, *Agreeing to Fair Process: The Problem with Contractarian Theories of Procedural Fairness*, 83 B.U. L. REV. 485, 526–29 (2003) [hereinafter Bone, *Agreeing to Fair Process*] (criticizing the extreme assumptions about information access that underlie typical *ex ante* arguments).

44. Cf., e.g., ROBERT G. BONE, CIVIL PROCEDURE: THE ECONOMICS OF CIVIL PROCEDURE 85–91 (2003) [hereinafter BONE, CIVIL PROCEDURE] (describing how factors such as litigation costs, divergent expectations of success, and mutual optimism influence the “settlement surplus”).

45. See Kapeliuk & Klement, *Contractualizing Procedure*, *supra* note 21, at 17 (evaluating the likelihood of cooperation, willingness to make procedural arrangements, and ability to make transfer payments at the pre-dispute-contracting stage compared to the stage after a dispute has arisen).

parties could trade benefits under one rule or action for burdens under another. A plaintiff and defendant, for instance, might trade a broader discovery scope for a staged discovery process and an early summary judgment evaluation of the case. The plaintiff would benefit from the broader scope, and the defendant would benefit from an earlier summary judgment check of the merits.

To sum up, the parties might make procedural rules for their own cases through agreement by committing to a specific action (Type II) or a general rule (Type III) in advance of when the action or rule must be implemented. But the judge can force renegotiation of the agreement when she has the power to review it *ex post*. Committing to an action in advance can make both parties better off in expectation, but it also can raise concerns about the bargaining conditions that give rise to the agreement. Committing to a general rule raises additional and possibly more serious concerns when the chosen rule deviates sharply from the background rules in place.

B. Current Law on Party Rulemaking

Parties have some latitude to choose procedural rules under current law. Most of the cases that I reviewed involve Type II rulemaking, with parties committing to a procedural action in advance as part of a commercial or consumer contract. There are not many examples of Type III. The most common example involves parties choosing a rule system by choosing a forum with the desired rules, and the following discussion examines contractual forum choice. But there are very few examples of parties designing their own rules at odds with the official rules that would otherwise apply.

As we shall see, the cases usually involve parties reaching agreement *before* a dispute arises. I found very few examples of agreements entered into after filing, other than the usual stipulations for additional time and the like. One possible reason is that procedural options after filing are treated as bargaining chips in settlement negotiation, so any agreement takes the form of a settlement ending the suit.

The following discussion briefly reviews examples of party rulemaking in the context of forum choice, pretrial procedure, trial procedure, remedies, and appeals.⁴⁶ While the main focus of this Article is on federal civil

46. The cognovit note is also worth mentioning in this connection. A debtor who executes a cognovit note waives all defenses in case of default and agrees in advance that the creditor can obtain a judgment without any notice or hearing at all. *D.H. Overmyer Co. v. Frick Co.*, 405 U.S. 174, 176 (1972). These agreements, which waive all due process rights, are enforceable provided they are made "voluntarily, intelligently, and knowingly." *Id.* at 187. The cognovit, however, is an "ancient legal device" and firmly entrenched as a longstanding method to regulate credit risk. *Id.* at 176-78. As such, it is an exceptional form of waiver. It would be a mistake to infer from the existence of cognovit notes a more general power to waive specific procedural rules in all cases or a power to engage in Type III contractual modification of existing rules. *But see* Noyes, *supra* note 21, at 603 (suggesting this argument but not actually making it clearly).

procedure, its analysis applies more generally, and the following discussion reviews both federal and state examples.

1. *Forum Choice*.—Courts routinely enforce agreements that select a forum. One example is a contract to arbitrate, which removes a dispute entirely from adjudication.⁴⁷ Another example, more centrally relevant to the concerns of this Article, is a forum-selection clause that chooses a particular court. There are two types of forum-selection clauses.⁴⁸ A “consent-to-jurisdiction clause” is strictly permissive; it commits the parties to submit to a particular court’s jurisdiction and venue if the suit is filed there, but it does not exclude other possible locations.⁴⁹ A mandatory clause—also called a “prorogation clause”—commits the parties to a single forum exclusively.⁵⁰ Judges routinely enforce both types of agreements provided they are reasonable.⁵¹ And reasonableness is liberally construed in favor of enforcement.⁵²

47. Arbitration agreements are liberally enforced. Indeed, the Supreme Court has construed the Federal Arbitration Act, 9 U.S.C. §§ 1–307 (2006), to embody a strong federal policy in favor of arbitration. *See, e.g.,* *Moses H. Cone Mem’l Hosp. v. Mercury Constr. Corp.*, 460 U.S. 1, 24 (1983) (“Section 2 [of the Federal Arbitration Act] is a congressional declaration of a liberal federal policy favoring arbitration agreements . . .”); Aaron-Andrew P. Bruhl, *The Unconscionability Game: Strategic Judging and the Evolution of Federal Arbitration Law*, 83 N.Y.U. L. REV. 1420, 1426–32 (2008) (describing the expansion of federal arbitration law as reflecting a strong federal pro-arbitration policy).

48. *See generally* 14D WRIGHT, MILLER & COOPER, *supra* note 22, § 3803.1, at 123–31 & nn.88–89 (explaining the differences between “mandatory” and “permissive” forum-selection clauses); Robert C. Casad, *Territorial Jurisdiction: Jurisdiction over Persons and Property*, in 16 JAMES WM. MOORE, MOORE’S FEDERAL PRACTICE § 108.53, at 108-95 to -103 (3d ed. 2011) (same).

49. Casad, *supra* note 48, § 108.53[4], at 108-98. Parties sometimes couple consent-to-jurisdiction clauses with consent to a specified mode of service. *See, e.g.,* *Nat’l Equip. Rental, Ltd. v. Szukhent*, 375 U.S. 311, 315–16 (1964) (allowing contracting parties to “agree in advance to submit to the jurisdiction of a given court” and modify or waive service operation); *Beautytuft, Inc. v. Factory Ins. Ass’n*, 48 F.R.D. 15, 27 (E.D. Tenn. 1969) (considering it “well settled” that parties may agree to personal jurisdiction and modifications to service by contract).

50. Casad, *supra* note 48, § 108.53[5], at 108-100.

51. *See, e.g.,* *Carnival Cruise Lines, Inc. v. Shute*, 499 U.S. 585, 593 (1991) (permitting “reasonable forum clause[s]” in form contracts); *The Bremen v. Zapata Off-Shore Co.*, 407 U.S. 1, 15 (1972) (advocating a default rule of specific enforcement of forum-selection clauses unless a contesting party can show the specific clause to be “unreasonable and unjust”); *see also* 14D WRIGHT, MILLER & COOPER, *supra* note 22, § 3803.1, at 51–52 (“[T]he common understanding is that these provisions are prima facie valid and should be enforced unless enforcement is shown to be unreasonable under the circumstances of the particular contract.”). This is especially true of commercial contracts, but as the *Carnival Cruise Lines* decision makes clear, a liberal approach applies to consumer contracts too. As long as the consumer has notice of the forum-selection clause and there is no bad faith or fraud, the clause will be enforced even if it is part of a contract of adhesion. *See Carnival Cruise Lines*, 499 U.S. at 595 (enforcing a forum-selection clause against cruise line customers because there was no evidence of bad faith or fraud on the part of the company).

This was not always so. Historically, courts were hostile to mandatory forum-selection clauses. They developed the so-called ouster doctrine, which rejected pre-suit agreements that ousted a court’s jurisdiction. *See, e.g.,* *Ins. Co. v. Morse*, 87 U.S. (20 Wall.) 445, 451 (1874) (“[A]ny citizen

A permissive consent-to-jurisdiction clause is an example of Type II rulemaking. The parties in effect commit to a forum choice that they could have chosen noncooperatively at the time of suit. Noncooperative choice would have been possible if the defendant had consented to personal jurisdiction and venue or waived any objections.⁵³ Thus, a consent-to-jurisdiction clause merely moves a future defendant's forum-specific consent or waiver to an earlier point in time.⁵⁴ This does not necessarily mean enforcement of the clause is unproblematic, but it does mean that any problems must flow from the specific facts of the case, such as the time shift or the contracting circumstances.⁵⁵

At first glance, the mandatory prorogation clause might seem more like Type III rulemaking than Type II. It locks both sides into a single forum, a result that could not be achieved without agreement. Suppose *A* and *B* agree to a clause that identifies the Southern District of New York as the exclusive forum for future federal litigation. If *A* files in the District of Massachusetts, *B* can move to dismiss the suit, relying on the forum-selection clause.⁵⁶ In the absence of such a clause, *B* would have to bring a motion to transfer under § 1404(a),⁵⁷ and the judge would apply a balancing test that gives great weight to the plaintiff's forum choice.⁵⁸ Thus, *B* can effectively lock *A* into the Southern District of New York with a mandatory forum-selection clause, but it cannot achieve the same result noncooperatively by using § 1404(a).⁵⁹

may no doubt waive the rights to which he may be entitled. He cannot, however, bind himself in advance by an agreement . . . to forfeit his rights at all times and on all occasions, whenever the case may be presented.”).

52. See 14D WRIGHT, MILLER & COOPER, *supra* note 22, § 3803.1, at 52–58 (“Generally, a strongly hospitable judicial attitude toward these clauses prevails.”).

53. FED. R. CIV. P. 12(h)(1). By contrast, the parties can neither commit to nor choose noncooperatively a federal court when there is no federal subject matter jurisdiction over the suit.

54. See, e.g., *Nw. Nat'l Ins. Co. v. Donovan*, 916 F.2d 372, 375 (7th Cir. 1990) (making this point in the course of enforcing a forum-selection clause).

55. Or perhaps a judge might decide that for some reason the particular choice is better made noncooperatively than through bargaining.

56. See 14D WRIGHT, MILLER & COOPER, *supra* note 22, § 3803.1, at 79–80, 104–05 (noting that most federal courts dismiss if the forum-selection clause is reasonable, but criticizing that approach).

57. 28 U.S.C. § 1404(a) (2006).

58. See *Piper Aircraft Co. v. Reyno*, 454 U.S. 235, 255 (1981) (acknowledging that “there is ordinarily a strong presumption in favor of the plaintiff's choice of forum” in § 1404(a) transfer determinations).

59. It is worth noting that the authors of the Wright, Miller, and Cooper treatise see the matter differently. They read the Supreme Court's decision in *Stewart Organization, Inc. v. Ricoh Corp.*, 487 U.S. 22 (1988), to hold that in diversity and federal question cases, where the contractual forum has personal jurisdiction and venue, a motion to dismiss should be treated as a motion to transfer under § 1404(a) and the forum-selection clause should be a “significant factor that figures centrally” in the § 1404(a) analysis but should not be determinative. 14D WRIGHT, MILLER & COOPER, *supra* note 22, § 3803.1, at 75–79 (quoting *Stewart*, 487 U.S. at 29). If this approach were followed, then the only difference between a scenario with a mandatory forum-selection clause choosing a federal court and one without such a clause would be the additional weight given to the contractual choice in the § 1404(a) analysis. The contractually chosen forum would still have to qualify independently

However, this distinction does not make mandatory clauses a form of Type III rulemaking.⁶⁰ A mandatory clause, like a permissive one, is simply a reflection of party choices at an earlier stage. Parties can settle on a forum if the plaintiff files there and the defendant consents to the forum or waives objections. A mandatory clause merely makes that consent bilateral and moves it up in time to a point before the lawsuit materializes. Moreover, a mandatory clause does not create a general rule that then guides or constrains strategic choice at the time of litigation; instead, it names a specific forum. Therefore, even a mandatory clause is better seen as an example of Type II rulemaking than of Type III.

2. *Pretrial*.—There are numerous pretrial matters that might be the subject of agreement. In theory, parties could agree to a different pleading standard, different timing and other conditions for raising defenses, limitations on joinder of additional parties, limitations on discovery, different summary judgment standards, shortened time for the pretrial stage, and so on. It turns out, however, that the cases cover a much more limited range. The following discussion examines the only two pretrial matters that receive substantial attention in the case law and scholarly commentary: agreements concerning discovery and agreements modifying the applicable statute of limitations.

a. *Discovery*.—Rule 29 of the Federal Rules of Civil Procedure gives parties wide latitude, with court oversight, to stipulate to discovery procedures different from those provided in the Federal Rules.⁶¹ Moreover, other features of the Federal Rules give parties additional freedom to shape discovery.⁶²

for personal jurisdiction and venue. The balancing test would then “place[] a significant limit on the extent to which private parties, even through a freely bargained contract, are permitted to refashion the laws of federal procedure.” *Id.* § 3803.1, at 76. However, the authors concede that their approach is not widely followed. *See id.* § 3803.1, at 79 (noting that, given the “confusion surrounding forum-selection clause analysis, many lower federal courts have failed to distinguish between the approach taken in *Carnival Cruise* and that taken in *Stewart*”).

60. Even so, it might support closer scrutiny of mandatory clauses than permissive ones.

61. *See* FED. R. CIV. P. 29 (providing that “[u]nless the court orders otherwise, the parties may stipulate” that certain aspects of depositions will be conducted in particular ways and that “other procedures governing or limiting discovery be modified”); Patrick E. Higginbotham, *Duty to Disclose: General Provisions Governing Discovery*, in 6 JAMES WM. MOORE, MOORE’S FEDERAL PRACTICE § 26.04[1], at 26-35 (3d ed. 2008) (“Parties may mutually stipulate to use procedures for discovery that vary from the rules . . .”); 8A CHARLES ALAN WRIGHT, ARTHUR R. MILLER & RICHARD L. MARCUS, FEDERAL PRACTICE AND PROCEDURE §§ 2091–2092 (3d ed. 2010) (delineating the parameters of the ability of litigants to stipulate discovery procedure).

62. This freedom is subject to trial judge oversight and, in some instances, approval. *See, e.g.*, FED. R. CIV. P. 26(a)(1) (initial disclosure); FED. R. CIV. P. 26(f) (discovery conference); FED. R. CIV. P. 33(a)(1) (interrogatories); FED. R. CIV. P. 33(b)(2) (same); FED. R. CIV. P. 34(b)(2)(A) (document production); FED. R. CIV. P. 34(b)(2)(E) (same); FED. R. CIV. P. 36(a)(3) (admissions).

However, there are very few examples in the case law of parties entering into formal agreements committing to discovery limits or modified discovery procedures.⁶³ This is not surprising for the post-filing stage. The Federal Rules of Civil Procedure establish a formal process for planning discovery.⁶⁴ Rule 26 requires the parties to meet and confer over a discovery plan shortly after a lawsuit is filed and to submit a report to the court summarizing their efforts.⁶⁵ The trial judge then holds a scheduling conference, consults with the parties' attorneys, and enters an order outlining a plan for discovery.⁶⁶ It would not be surprising if parties relied on this process rather than formal discovery agreements, especially as the judge would likely review the terms of their agreement in any event.

It is a bit more puzzling that there is only meager evidence in the case law of discovery agreements before litigation. The conventional wisdom repeated in treatises and commentaries is that parties have broad power to contract for discovery limits *ex ante*,⁶⁷ but these claims rely on flimsy case law support. Most of the secondary sources rely on a single case, *Elliott-McGowan Productions v. Republic Productions, Inc.*,⁶⁸ in which the district judge enforced a pre-suit agreement placing limits on who could inspect documents and when notice of inspection had to be given.⁶⁹ Perhaps the availability of arbitration explains the paucity of cases involving *ex ante*

63. However, the advisory committee's note to the 1970 amendment of Rule 29 observed that "[i]t is common practice for parties to agree on" variations of the "procedures by which methods of discovery other than depositions are governed" and that Rule 29 was meant to recognize such agreements and give them formal effect. FED. R. CIV. P. 29 advisory committee's note (1970).

64. This formal process has been in full effect only since 1993. See Steven S. Gensler, *Judicial Case Management: Caught in the Crossfire*, 60 DUKE L.J. 669, 676-80 (2010) (describing the evolution of the case-management process and, in particular, the 1993 amendment to Rule 26(f), mandating discovery planning conferences for the first time). One might expect to find earlier cases dealing with discovery agreements, but I was not able to locate any.

65. FED. R. CIV. P. 26(f).

66. FED. R. CIV. P. 16(a)-(b).

67. See, e.g., 11 MATTHEW BENDER & CO., BENDER'S FORMS OF DISCOVERY TREATISE § 1.04 (2011) ("[P]rovided there is no inequality of bargaining power, [parties] may also contractually limit discovery with respect to future litigation."); Higginbotham, *supra* note 61, § 26.04[1], at 26-35 (stating the same); Noyes, *supra* note 21, at 609-10 (assuming that parties can contract for discovery limits *ex ante*); Thornburg, *supra* note 22, at 202 (stating that parties can limit discovery by contract); Note, *Discovery Abuse Under the Federal Rules: Causes and Cures*, 92 YALE L.J. 352, 364 (1982) (assuming that parties can "make judicially enforceable private agreements concerning discovery"). But see 8 WRIGHT, MILLER & MARCUS, *supra* note 22, § 2005, at 52 (criticizing this practice as "hardly an appropriate means for disregarding rules of court devised to serve the public interest in bringing out all the facts prior to trial"). The benefits are well-known. Parties can escape the Pareto-inferior equilibrium of the prisoner's dilemma and achieve gains through cooperation, such as reduced litigation costs. Kapeliuk & Klement, *Contractualizing Procedure*, *supra* note 21, at 16.

68. 145 F. Supp. 48 (S.D.N.Y. 1956).

69. *Id.* at 50; see also 8 WRIGHT, MILLER & MARCUS, *supra* note 22, § 2005, at 52 (criticizing *Elliott-McGowan*); *Developments in the Law—Discovery*, 74 HARV. L. REV. 940, 979 (1961) (same).

agreements: contracting parties might just switch to arbitration when they are concerned about excessive or abusive discovery in court.

Among the few examples of discovery agreements, some are instances of Type II rulemaking while others are more properly treated as Type III. For example, an agreement not to take any depositions or to use a particular method to review documents in response to a document request is an example of Type II rulemaking. In these cases, the parties are merely committing to choices they could make noncooperatively at the point of implementation: each party can choose not to take any depositions and one party can use a special document-review method, assuming no objection from the other side. An agreement to set a limit on depositions that leaves room for choice seems more like Type III rulemaking since it alters the limits already set by the Federal Rules. But this is not a particularly dramatic example since Rule 29 already allows parties to modify discovery rules by stipulation.⁷⁰ In fact, in view of Rule 29, discovery agreements are probably more akin to Type II rulemaking than Type III: parties achieve by agreement before filing what the existing rules give them freedom to accomplish after filing.⁷¹

b. Statute of Limitations.—The law is relatively clear about contractual modification of statutes of limitations.⁷² Parties are free to shorten an applicable statute by agreement as long as the shorter period is reasonable.⁷³ However, parties have much less freedom to lengthen a statute of limitations or waive the defense in advance of suit.⁷⁴ The reason courts give for the difference has to do with the legislative policy of avoiding stale

70. See *supra* note 61 and accompanying text.

71. A clear example of Type III rulemaking would be an agreement to alter Rule 26's provisions on scope, expert discovery, protective orders, and the like. Yet it is not at all clear that judges would be willing to enforce that type of agreement.

72. One survey of contracts filed with the SEC revealed that contractual modification of statutes of limitations was rather common. Kapeliuk & Klement, *Contractualizing Procedure*, *supra* note 21, at 6–7. In addition, parties sometimes waive other related defenses such as estoppel and laches. *Id.* at 7 & n.25.

73. See, e.g., *Gifford v. Travelers Protective Ass'n of Am.*, 153 F.2d 209, 211 (9th Cir. 1946) (confirming the ability of the parties to shorten the limitations period but warning that “the interval may not be unreasonable”); *Shaw v. Aetna Life Ins. Co.*, 395 A.2d 384, 386 (Del. Super. Ct. 1978) (explaining that an expressly abbreviated limitations period is permitted because it “hastens the enforcement and complements the policy behind the statute of limitations”); *Keiting v. Skauge*, 543 N.W.2d 565, 567 (Wis. Ct. App. 1995) (“[T]he right to contract for a shortened limitations period is . . . supported by public policy.”); 15 GRACE MCLANE GIESEL, CORBIN ON CONTRACTS: CONTRACTS CONTRARY TO PUBLIC POLICY § 83.8, at 287 (Joseph M. Perillo ed., rev. ed. 2003) (reporting that contractually shortening the limitations period “is not contrary to public policy but rather assists the public policy behind statutes of limitations: preventing stale claims”).

74. See, e.g., *Shaw*, 395 A.2d at 386–87 (concluding that a contract that extends the limitations period beyond the statute violates public policy and impermissibly circumvents the law); 15 GIESEL, *supra* note 73, § 83.8, at 289–90 (noting that “courts do not enforce parties’ agreements to lengthen the limitations period” and that “general agreements in advance to waive or not plead the applicable statute of limitations are void”). But see *Collins v. Envtl. Sys. Co.*, 3 F.3d 238, 241 (8th Cir. 1993) (holding that Minnesota law allows the parties to agree to waive the statute of limitations but only for a reasonable period of time).

claims: lengthening a statute of limitations increases the risk of stale claims, while shortening it does not.⁷⁵

An agreement to lengthen a statute of limitations is a form of Type II rulemaking. The statute of limitations is a waivable defense,⁷⁶ so the same result could be achieved noncooperatively by waiver after the plaintiff files. Thus, an agreement to lengthen the statute of limitations is functionally equivalent to an agreement to waive it conditioned on suit being filed before the later date.⁷⁷

It is more difficult to classify an agreement to shorten a statute of limitations. If it is treated simply as a promise to file suit at an earlier time, it is a form of Type II rulemaking. Filing earlier is, of course, something the plaintiff can do noncooperatively. However, this interpretation misses the fact that the parties have not merely arranged for a filing date; rather, they have agreed to alter a general rule constraining their future strategic choices, and they have done so in a way that is not possible to achieve through provisions like waiver that are built into the rule itself. The plaintiff is still free to choose when to sue, but she is constrained to a shorter period of time. This interpretation of the agreement makes it a form of Type III rulemaking.

3. *Trial.*—The clearest example of party rulemaking aimed at the trial stage is the contractual waiver of jury trial. According to the prevailing view, prelitigation agreements to waive jury trial are enforceable as long as they are made in a knowing, voluntary, and intelligent manner.⁷⁸ This is a

75. See, e.g., *Shaw*, 395 A.2d at 386 (explaining that extending a statute of limitations contradicts the device's public policy purpose of "discourag[ing] the litigation of old or stale demands" (citation omitted)). Even so, parties are relatively free to extend a statute of limitations after the cause of action has accrued, subject once again to a reasonableness constraint. 15 GIESEL, *supra* note 73, § 83.8, at 290. The argument for distinguishing lengthening from shortening is flawed, however, because it ignores one side of the policy balance. Any statute of limitations seeks to strike a balance between allowing adjudication of meritorious claims and screening claims that are weak because of stale evidence. Lengthening the statute of limitations jeopardizes the latter policy. But shortening it jeopardizes the former. After all, it is the risk of screening meritorious suits that prompts courts to impose a reasonableness limitation when shortening the limitations period.

76. See, e.g., *Thompson v. Volini*, 849 S.W.2d 48, 50 (Mo. Ct. App. 1993) (holding that parties can waive the statute of limitations defense after the statutory period has expired); *Duncan v. Lisenby*, 912 S.W.2d 857, 858 (Tex. App.—Houston [14th Dist.] 1995, no writ) ("Parties may agree to waive the statute of limitations before the statutory bar has fallen.").

77. There is another way to see this point. The statute of limitations rule is actually a complex set of requirements and conditions, something like the following:

A lawsuit shall be commenced (where *commence* is defined legally) within *X* years of the accrual of the cause of action (where *accrual* is defined legally), and if the lawsuit is not commenced within the requisite period of time, it shall be dismissed unless the defendant consents to the suit going forward or waives the objection.

When parties agree to lengthen the statute of limitations, it might seem as though they are altering the rule by changing *X*, the number of years, but they are actually invoking the waiver provision in advance.

78. See, e.g., *Great Earth Int'l Franchising Corp. v. Milks Dev.*, 311 F. Supp. 2d 419, 437 (S.D.N.Y. 2004) ("Unquestionably, the Seventh Amendment right to trial by jury may be waived,

straightforward example of Type II rulemaking. Parties to a suit can choose a bench trial noncooperatively if both waive their rights independently.⁷⁹ Thus, contractual waiver simply moves the choice to a point earlier in time.

There is some authority for the proposition that pretrial agreements to waive evidence objections are enforceable.⁸⁰ Many of the cases, however, deal with stipulations during the course of litigation rather than contractual commitments entered into before litigation begins or a claim arises.⁸¹ Moreover, these cases tend to involve stipulations to the admissibility of case-specific evidence or stipulations to waive evidence-specific objections that could be waived at trial, rather than agreements to alter the general evidence rules themselves.⁸² In other words, the cases recognizing party control over admissibility of evidence tend to focus on Type II rulemaking. They enforce agreements that commit parties to actions they could have taken noncooperatively at trial.⁸³

and the waiver is enforceable so long as it is made 'knowingly and voluntarily.'" (citing *Morgan Guar. Trust Co. of N.Y. v. Crane*, 36 F. Supp. 2d 602, 603 (S.D.N.Y. 1999)); *RDO Fin. Servs. Co. v. Powell*, 191 F. Supp. 2d 811, 813 (N.D. Tex. 2002) ("The federal standard for determining the validity of a contractual waiver of the right to a jury trial is . . . whether the waiver was made in a knowing, voluntary, and intelligent manner."); *see also* 9 CHARLES ALAN WRIGHT & ARTHUR R. MILLER, *FEDERAL PRACTICE AND PROCEDURE* § 2321, at 278 (3d ed. 2008) (noting that contractual waivers of jury trial "will be strictly construed").

79. *See* FED. R. CIV. P. 38 (providing that any party who fails to properly serve and file its demand for a jury trial waives that right).

80. *See, e.g.,* *United States v. Mezzanatto*, 513 U.S. 196, 202–03 (1995) (noting that there is a "presumption of waivability"); 21 CHARLES ALAN WRIGHT & KENNETH W. GRAHAM, JR., *FEDERAL PRACTICE AND PROCEDURE* § 5039.5, at 860 (2d ed. 2005) (noting that courts traditionally had resisted allowing bargaining over the rules of evidence but that recently courts have been more receptive); Noyes, *supra* note 21, at 607 ("It is generally acknowledged that *ex ante* contracts to alter the rules of evidence are enforceable."); Taylor & Cliffe, *supra* note 22, at 1086 n.5 ("[A] contract may specify that evidence in the form of hearsay that would otherwise be admissible pursuant to a hearsay exception would be inadmissible unless the declarant were unavailable to testify.").

81. *See, e.g.,* *United States v. Bonnett*, 877 F.2d 1450, 1458–59 (10th Cir. 1989) (enforcing a stipulation made during prosecution of a criminal case); *Tupman Thurlow Co. v. S.S. Cap Castillo*, 490 F.2d 302, 309 (2d Cir. 1974) (enforcing a stipulation made during litigation); *see also Mezzanatto*, 513 U.S. at 203 & n.3 (enforcing a waiver agreement made during plea bargaining in a criminal case but assuming in dictum that "extrajudicial contracts made prior to litigation [might] trigger closer judicial scrutiny").

82. *Compare, e.g., Bonnett*, 877 F.2d at 1458–59 (enforcing a stipulation as to admissibility of certain evidence over a hearsay objection); *Tupman Thurlow*, 490 F.2d at 309 (enforcing a stipulation that certain documents were admissible), *with* *People v. Baynes*, 430 N.E.2d 1070, 1077 (Ill. 1981) (refusing to enforce a stipulation to the admissibility of polygraph evidence in a criminal case when the evidence would otherwise have been excluded as unreliable, noting that "[t]he stipulation attempts to change the legal standard for admissibility [and] [t]his court cannot accept such a result").

83. Many of the sources that report broad party freedom to enter into contracts relating to evidence rely on a 1932 *Harvard Law Review* note. *See, e.g., Mezzanatto*, 513 U.S. at 202 (citing Note, *Contracts to Alter the Rules of Evidence*, 46 HARV. L. REV. 138 (1932), for the proposition that *ex ante* contracts to alter the rules of evidence are generally enforceable); Noyes, *supra* note 21, at 607 (same). That note cites cases where courts have allowed parties to waive or dispense with obstacles to the admission of certain evidence, but it observes that courts have been more wary of contracts or stipulations that do away with the hearsay rule in general. Note, *supra*, at 139–40.

Finally, it is worth mentioning that parties sometimes contract in advance to reassign the burden of proof.⁸⁴ One might view this as contracting for a substantive rule, but burdens also have a strong procedural dimension. Insofar as burden modification counts as altering procedure, it is probably best understood as Type III rulemaking since it is not clear how the parties could effect the same result noncooperatively at trial.

4. *Remedies.*—Parties can control a number of substantive aspects of remedies. For example, parties can contract for liquidated damages within limits,⁸⁵ and they have broad, though not unlimited, power to choose the substantive law applicable at the remedy as well as the liability stage.⁸⁶ Moreover, parties are free to enter into high–low agreements, which impose a floor and a ceiling on the amount of any damages recovery.⁸⁷ And they sometimes agree *ex ante* to forego damages altogether, thereby limiting themselves to injunctive and declaratory relief.⁸⁸

On the more procedural (though still rather substantive) side, parties sometimes commit in advance to the entry of a preliminary injunction or provide contractually for fee shifting.⁸⁹ The former best fits under Type II rulemaking because the defendant could do the same noncooperatively during litigation, and the latter fits within Type III since it alters the generally applicable fee rule that applies at trial.

Interestingly, the note also distinguishes between agreements that allow evidence to be admitted that would otherwise have been excluded and agreements that exclude evidence that would otherwise be admitted. *Id.* at 142–43. The former are mostly unobjectionable, but the latter are an “impediment to ascertaining the facts.” *Id.*

84. See Scott & Triantis, *supra* note 21, at 866–78 (discussing different ways that parties can and do reassign the burden of proof by contract).

85. See, e.g., Ashcraft & Gerel v. Coady, 244 F.3d 948, 954 (D.C. Cir. 2001) (acknowledging the validity of liquidated damages provisions so long as “the amount agreed to by the parties prior to the breach is reasonable”); Wassenaar v. Panos, 331 N.W.2d 357, 361 (Wis. 1983) (“The overall single test of validity is whether the [stipulated-damages] clause is reasonable under the totality of circumstances.”); 15 GIESEL, *supra* note 73, § 83.7, at 286 (“As long as the stipulated damage amount is a true liquidated damage amount and not simply a penalty, the courts have enforced the stipulation.”).

86. See O’HARA & RIBSTEIN, *supra* note 27, at 56–60 (describing the broad power parties have to choose the substantive law they prefer).

87. See Moffitt, *supra* note 21, at 496–97 (describing high–low agreements and noting that courts usually enforce them).

88. I am not aware of any case law reviewing the enforceability of such agreements, but one survey of contracts filed with the SEC found evidence for this type of clause. Kapeliuk & Klement, Contractualizing Procedure, *supra* note 21, at 7–8.

89. Kapeliuk and Klement found examples of these provisions. *Id.* at 8 nn.28–29 (providing examples of agreements with fee and cost shifting); *id.* at 9 n.32 (providing examples of agreements with preliminary injunctions). I am aware of no legal authority reviewing the enforceability of agreements for entry of preliminary injunctions, and the only authority I know addressing contractual fee-shifting provisions is a passage in *Alyeska Pipeline Service Co. v. Wilderness Society*, 421 U.S. 240 (1975), noting generally that exceptions to the American rule can be created by enforceable contract. *Id.* at 257.

5. *Appeals*.—In theory, parties might have an interest in contracting about various aspects of appeal, including timing, scope of review, and so on. The case law, however, is quite thin. Courts enforce agreements to forego appeal as part of a settlement.⁹⁰ And some commentators claim that parties can agree to forego appeal rights in advance.⁹¹

It is unclear why there are not more cases involving agreements to customize appellate procedure. One reason might have to do with restricted appeal rights in arbitration. Perhaps parties who wish to limit appeal do so by choosing arbitration rather than contracting for limitations in adjudication. In any event, contractual waiver of appeal rights is clearly an example of Type II rulemaking.

6. *Summary*.—Three general points emerge from this brief overview. First, it is notable how little case law authority there is on the subject of party rulemaking. Parties can benefit from modifying procedural rules in a variety of different ways both before and after a lawsuit is filed. As Part III explains, doing so can reduce litigation costs, signal good faith, and bond a promise. Yet the cases that I found involve only a limited range of party rulemaking. To be sure, reported cases might not accurately reflect what is actually going on. Nevertheless, one would expect to see precedent supporting enforceability if the practice were widespread. For without formal assurance of legal enforcement, parties would have trouble making credible commitments.

Second, most of the precedents deal with Type II rulemaking. There are very few examples of Type III rulemaking that changes otherwise applicable rules.⁹² Altering the general rules of discovery is a possible example of Type III,⁹³ as are agreements to shorten the statute of limitations, to alter the American rule on attorneys' fees, and possibly to reassign the burden of proof. But this seems to be the extent of it.

Third, a court is not bound to enforce an agreement committing to a procedural choice or creating a procedural rule. The judge can decline enforcement on the usual grounds of validity and enforceability applicable to all contracts.⁹⁴ In addition, it appears that judges have some power to refuse enforcement beyond the general rules of contract law. For example, a contractual waiver of jury trial must be made in a knowing, voluntary, and

90. 15A CHARLES ALAN WRIGHT, ARTHUR R. MILLER & EDWARD H. COOPER, FEDERAL PRACTICE AND PROCEDURE § 3901, at 18–19 (2d ed. 1992).

91. *E.g.*, Noyes, *supra* note 21, at 612–13; Thornburg, *supra* note 22, at 202.

92. *See* Noyes, *supra* note 21, at 608 (“Beyond Rule 38, there have been relatively few judicial decisions requiring a federal court to decide whether to enforce a contractual agreement to alter the Federal Rules of Civil Procedure.”).

93. This might actually be an example of Type II rulemaking insofar as the parties simply implement in advance what Rule 29 of the Federal Rules of Civil Procedure already allows them to do during litigation. *See supra* note 61 and accompanying text.

94. Moreover, some of those grounds, such as the defense of unconscionability, are sufficiently vague and elastic that they can be used as a platform for active judicial oversight.

intelligent manner.⁹⁵ Still, judges seem quite willing to enforce most agreements as long as they deal with the set of procedures recognized as suitable for *ex ante* specification.

One possible reason for the scarcity of cases involving party rulemaking might have to do with the availability of arbitration.⁹⁶ But this explanation just poses the salient question in sharp relief: *If parties can and do cooperatively design their own procedures in arbitration, why should they not have broad power to do so in adjudication as well?*

III. The Normative Dimension: Conventional Arguments

The remainder of this Article addresses the normative question of how much power parties should have to make their own procedural rules. This part summarizes and critically reviews the conventional arguments for and against party rulemaking. These arguments are important and helpful, but they are also seriously incomplete because they lack a convincing normative account of adjudicative legitimacy. Part IV attempts to fill that gap. But first we need to see just where and why the existing arguments fall short.

Although the following discussion focuses mainly on policy, there is an important legal question that should be addressed at the outset. Some critics argue that it is inconsistent with the Rules Enabling Act to give private parties the power to alter official rules adopted through the congressionally authorized rulemaking process.⁹⁷ Section III(B)(2) below responds to this argument at the policy level by explaining how party rulemaking is consistent with the values and policies served by the rulemaking process. Here, I

95. See *supra* note 78 and accompanying text.

96. Another reason might have to do with the rise of trial judge case management. It is well known that the Federal Rules of Civil Procedure delegate broad discretion to trial judges. See Bone, *Procedural Discretion*, *supra* note 7, at 1967–70 (examining ways the Federal Rules facilitate trial judge discretion). If judges frequently shape case-specific procedure without much constraint from general rules, parties would have very weak incentives to make rules *ex ante* knowing that those rules would just end up being renegotiated with the judge *ex post*. There are two problems with this explanation. First, case management has arisen in response to concerns about case backlog and delay, so one would expect trial judges to be receptive to agreements that limit procedure to save costs. See Bone, *The Process of Making Process*, *supra* note 13, at 904 (noting the example of the Civil Justice Reform Act of 1990, which required courts to develop case-management plans in an effort to control case backlog and delay). Yet there are very few cases evidencing those agreements. For example, one does not see parties agreeing to stricter pleading standards even though they could benefit from such an agreement and the agreement should be appealing to trial judges concerned about backlog and frivolous suits. Second, aggressive trial judge case management is a phenomenon of the past thirty years, but there is no evidence of broader party rulemaking before that time. See Gensler, *supra* note 64, at 670–72 (noting the evolution from passive to active case management by federal judges in the last thirty years, including amendments to the Federal Rules of Civil Procedure giving judges an “ever-expanding set of case-management tools”).

97. See, e.g., Taylor & Cliffe, *supra* note 22, at 1098–104 (criticizing judicial recognition of private agreements to alter procedural rules as circumventing the deliberative decisions of Congress and undermining the goals of uniformity and resolution based upon the merits rather than upon procedural technicalities).

want to address a different concern. The Rules Enabling Act authorizes “general rules of practice and procedure.”⁹⁸ If this phrase is understood as a limitation, then one might argue that the rulemaking committees have no power to authorize party rulemaking because it creates case-specific rather than “general” rules.⁹⁹ In other words, by adopting a master rule authorizing party rulemaking, the rulemaking committees would in effect give power to private parties to make rules that the committees could not make themselves.¹⁰⁰

However, the Rules Enabling Act’s “general rules” requirement, whatever it means, cannot possibly bar case-specific procedure. As we have seen, Rule 29 of the Federal Rules of Civil Procedure already gives parties power to alter the generally applicable discovery rules, and to my knowledge, no one has suggested that this Rule violates the Enabling Act.¹⁰¹ Moreover, both trial judges and parties have broad power to tailor procedures to the specific circumstances of particular cases under the current Federal Rules, and no one claims that this power is inconsistent with the Enabling Act.¹⁰² To be sure, these are examples of Type II rulemaking, but I can see no principled reason to distinguish between Type II and Type III insofar as application of the “general rules” provision is concerned. Indeed, it seems reasonable to construe the phrase “general rules” to refer to a uniformly applied set of rules, even if some of those rules authorize judges or parties to make more specific rules for individual cases.¹⁰³

Whether federal courts have common law power to authorize party rulemaking inconsistent with the Federal Rules is a more complicated question.¹⁰⁴ The Supreme Court has noted on more than one occasion that

98. Rules Enabling Act § 401(a), 28 U.S.C. § 2072(a) (2006).

99. See Stephen B. Burbank, *Pleading and the Dilemmas of “General Rules,”* 2009 WIS. L. REV. 535, 541–43 (noting that the drafters intended the phrase *general rules* to ensure interdistrict uniformity and probably also assumed uniformity vis-à-vis the substantive character of cases).

100. One might also object that this places rulemaking authority in the hands of private parties, but the parties make rules only for their own cases.

101. See *supra* note 61 and accompanying text.

102. See Bone, *Procedural Discretion*, *supra* note 7, at 1967 (noting the broad scope of case-specific discretion that the Federal Rules of Civil Procedure give district judges).

103. It is an interesting question whether a rule made by the parties pursuant to an explicit grant of party-rulemaking power in the Federal Rules could ever violate the Rules Enabling Act proviso prohibiting rules that “abridge, enlarge or modify any substantive right.” Rules Enabling Act § 401(a), 28 U.S.C. § 2072(b). If the Federal Rules were to authorize Type III rulemaking and parties were to adopt a rule for the purpose of indirectly limiting or expanding substantive rights (such as a rule shifting the burden of persuasion or eliminating judgment as a matter of law), one might argue that the proviso has been violated. However, a Federal Rule authorizing Type III rulemaking furthers a procedural purpose—enabling party choice in litigation—and therefore is not obviously substantive. The situation might be different if private parties routinely contracted for a particular procedural rule that had major substantive effects contrary to congressional policy. In that case, the Federal Rule that licensed this predictable pattern might be invalid as violating the Rules Enabling Act proviso.

104. Federal courts do have power to make common law procedural rules, and they have done so in several different areas, such as forum non conveniens and preclusion. See Amy Coney

the federal courts do not have power to adopt common law procedural rules inconsistent with the Federal Rules.¹⁰⁵ A common law rule favoring enforcement of Type III rulemaking agreements would not itself be inconsistent with the Federal Rules unless the Federal Rules were construed to be mandatory. But a common law rule might be inconsistent if its specific aim were to license private parties to deviate from the Federal Rules.

We need not dwell on these issues any further. The following discussion assumes that if party rulemaking is justified on policy grounds, it can be implemented by Federal Rule, common law, or congressional statute.

A. Considerations Favoring Party Rulemaking

1. *The Flawed Argument from Arbitration.*—Some courts and commentators defend broad freedom to shape procedure in adjudication by comparing adjudication with arbitration. One version of this argument holds that party-made procedures cannot be unfair or inefficient if they are tolerated in arbitration.¹⁰⁶ Another version holds that there is little point in disallowing party rulemaking if the result will be that parties exit adjudication and create the same procedures in arbitration.¹⁰⁷ And a third version argues that it is a good idea to allow parties to design their own procedures in adjudication as a way to discourage them from escaping adjudication for arbitration.¹⁰⁸

All of these arguments are extremely weak. The first assumes that arbitration and adjudication are normatively comparable: what is unfair or inefficient for adjudication must also be unfair or inefficient for arbitration. But this is not necessarily correct. Civil adjudication performs different functions than arbitration and draws on different sources for its institutional legitimacy. In Part IV, I argue that the distinctive feature of adjudication is its commitment to a particular form of principled reasoning and that this commitment is essential to its institutional legitimacy. The second and third versions of the argument are also inadequate. The risk that parties will turn

Barrett, *Procedural Common Law*, 94 VA. L. REV. 813, 822–32 (2008) (discussing five different areas of procedural common law, including forum non conveniens and preclusion). But none of these rules conflict with the Federal Rules of Civil Procedure.

105. See, e.g., *Swierkiewicz v. Sorema N.A.*, 534 U.S. 506, 515 (2002) (stating that a heightened pleading standard in discrimination suits “must be obtained by amending the Federal Rules, and not by judicial interpretation” (quoting *Leatherman v. Tarrant Cnty. Narcotics Intelligence & Coordination Unit*, 507 U.S. 163, 168 (1993))); *Leatherman*, 507 U.S. at 168 (stating that an added specificity requirement for claims against municipalities would require amending the Federal Rules).

106. Noyes, *supra* note 21, at 594, 620.

107. See, e.g., *In re Prudential Ins. Co. of Am.*, 148 S.W.3d 124, 132 (Tex. 2004) (arguing that an *ex ante* contractual waiver of a jury trial should be enforced because otherwise the parties would have to go to arbitration).

108. *Id.* at 132; Moffitt, *supra* note 21, at 490–91; cf. Noyes, *supra* note 21, at 594 (arguing for more contractual flexibility to design court procedures to facilitate party access to the superior features of adjudication compared to arbitration).

to arbitration is not a good enough reason by itself to make adjudication more party friendly. It all depends on the costs and benefits of the change. Indeed, this argument taken to its logical extreme could justify converting adjudication completely into arbitration.

Still, the comparison to arbitration is illuminating. If it makes sense to bar some types of party rulemaking in adjudication even when those types are permissible in arbitration, it must be because adjudication is different in a normatively relevant way. Part IV develops that insight in greater depth.

2. *The Benefits of Party Rulemaking.*—Proponents of party rulemaking cite a number of benefits, and the following discussion addresses each in turn.

a. *Outcome Benefits.*—Parties can use *ex ante* agreements to solve collective-action problems that produce costly and wasteful litigation investment *ex post*.¹⁰⁹ A good example is discovery. As many commentators have pointed out, parties to a lawsuit face a collective-action problem at the discovery stage. In one version of the problem, it is a classic prisoner's dilemma: each party anticipates that the other will use discovery abusively, so each responds in kind, and the result is an equilibrium in which both sides are worse off than if they had exercised restraint.¹¹⁰ By agreeing to limit discovery in advance, the parties are able to commit to mutual restraint—that is, provided courts routinely enforce discovery agreements *ex post*. In a different version, the parties adopt substitute strategies in equilibrium: one side invests much more aggressively than the other.¹¹¹ This is less likely to produce wasteful litigation costs, but it is more likely to produce skewed settlements. By contracting for discovery limits before a dispute arises, when neither party knows which side of a future lawsuit she will occupy, the

109. See Kapeliuk & Klement, Contractualizing Procedure, *supra* note 21, at 17–19 (discussing these benefits); see also Steven Shavell, *Alternative Dispute Resolution: An Economic Analysis*, 24 J. LEGAL STUD. 1, 5–21 (1995) (concluding that alternative dispute resolution (ADR) that is contractually chosen before a dispute arises provides economic benefits while ADR chosen after a dispute arises does not).

110. BONE, CIVIL PROCEDURE, *supra* note 44, at 219–24; Ronald J. Gilson & Robert H. Mnookin, *Disputing Through Agents: Cooperation and Conflict Between Lawyers in Litigation*, 94 COLUM. L. REV. 509, 514–15 (1994); John K. Setear, *The Barrister and the Bomb: The Dynamics of Cooperation, Nuclear Deterrence, and Discovery Abuse*, 69 B.U. L. REV. 569, 584–86 (1989).

111. See Chris William Sanchirico, *Harnessing Adversarial Process: Optimal Strategic Complementarities in Litigation 2–3* (Jan. 2006) (unpublished manuscript), available at <http://ssrn.com/abstract=788564> (arguing that parties will “strategically substitute” and citing empirical research on discovery showing that plaintiffs retreat in response to defendant aggression). The reason why the parties adopt opposite strategies is easy to understand. Both parties know that abusive expenditures on discovery will simply cancel out and thus confer no gain. Given this, the parties are better off adopting opposite strategies. When *A* adopts a strategy of retreat in the face of aggression, for example, *B* will act less aggressively since *A*'s retreat reduces the marginal benefit to *B* of additional aggression. And less aggression by *B* makes *A* better off.

parties can commit to mutual restraint and thus reduce the likelihood of a one-sided result.

Furthermore, party rulemaking can influence the parties' pre-suit behavior in ways that enhance their joint welfare. To illustrate, suppose *A* enters into a contract with *B* for the design and construction of a new technological component.¹¹² The technology is complex, and *A* and *B* worry about a significant risk of jury error in the event of a future lawsuit for breach of contract. The parties can reduce this risk substantially by agreeing in advance to waive jury trial and appoint an expert who will assess compliance with the contract specifications in the event of an alleged breach. Doing so increases the value of the contract to *A* and *B*, and a more valuable contract can elicit greater investment in the venture, which in turn increases the contract's value even more.

There is a third way that parties can reap benefits by using party rulemaking. Their *ex ante* procedural choices can signal private information that is difficult to communicate credibly in any other way.¹¹³ For example, suppose that *A*, a trade secret owner, and *B*, a licensee, enter into a long-term trade-secret licensing agreement. Suppose that in arrangements of this kind, some trade secret owners bring frivolous suits during the licensing period in order to pressure settlement for better licensing terms. *A* assures *B* that it is not that kind of trade secret owner, but *B* is skeptical. *A* might agree to a strict pleading rule that makes it more difficult to file a frivolous suit. In equilibrium, *B* will infer from *A*'s willingness to include such a rule that *A* is not the type of firm that files frivolous suits. As a result, *B* should be willing to pay more for the license and perhaps invest more in the arrangement.

Party rulemaking can also generate benefits beyond the parties to the agreement. When party-chosen rules enhance the value of a contract, there are likely to be spillover benefits for others. More directly related to litigation, party rulemaking can reduce the public costs of the court system when it reduces the private costs of litigation insofar as private and public costs are correlated. Public benefits can also be produced in more complicated ways. Suppose *A* and *B* agree to a strict pleading rule that screens frivolous suits. If the presence of frivolous suits in litigation makes it more difficult for parties to settle meritorious suits, as is likely, a strict pleading rule in a case between *A* and *B* should make it easier for parties to settle and thereby save the public cost of a trial.¹¹⁴

112. I adopt this example from Kapeliuk & Klement, *Contractualizing Procedure*, *supra* note 21, at 21.

113. *Id.* at 23–25.

114. See Robert G. Bone, *Modeling Frivolous Suits*, 145 U. PA. L. REV. 519, 547–50, 559–63 (1997) [hereinafter Bone, *Modeling Frivolous Suits*] (describing equilibria under conditions of asymmetric information where meritorious suits go to trial rather than settle because of the presence of meritless suits). See generally Avery Katz, *The Effect of Frivolous Lawsuits on the Settlement of Litigation*, 10 INT'L REV. L. & ECON. 3 (1990) (analyzing the case where asymmetric information favors the plaintiff).

However, this rosy picture assumes that we can predict the effects of a party-selected procedural rule. Prediction, it turns out, is very tricky in the complex strategic environment of litigation. A party-made rule that seems to reduce public costs might actually increase those costs. For example, an agreement to limit discovery could increase public costs if the expectation of a less onerous discovery burden and limited access to information reduced the size of the settlement surplus and with it the likelihood of settlement, thereby increasing the risk of trial.¹¹⁵ Also, by restricting access to information, discovery limits could generate trial or settlement outcomes with a higher-than-optimal error risk, thereby undermining deterrence goals.¹¹⁶ To be sure, parties will take account of private costs when they negotiate their contract, but there is no reason for them to take account of public costs like these. We shall return to this point later.

b. Autonomy Values.—One might argue in support of party rulemaking that giving freedom of choice enhances party autonomy, which furthers one of the core values of the adversary system.¹¹⁷ This argument has superficial appeal. Party rulemaking, after all, does give litigants a larger menu of choices. But it turns out to be quite problematic on closer analysis.

Party autonomy serves two main functions in litigation. First, it promotes accurate outcomes by harnessing private incentives to develop the facts and the law. As the previous subsection suggests and as is developed in greater detail later, giving the parties a broader range of choices may or may

115. This can happen in at least two ways. First, limited discovery, while it reduces litigation costs, also restricts access to private information useful for aligning the parties' estimates of likely success. Divergent estimates are a major impediment to settlement. See Robert D. Cooter & Daniel L. Rubinfeld, *An Economic Model of Legal Discovery*, 23 J. LEGAL STUD. 435, 439-44 (1994) (discussing the ways in which information exchange through discovery increases the probability of settlement). Second, limited discovery reduces total discovery costs and thus the amount that the parties can save by settling. With less to save, the parties might be less inclined to settle. See BONE, CIVIL PROCEDURE, *supra* note 44, at 71-76 (describing the basic economic model of settlement).

116. The parties do not internalize all the social benefits of increased deterrence, and therefore they might agree to discovery limits that make them better off but weaken deterrence from a social perspective. See generally Steven Shavell, *The Social Versus the Private Incentive to Bring Suit in a Costly Legal System*, 11 J. LEGAL STUD. 333 (1982) (describing the divergence between social and private benefits and costs of litigation).

117. Versions of this argument appear in the literature dealing with procedural contracting. For example, Professor Moffitt argues that procedural customization through contract promotes procedural justice values by furthering party participation and control. Moffitt, *supra* note 21, at 479-81. I discuss this procedural justice argument in the text below. Furthermore, Professor Noyes argues that liberal enforcement of procedural contracts reflects a commitment to party autonomy, promotes the idea that "parties own their disputes and may design their own dispute resolution rules," and preserves "the concept of freedom of contract." Noyes, *supra* note 21, at 598 n.78, 620-21 (citation omitted). In addition, autonomy values are often used to justify arbitration. See, e.g., Edward Brunet, *The Core Values of Arbitration*, in EDWARD BRUNET ET AL., ARBITRATION LAW IN AMERICA: A CRITICAL ASSESSMENT 3, 11, 28 (2006) (identifying party autonomy as one of the key values of arbitration).

not enhance outcome quality, depending on party incentives, externalities, and public costs.¹¹⁸

Party autonomy is also thought to serve an intrinsic function. There are two different versions of this claim. One version focuses on the psychological benefits of participation. Those who subscribe to this version point to a body of empirical findings, collectively known as the procedural justice literature, that shows that parties tend to be more satisfied with the justice of the outcome and the fairness of the process when they have had a chance to participate personally, even if they lose in the end.¹¹⁹ At first glance, this literature might seem to offer support for party rulemaking. Perhaps feelings of just and fair treatment increase with the expanded control opportunities that procedural contracting affords.¹²⁰

However, this argument is weak. A party might feel better at the time of contracting, but she is likely to feel frustrated at the time of enforcement whenever her *ex ante* choice turns out to be adverse to her *ex post* interests. Indeed, after she knows the facts of her particular case, she might even think that it is unjust to apply the procedure she agreed to *ex ante*. Furthermore, even if procedural contracting does enhance feelings of just treatment, that fact alone offers no normative reason to support party rulemaking. It must be connected to a normative theory that gives weight to party preferences.¹²¹

The second version of the intrinsic value claim is normative. It focuses on a Kantian principle of respect for persons and holds, roughly speaking, that respect for the autonomy and dignity of individual litigants requires giving each party a right to participate in and exercise some control over proceedings that significantly affect her life.¹²² I have referred to this in other writing as a “process-oriented” theory of participation because it focuses on the value of participation to the litigation process apart from any

118. See *infra* section III(B)(3).

119. See, e.g., E. ALLAN LIND & TOM R. TYLER, *THE SOCIAL PSYCHOLOGY OF PROCEDURAL JUSTICE* 26–40, 93–106 (1988) (discussing the empirical literature); JOHN THIBAUT & LAURENS WALKER, *PROCEDURAL JUSTICE: A PSYCHOLOGICAL ANALYSIS* 68–80 (1975) (presenting some of the findings).

120. See Moffitt, *supra* note 21, at 479–81 (arguing from “[t]he lessons of procedural justice research” that “[p]roviding disputants with process control increases their perception of justice”).

121. As I have argued elsewhere, the most suitable normative theory for this purpose is utilitarian, and the aggregative calculus of utilitarianism gives positive feelings about process or outcome no particular priority over any other feelings. Bone, *Agreeing to Fair Process*, *supra* note 43, at 505–07.

122. See, e.g., JERRY L. MASHAW, *DUE PROCESS IN THE ADMINISTRATIVE STATE* 158–200 (1985) (defending a dignitary-values theory of due process). Other scholars argue that participation is essential to the legitimacy of adjudication as a source of binding judgments, just as participation is essential to the legitimacy of legislation and other forms of government action in a liberal democracy. See, e.g., Lawrence B. Solum, *Procedural Justice*, 78 S. CAL. L. REV. 181, 273–77 (2004) (arguing that the legitimacy of final adjudications is predicated on the right of parties to participate and that this value cannot be reduced to accuracy or efficiency).

value it might have for outcome quality.¹²³ Assuming that process-oriented theory is coherent—and it is not at all clear that it is¹²⁴—the question is whether the value of process-oriented participation is furthered by party rulemaking. The answer is unclear. Although allowing enforceable rulemaking contracts expands the *ex ante* choice set, it also restricts participation opportunities and limits party control *ex post*. A party who would rather employ a different procedure at the time of litigation is unable to do so when she is bound by a pre-suit agreement.

Thus, the critical question for the autonomy argument is which perspective—*ex ante* or *ex post*—is the right one to use to evaluate party autonomy and litigant control. A good argument can be made in favor of crediting control only when it is exercised *ex post* in the context of a specific lawsuit and not *ex ante* before any dispute arises. To see this point, consider an analogy to voting. The right to vote has been justified on intrinsic as well as instrumental grounds.¹²⁵ Like guaranteeing party control over litigation, assuring the right to vote is said to respect individual autonomy and dignity by giving each person an opportunity to participate in elections that affect her life and on an equal basis with all other citizens.¹²⁶ But the autonomy values underlying the right to vote are not advanced by allowing voters to commit before an election to vote in a particular way and then enforcing those commitments on election day when the voters have changed their minds. Voting has value not simply as another way to exercise freedom of choice. It has value primarily as a way for individuals to express their political preferences within the institutional framework of electoral politics, which means at the ballot box.

Similarly, the value of party autonomy in litigation depends on how participation works within the framework of adjudication, and this depends in turn on a theory of adjudicative participation. If participation has intrinsic value as a means of exercising control during litigation, *ex ante* agreements restricting choices *ex post* would not necessarily further party autonomy. In any event, it will not do to argue that procedural contracting furthers autonomy in a relevant way just because it expands party choice by adding a contract option.

123. Robert G. Bone, *Rethinking the "Day in Court" Ideal and Nonparty Preclusion*, 67 N.Y.U. L. REV. 193, 201–02 (1992) [hereinafter Bone, *Rethinking*].

124. *See id.* at 279–88 (detailing a number of ways in which the process-oriented view is problematic).

125. *See, e.g., Reynolds v. Sims*, 377 U.S. 533, 561–62 (1964) (arguing that the right to vote is instrumental in securing “basic civil and political rights”); MASHAW, *supra* note 122, at 163 (arguing that enfranchisement has an intrinsic value linked to autonomy out of all proportion to the minuscule amount of political power it actually confers); Solum, *supra* note 122, at 277 (noting the intrinsic value of the right to an equal vote independent of outcomes).

126. JOHN RAWLS, *A THEORY OF JUSTICE* 194–200 (rev. ed. 1999).

B. Considerations Opposing Party Rulemaking

1. *The Problem of Consent.*—One of the major criticisms of party rulemaking has to do with defective consent.¹²⁷ According to this argument, consent can never be meaningful when bargaining power is seriously skewed, as it is for consumer, employment, and other similar contracts.¹²⁸ Over the past decade, this complaint has focused mainly on contracts of adhesion in the arbitration setting.¹²⁹ This is largely because arbitration has been the primary locus of procedural contracting. However, the complaints readily carry over to adjudication as well.¹³⁰ Therefore, we can learn a great deal about the strengths and limits of the consent argument in general by examining its merits in the arbitration setting.

The major focus of attention these days is consumer arbitration, but arbitration of employment disputes not covered by a collective bargaining agreement has also come into heavy criticism.¹³¹ These two types of

127. See Taylor & Cliffe, *supra* note 22, at 1105 (identifying an imbalance of bargaining power as “[o]ne of the more troubling aspects of” prelitigation agreements); Thornburg, *supra* note 22, at 209–10 (identifying valid consent as an important factor in enforcing a procedural contract).

128. See Taylor & Cliffe, *supra* note 22, at 1105 (“[T]he opportunity for bargaining is not realistically present . . . in employment contracts, franchise agreements, and consumer transactions, where one party is largely at the disposal of the other in entering the contract . . .”).

129. See, e.g., Linda J. Demaine & Deborah R. Hensler, “Volunteering” to Arbitrate Through Predispute Arbitration Clauses: The Average Consumer’s Experience, 67 LAW & CONTEMP. PROBS. 55, 72–74 (2004) (noting the lack of voluntary agreement to many consumer arbitration contracts); Charles L. Knapp, *Taking Contracts Private: The Quiet Revolution in Contract Law*, 71 FORDHAM L. REV. 761, 788–89 (2002) (noting bargaining-power problems in consumer arbitration agreements); Jean R. Sternlight, *Creeping Mandatory Arbitration: Is It Just?*, 57 STAN. L. REV. 1631, 1648–53 (2005) (describing criticisms of mandatory binding arbitration, including the lack of meaningful consent and the unfair terms and oppressive tactics that it facilitates).

130. In fact, many commentators focus on problematic consent as a ground for rejecting the Supreme Court’s generous approach to enforcing forum-selection clauses. See, e.g., Carrington & Haagen, *supra* note 16, at 350–57 (lamenting the lower courts’ enforcement of the *Bremen* rule in cases where the forum-selection clause was buried in fine print and criticizing the flawed argument that Carnival Cruise Line’s forum-selection clause will result in savings to those who purchase tickets for a cruise); Mullenix, *supra* note 16, at 362–63 (arguing that courts almost always find fair bargaining—without regard to the relative sophistication of the respective parties—by presuming that the party opposing the forum-selection clause “received consideration or a contractual concession in return for the provision”).

131. See, e.g., David Sherwyn et al., *Assessing the Case for Employment Arbitration: A New Path for Empirical Research*, 57 STAN. L. REV. 1557, 1558–60 (2005) (noting the explosion of literature about employment arbitration and the sharp debate over its merits). There is also some criticism of arbitration in franchise agreements. See, e.g., Jennifer L. Gehrig, *Arbitration: A Franchisee’s Perspective*, 22 FRANCHISE L.J. 121, 121 (2002) (acknowledging that arbitration clauses have “allowed franchisors to shield themselves from class action suits by franchisees, limit available damages, select a forum favorable to the franchisor, and eliminate franchisees’ right to a jury trial,” but that, at times, arbitration clauses can be favorable to franchisees); William L. Killion, *An Informal Study of Arbitration Clauses Reveals Surprising Results*, 22 FRANCHISE L.J. 79, 79 (2002) (noting that “arbitration clauses in typical franchise agreements probably deprive franchisees of the ability to band together in suing their franchisor” but that arbitration clauses are still beneficial to franchisees). But franchise agreements are not really adhesion contracts. The franchise agreement involves relatively high stakes and the parties are businesspeople who are

arbitration share common features that can be problematic. Both involve repeat players—the employer in one case and the seller in the other—who enjoy relative advantages in crafting the arbitration contract and prosecuting an arbitration proceeding.¹³² Moreover, the employment agreement, like the consumer contract, can be adhesive in nature, although employment agreements are less likely to be adhesive than consumer contracts.¹³³ Of course, the fact that an employment contract is adhesive is not necessarily a problem in itself. After all, many features of the employment relationship are imposed on a take-it-or-leave-it basis.¹³⁴ In addition, labor-market competition can ameliorate some of the potentially adverse effects of adhesive employment contracts in much the same way as competition can for consumer contracts, as described below. However, competition works only if a sufficiently large number of employees read and understand these agreements and can shop for jobs, both of which are questionable for employees in the lower ranks.

These problematic features are certainly cause for concern. However, an employee can also benefit from arbitration even when she does not expressly bargain for it.¹³⁵ For example, sometimes both the employee and the employer desire the confidentiality that arbitration confers, especially when the dispute involves sensitive matters. Also, the employee can benefit from speedier and less costly dispute resolution insofar as arbitration is faster and less expensive than adjudication. Moreover, employment disputes are less likely than consumer disputes to involve the small-claims problem discussed below, so individual arbitration proceedings are more likely to be cost

likely to consult lawyers and bargain for their contract terms. See Christopher R. Drahozal & Keith N. Hylton, *The Economics of Litigation and Arbitration: An Application to Franchise Contracts*, 32 J. LEGAL STUD. 549, 581–82 (2003) (noting that arbitration is frequently criticized as coercive but asserting that “[t]he coercion claim is weak in the franchising context, since . . . both parties are businesses that consult with lawyers”); Christopher R. Drahozal, “Unfair” Arbitration Clauses, 2001 U. ILL. L. REV. 695, 766 (“Many franchisees are sophisticated business people who can and do shop around for franchise opportunities.”).

132. For the classic discussion of repeat-player versus one-shotter dynamics, see generally Marc Galanter, *Why the “Haves” Come Out Ahead: Speculations on the Limits of Legal Change*, 9 LAW & SOC’Y REV. 95, 97–114 (1974). For a discussion of repeat-player problems in employment arbitration, see generally Lisa B. Bingham, *Employment Arbitration: The Repeat Player Effect*, 1 EMP. RTS. & EMP. POL’Y J. 189 (1997). But see Sherwyn et al., *supra* note 131, at 1570–71 (reviewing the criticisms of Bingham’s studies and concluding that she fails to support arbitrator bias in the employment arbitration setting).

133. The consumer contract in mass marketing is clearly a take-it-or-leave-it arrangement, whereas employees, especially those at higher levels of the company, sometimes have the ability and the sophistication to bargain for contract terms. See, e.g., Randall Thomas et al., *Arbitration Clauses in CEO Employment Contracts: An Empirical and Theoretical Analysis*, 63 VAND. L. REV. 959, 960 (2010) (noting that CEOs negotiate their employment agreements).

134. See Sherwyn et al., *supra* note 131, at 1563–64 (making this point and noting that the contract-of-adhesion criticism is mostly “an issue of perception”).

135. See *id.* at 1578–81 (reviewing the empirical studies current as of 2005 and conducting their own empirical study, and theorizing that while some employers might use arbitration to avoid courts and undermine employee rights, many use arbitration to provide a nonadversarial, low-cost forum for low-value claims).

effective. And the data on win rates do not show a strong skewing effect in favor of employers compared to outcomes in adjudication.¹³⁶ In any event, it is clear that lack of consent associated with adhesion employment contracts is not itself a reason to deny party rulemaking anymore than it is a reason to deny all employment arbitration.¹³⁷ There may be situations where the procedures are obviously unfair, and in those cases the judge can refuse enforcement.¹³⁸

The critics of employment arbitration often cite a different problem, one that has little to do with lack of employee consent. They argue that adjudication in court is superior to arbitration for enforcing the broad public interest in statutory civil rights claims, such as Title VII claims, Age Discrimination in Employment Act claims, and the like.¹³⁹ However, this objection is much weaker for party rulemaking than for arbitration since party rulemaking keeps the dispute in a federal court and assures that it will be public to some extent. Moreover, the judge has the power to deny enforcement to any procedures that sharply conflict with a statute's public goals.

Consumer arbitration is more controversial than employment arbitration. This category includes consumer contracts with discount-securities brokers, credit card companies, telecommunications providers, and so on.¹⁴⁰ A recent Supreme Court case, *AT&T Mobility LLC v. Concepcion*, nicely illustrates the problem. In that case, AT&T Mobility included an arbitration clause combined with a class-action-waiver clause in wireless-service contracts offered to its subscribers.¹⁴¹ The plaintiffs complained that

136. At least with the empirical studies available as of 2005. *Id.* at 1567–72. Sherwyn et al. also conclude that the empirical results on comparative damages amounts are inconclusive. *Id.* at 1573–78. *But see* David S. Schwartz, *Mandatory Arbitration and Fairness*, 84 NOTRE DAME L. REV. 1247, 1283–315 (2009) (criticizing empirical studies that purport to show the fairness of arbitration).

137. *See, e.g.*, Thomas et al., *supra* note 133, at 963 (conducting a study of CEO contracts showing that the presence of arbitration clauses turns on economic factors and concluding that “[i]n the context of the contracts-of-adhesion theory, . . . while most employees are unable to bargain for arbitration when those nuanced factors indicate that it is desirable, a flat ban on arbitration clauses in all employment contracts may be too blunt a policy tool”).

138. One empirical study identified circumstances where bargained-for CEO agreements included arbitration clauses and generalized from these results to identify circumstances where other employees might have agreed to arbitration had they been able to bargain. *See id.* at 962 (noting, for example, that arbitration clauses were used in industries subject to rapid change (where speedy resolution is desirable), and in firms with lower profitability (where low-cost procedures are advantageous) and also when complicated issues are involved (since those are better decided by expert decision makers)).

139. *See, e.g.*, Sternlight, *supra* note 129, at 1664–65 (“[W]e care more when federal statutory claims such as employment discrimination are taken away from the public eye than when a dispute over the quality of soybeans shipped from Missouri to Nevada is handled privately.”).

140. *See supra* note 129 (collecting relevant sources).

141. *AT&T Mobility LLC v. Concepcion*, 131 S. Ct. 1740, 1744 (2011). It is quite common for sellers to include arbitration clauses in consumer contracts and combine them with clauses waiving class arbitration. *See* Theodore Eisenberg et al., *Arbitration's Summer Soldiers: An*

the agreement deprived consumers of the only viable method of obtaining relief for small claims—the class action—and that it did so through a contract of adhesion offered on a take-it-or-leave-it basis without any opportunity for meaningful consent.¹⁴²

At the outset, it is important to note the limited scope of this criticism. It does not apply to cases where sophisticated parties bargain for contract terms. There might still be reasons to oppose enforcement in these situations, but those reasons are unlikely to include defective consent.

Moreover, even in the context of adhesion contracts, the policy analysis is more complex and the results more nuanced than many of the most vocal critics assume. There is no question that wealth disparities, informational asymmetries, and bounded-rationality constraints produce some problematic agreements.¹⁴³ But these same factors also operate in ordinary litigation to skew settlements and trial judgments. It is true that consumers have legal representation when they file lawsuits but not when they buy products. However, agency problems weaken the attorney–client bond. Indeed, consumer suits and other class actions featuring mostly small claims are notorious for high agency costs.¹⁴⁴ In short, it is a mistake to romanticize civil litigation, just as it is a mistake to romanticize arbitration. If the criticism is about arbitration as it actually exists and arbitration agreements as they actually operate, then the proper comparison is to litigation as it is actually conducted.

Furthermore, the fact that a contract does not involve bargaining is not a reason by itself to condemn it. Standard form contracts are pervasive these days, and mass markets could not function effectively without them.¹⁴⁵

Empirical Study of Arbitration Clauses in Consumer and Nonconsumer Contracts, 41 U. MICH. J.L. REFORM 871, 882–85, 888 (2008) (concluding that sellers use consumer arbitration mainly to avoid the risk of aggregate dispute resolution).

142. Brief for Respondents at 39–44, *Concepcion*, 131 S. Ct. 1740 (No. 09-893), 2010 WL 4411292, at *39–44. This is essentially the basis for the Ninth Circuit’s unconscionability holding in the case, which was later reversed by the Supreme Court. *Laster v. AT&T Mobility LLC*, 584 F.3d 849, 853–54 (9th Cir. 2009), *rev’d sub nom. AT&T Mobility LLC v. Concepcion*, 131 S. Ct. 1740, 1753 (2011). Moreover, this is one of the main arguments that critics use to attack consumer arbitration agreements and class action waivers more generally. See generally Sternlight, *supra* note 129, at 1648–53 (summarizing the critics’ arguments).

143. See, e.g., Russell Korobkin, *Bounded Rationality, Standard Form Contracts, and Unconscionability*, 70 U. CHI. L. REV. 1203, 1217–18 (2003) (“[T]he fundamental cause of inefficient terms in form contracts lies in the boundedly rational approaches buyers use to evaluate information . . .”).

144. See Jonathan R. Macey & Geoffrey P. Miller, *The Plaintiffs’ Attorney’s Role in Class Action and Derivative Litigation: Economic Analysis and Recommendations for Reform*, 58 U. CHI. L. REV. 1, 19–27 (1991) (“The existing regulatory system cannot effectively deal with agency costs that arise in class action and derivative litigation because plaintiffs in the class action and derivative context are often completely incapable of monitoring the attorney.”).

145. See, e.g., *Hill v. Gateway 2000, Inc.*, 105 F.3d 1147, 1149 (7th Cir. 1997) (discussing the problems with providing notice of contractual terms to consumers in advance of a purchase and the efficiencies of relying instead on the availability of postpurchase rejection once the consumer has received the enclosed legal terms with the product); Mark R. Patterson, *Standardization of*

Indeed, enforcing procedural contracts can improve consumer welfare even when those contracts are adhesionary and consent is problematic.¹⁴⁶ The reason is easy to see. Suppose most consumers would prefer to accept limited procedures in a future lawsuit in return for a lower product price. If Firm *A* tries to force the limited procedures through an adhesion contract without reducing price, competitors have an incentive to offer the same package at the lower price. Moreover, if Firm *A* includes draconian procedures that consumers would reject if they had a chance, competitors should step in and offer better terms for a somewhat higher price and aggressively publicize Firm *A*'s exploitative conduct and the comparative benefits of the competitor's terms.¹⁴⁷ Therefore, to some extent, market forces limit the ability of sellers to foist upon consumers grossly one-sided procedural terms.¹⁴⁸

In fact, by enforcing procedural contracts *ex post*, sellers can reliably offer different procedural packages to cater to different consumer tastes. Consumers who value procedure highly will be able to purchase a package with the robust procedures they want at a higher price, and consumers who do not care much about procedure will be able to purchase a suitable package at a lower price. Moreover, if the different packages are advertised aggressively, consumers might receive helpful information about the relative advantages of different procedural opportunities.

Not all consumers need know about or understand the price-procedure package to drive this type of competition, as long as enough do and the seller cannot distinguish one type of consumer from the other. The less knowledgeable consumers can free ride on the more knowledgeable.¹⁴⁹ And if

Standard-Form Contracts: Competition and Contract Implications, 52 WM. & MARY L. REV. 327, 331 & n.1 (2010) (noting the widespread use of standard form contracts).

146. See generally O'HARA & RIBSTEIN, *supra* note 27, at 34–36, 133–45 (analyzing the benefits and costs of enforcing choice-of-law and choice-of-forum clauses in consumer contracts). In fact, consumers give consent in a competitive market that offers a variety of product-procedure packages by shopping for the package that best meets their preferences. For discussion of these points in the arbitration context, see Stephen J. Ware, *Replies to Professor Sternlight*, in BRUNET ET AL., *supra* note 117, at 327, 327–34.

147. See David Gilo & Ariel Porat, *Viewing Unconscionability Through a Market Lens*, 52 WM. & MARY L. REV. 133, 139 & n.9 (2010) (“[A]s long as there is a credible threat that competitors . . . bring consumers’ attention to suppliers’ inefficient or unfair terms, no supplier would incorporate such terms in its contract in the first place.”). But see Xavier Gabaix & David Laibson, *Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets*, 121 Q.J. ECON. 505, 505–07 (2006) (arguing that the existence of “myopic” (or unaware) consumers creates an environment that can be at least partially “immune to such competitive pressure”).

148. Also, reputation markets can exert a disciplining force. See Drahozal, *supra* note 131, at 767–69 (observing that in some contexts a firm’s interest in maintaining a good reputation should act as a deterrent to abusive arbitration practices, at least when reputational information is widely distributed).

149. See Keith N. Hylton, *Agreements to Waive or to Arbitrate Legal Claims: An Economic Analysis*, 8 SUP. CT. ECON. REV. 209, 253 n.98 (2000) (noting that this effect depends on the proportion of informed and uninformed consumers); Alan Schwartz & Louis L. Wilde, *Intervening*

consumers value the procedural component of the package highly enough, sellers will have incentives to advertise its advantages, thereby educating everyone. One commentator has argued that sellers are unlikely to offer better procedural terms where doing so will attract consumers with highly litigious preferences.¹⁵⁰ But this just depends on the price the seller is able to charge. As long as there are enough consumers who want the better terms and sellers can charge a price high enough to cover the additional risk, some seller should be willing to offer the option.¹⁵¹

This rosy picture assumes a robustly competitive market with no serious imperfections. Insofar as actual market conditions depart from the ideal, enforcing consumer contracts with limited procedures could be more problematic.¹⁵² But this means only that enforcement should depend on the circumstances; it does not mean that courts should deny enforcement altogether. In fact, it is perfectly rational for a consumer to accept very limited procedural opportunities in return for a lower product price. This is not to say that consumers are free of bounded-rationality constraints; they surely are not.¹⁵³ My point is only that the presence of limited procedures in a consumer contract is not in and of itself a sign that a consumer who accepts the terms must be acting irrationally.

To see the latter point clearly, let us look more closely at consumer choice. Two factors enter into a rational consumer's evaluation of procedural terms: the probability of a future lawsuit, and the expected recovery should such a lawsuit materialize. As for probability, a consumer might reasonably assume a relatively small likelihood of an unsatisfactory product and resulting litigation. Product safety and health are monitored to some extent by public agencies and private testing organizations (like *Consumer Reports*), and numerous websites report product evaluations and consumer

in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis, 127 U. PA. L. REV. 630, 637–38 (1979) (describing how the presence of informed consumers incentivizes companies to offer competitive terms to uninformed consumers as well); *cf.* Drahozal, *supra* note 131, at 766 (making this point for franchisor–franchisee arbitration agreements).

150. Dodge, *supra* note 22, at 761.

151. On the other hand, if only a few consumers want the better procedural package, then the package with limited procedures should satisfy most consumers.

152. There is some evidence that the frequency of pro-seller conflict resolution terms varies with industry concentration and thus the degree of competition in the market. *See, e.g.*, Eisenberg et al., *supra* note 141, at 891–92 (comparing industries with high and low concentrations and correspondingly high and low rates of arbitration clauses). *But see* Florencia Marotta-Wurgler, *Competition and the Quality of Standard Form Contracts: The Case of Software License Agreements*, 5 J. EMPIRICAL LEGAL STUD. 447, 467–74 (2008) (finding no evidence that concentration of software industries leads to worse terms for consumers in end-user license agreements).

153. *See* Christine Jolls, Cass R. Sunstein & Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1505–08 (1998) (examining the possible effect of bounded rationality on the interaction between mandatory contract terms and consumer choice); Korobkin, *supra* note 143, at 1206 (discussing bounded-rationality constraints on information acquisition in standard-form-contract sales).

complaints. Given this intense level of scrutiny, reputable sellers have incentives to market reasonably safe goods and services, and when they do not, consumers are likely to learn about the defects. As for expected recovery, a rational consumer will anticipate receiving very little from a consumer lawsuit unless she suffers serious personal injury from a product defect. Individual losses are too small to justify separate suits, and each consumer's share of a class settlement is minimal. When the probability of a defect and resulting lawsuit is low and any recovery is likely to be small, the expected benefit of more robust procedural terms is also small. Indeed, this may be the reason there is not more competition over procedural packages: rational consumers simply do not care enough about the procedures that govern future lawsuits.

Furthermore, while the evidence is mixed, there are empirical studies that show consumers fare reasonably well in some important types of consumer arbitration.¹⁵⁴ Reputable organizations like the American Arbitration Association (AAA) have incentives to provide reasonably fair arbitration procedures in order to preserve a reputation for evenhandedness.¹⁵⁵ Moreover, reputable sellers have incentives to make use

154. See, e.g., Christopher R. Drahozal & Samantha Zyontz, *An Empirical Study of AAA Consumer Arbitrations*, 25 OHIO ST. J. ON DISP. RESOL. 843, 847–62, 916–18 (2010) (summarizing the results of previous empirical studies and reporting the results of a new empirical study of American Arbitration Association (AAA) arbitrations). One of the chief complaints about arbitration has to do with the fear that arbitrators might bias outcomes in favor of repeat players in order to attract future business. However, the empirical evidence does not clearly support this concern, and several respectable studies show no statistically significant effects. See Eisenberg et al., *supra* note 141, at 873 n.8 (collecting studies reporting no statistically significant difference between litigation and arbitration awards); *id.* at 894 (noting that studies do not show biased outcomes). But see Catherine A. Rogers, *The Arrival of the "Have-Nots" in International Arbitration*, 8 NEV. L.J. 341, 351 n.56 (2007) (collecting sources that suggest there is a repeat-player problem). The risk of bias no doubt depends on the arbitration association. While the AAA has strong incentives to use neutral arbitrators, there are organizations, such as the National Arbitration Forum (NAF), that might be less careful. See Schwartz, *supra* note 136, at 1286 (stating that the NAF "has come under heated criticism from consumer watchdog groups for creating a systematically biased arbitration forum for banking and consumer credit interests"). For a theoretical study of arbitrator bias based on a model of arbitrator incentives, see Alon Klement & Zvika Neeman, *Private Selection and Arbitrator Impartiality* 6–17 (Mar. 31, 2011) (unpublished manuscript), available at <http://ssrn.com/abstract=1800026>.

155. See, e.g., Drahozal, *supra* note 131, at 769–70 (noting incentives of arbitration organizations to provide procedures that judges will consider adequate to support enforcement of the organizations' arbitration awards); Rogers, *supra* note 154, at 355 (describing how the AAA and other organizations have responded to public criticism with pro-consumer changes in arbitration procedures). For an example of the results of these changes, see the AAA's Consumer Due Process Protocol. *Consumer Due Process Protocol*, AM. ARB. ASS'N (Apr. 17, 1998), <http://adr.org/sp.asp?id=22019>. It makes some sense to focus on the AAA because it handles much of the arbitration in the United States. See ALAN SCOTT RAU ET AL., *ARBITRATION* 30 (3d ed. 2002) (outlining the "central role [of the AAA] in the administration of much of the arbitration that takes place in this country"). However, it is important to add that the significance of AAA data in this connection depends on the fraction of consumer arbitrations that the AAA handles. See Schwartz, *supra* note 136, at 1285–86 (making this point and stating that it is uncertain how much consumer and franchise arbitration goes through the AAA).

of fair procedures in order to avoid developing a reputation for unfair dealing. Although reputation markets do not always work well, they are more effective today with the Internet as a low-cost medium to spread reputational information.¹⁵⁶ Moreover, when sellers go too far and impose severely restrictive procedural terms, courts can handle the problems on a case-specific basis.¹⁵⁷

Finally, the problem with consumer contracts in many cases has less to do with consent or individual unfairness and more to do with the adverse effect of the procedural term on deterrence.¹⁵⁸ *AT&T Mobility LLC v. Concepcion* is a good example. In that case, wireless-service subscribers alleged that AT&T Mobility offered “free phones” to anyone who agreed to a services contract but then turned around and charged sales tax on the phones.¹⁵⁹ Each subscriber’s loss—the amount of the tax on a cell phone purchase—was too small to justify a separate suit.¹⁶⁰ Given the small potential losses, it would have been completely rational for even a fully informed consumer to accept the arbitration and class-waiver terms at the time of purchase.¹⁶¹

156. See DANIEL J. SOLOVE, *THE FUTURE OF REPUTATION: GOSSIP, RUMOR, AND PRIVACY ON THE INTERNET* 4 (2007) (explaining that reputational information used to be “scattered, forgettable, and localized” but is now “permanent and searchable” because of the Internet). Still, it is not always easy to sort between reliable and unreliable information on the web. Moreover, a consumer who searches for information about an arbitration association is likely to pick up the association’s home page at the top of the search results and might never look further if the home page makes sufficiently strong assurances.

157. In the arbitration setting, courts use the unconscionability doctrine to invalidate oppressive arbitration agreements. The recent Supreme Court decision in *Concepcion* imposed limits, but those limits still leave room for judicial monitoring of particularly abusive arbitration terms. See *AT&T Mobility LLC v. Concepcion*, 131 S. Ct. 1740, 1746, 1753 (2011) (abrogating California precedent deeming most class-arbitration waivers unconscionable as “an obstacle to the accomplishment and execution of the full purposes and objectives of Congress” in promulgating the Federal Arbitration Act (FAA) but reaffirming that the FAA “permits agreements to arbitrate to be invalidated by ‘generally applicable contract defenses, such as fraud, duress, or unconscionability’” when such defenses do not “apply only to arbitration or . . . derive their meaning from the fact that an agreement to arbitrate is at issue” (quoting *Doctor’s Assocs., Inc. v. Casarotto*, 517 U.S. 681, 687 (1996))). But see Jean R. Sternlight, *Tsunami: AT&T Mobility v. Concepcion Impedes Access to Justice*, 90 OR. L. REV. (forthcoming 2012) (manuscript at 7), available at <http://ssrn.com/abstract=1924365> (reporting that “a number of judges are extending” *Concepcion*’s reasoning beyond class action waivers to limit the use of the unconscionability doctrine to invalidate other arbitration terms). Moreover, the Court’s holding is based on preemption under the FAA, so it would not apply to adjudication. Thus, in adjudication, a judge could simply refuse to enforce an abusive contract on public policy or other grounds.

158. If one believes that a procedural modification is normatively undesirable on substantive grounds, it is easy to impute one’s beliefs to other individuals and then conclude that those individuals must be victims of bargaining inequality when they consent to the undesirable terms. But the problem is in fact substantive, and the criticism is actually about the social costs of party rulemaking. We shall examine those costs in section III(B)(3), *infra*.

159. *Concepcion*, 131 S. Ct. at 1744.

160. The total loss for the *Concepcions* was \$30.22. *Id.*

161. In addition, AT&T Mobility guaranteed a conditional payment that had the effect of making arbitration feasible for individual subscribers, and it also provided pro-consumer arbitration procedures. See *id.* at 1753 (noting that AT&T Mobility agreed to “pay claimants a minimum of

Thus, the problem in *Concepcion* had little to do with lack of consent.¹⁶² As in most consumer arbitration cases, the problem in *Concepcion* was that the class waiver eliminated the most effective device for aggregating individual claims and attracting lawyers to serve as private attorneys general to enforce deterrence goals.¹⁶³ In short, the criticism is actually about the external costs of the class action waiver and in particular the adverse impact on private enforcement of consumer protection laws.¹⁶⁴

The experience with arbitration shows that one must be careful about using a defective-consent argument against party rulemaking in adjudication. The conditions for valid consent should take account of how consent actually operates in ordinary litigation. Moreover, enforcing a party-rulemaking agreement might be beneficial to a party even when that party's consent is problematic *ex ante*. And an argument ostensibly based on defective consent can really be about adverse substantive effects resulting from limits to private enforcement of public law, or—as in the employment context—about the superiority of adjudication for deciding civil rights and other public law issues. I do not mean to suggest that there are no cases where consent is a problem. But there are not so many that party rulemaking should be banned altogether, even for adhesion contracts.¹⁶⁵ Thus, the defective consent argu-

\$7,500 and twice their attorney's fees if they obtain an arbitration award greater than AT&T's last settlement offer" and that the district court judge found this sufficient to incentivize individual proceedings).

162. Someone who holds a very strong view of consumer autonomy might still object, but this extreme position is difficult to hold without rejecting all adhesion and form contracts.

163. It is noteworthy that critics of the *Concepcion* decision emphasize its adverse effect on substantive law enforcement and not its impairment of meaningful consent. See, e.g., Myriam Gilles & Gary Friedman, *After Class: Aggregate Litigation in the Wake of AT&T Mobility v. Concepcion*, 79 U. CHI. L. REV. (forthcoming 2012) (manuscript at 5), available at <http://ssrn.com/abstract=1928071> (arguing that the Court's decision will end up scuttling private consumer class actions and encouraging state attorneys general to take up the enforcement slack); Sternlight, *supra* note 157, at 1–2 (decrying the effect of the Court's decision on substantive law enforcement). To be sure, weaker deterrence will adversely affect consumers in the long run, but no single consumer is likely to give much weight to that cost, especially as deterrence is a public good and consumers have incentives to free ride.

164. It is possible to frame this in terms of consent, but doing so requires a stretch. One might argue that as a practical matter the real party in interest is the class attorney since she has the largest stake and effectively controls the suit, and therefore her consent is needed to make a valid agreement. Alternatively, one might argue that the relevant consent is that of the consumer class as a whole, and class consent cannot be reduced to the sum of individual consents because of free-rider problems.

165. Over the past several years, Congress has entertained bills that would ban arbitration clauses in consumer and employment contracts, but none has yet been adopted. See, e.g., Arbitration Fairness Act of 2011, H.R. 1873, 112th Cong. (2011) (proposing to amend Title 9 of the U.S. Code to prohibit predispute arbitration agreements in employment, consumer, and civil rights disputes). One might expect a similar response to party rulemaking if it became more available. However, an outright ban is unwise even though more limited regulation might be desirable. See generally Sarah Rudolph Cole, *On Babies and Bathwater: The Arbitration Fairness Act and the Supreme Court's Recent Arbitration Jurisprudence*, 48 HOUS. L. REV. 457, 460–62, 467–69 (2011) (opposing the Arbitration Fairness Act's proposed ban on consumer arbitration but recommending more refined adjustments).

ment at most justifies judicial refusals to enforce agreements in specific situations.

2. *The Baseline Argument.*—Some commentators argue that as a policy matter, contracting should be allowed only when it is sanctioned in some way by existing procedural rules. There is a strong and a weak version of this argument. The strong version supposes that procedural contracting should be allowed only insofar as it is explicitly authorized by existing rules.¹⁶⁶ The weak version expands permissible contracting a bit further to include Type II rulemaking: on this view, parties should be able to contract in advance of litigation to accomplish what the rule system allows them to do during litigation, even when the Federal Rules do not explicitly authorize *ex ante* contracting.¹⁶⁷ In other words, they should be able to use contract to change the timing, but only the timing, of a procedural choice. To illustrate the difference, the strong version of the argument would bar enforcement of a pre-suit jury trial waiver because pre-suit waiver is not explicitly authorized by the Federal Rules of Civil Procedure or statute, whereas the weak version would allow enforcement of the waiver because parties can waive jury trial during the course of litigation.¹⁶⁸

Each version of the argument relies on problematic assumptions about the formal rule system. One assumption is that the existing system reflects an optimal balance of competing social policies and that contractual departures necessarily upset the balance.¹⁶⁹ The other assumption is that the process for making procedural rules has important features, such as uniformity and public debate, that are not shared by party rulemaking and that should be part of any method for making procedural rules.¹⁷⁰ Both assumptions are flawed.

166. See Taylor & Cliffe, *supra* note 22, at 1088 (arguing for explicit congressional authorization of any party rulemaking).

167. See Dodge, *supra* note 22, at 783–85 (referring to this approach of rulemaking as a “rule of symmetry”).

168. These arguments, therefore, are different than the Rules Enabling Act argument discussed above. See *supra* notes 97–105 and accompanying text. The claim that party rulemaking violates the Rules Enabling Act depends on a legal argument about the proper interpretation of the Act. The claims discussed in this section depend instead on the policies and values underlying the rulemaking process authorized by the Act.

169. See Taylor & Cliffe, *supra* note 22, at 1090–91 (contending that the rules of public dispute resolution strike a balance between parties and that this balance is “fine tune[d]” through continual adjustments); Thornburg, *supra* note 22, at 207–08 (“The government-created rules of procedure represent . . . the system’s best efforts to find a correct balance between fairness and efficiency . . .”); see also Dodge, *supra* note 22, at 766–67, 770 (arguing that, within the structure provided by the Federal Rules, the courts have already weighed procedural rights and balanced the competing public and private rights of the litigation process).

170. See Taylor & Cliffe, *supra* note 22, at 1099–100 (recounting the beneficial features of the formal rulemaking process, including increased deliberation, the opportunity for public debate, and uniformity).

As for the first assumption, it is highly implausible that the current system of procedural rules reflects an optimal policy balance. Federal Rules are frequently amended and revised, a fact that is hard to square with optimality. For example, Rule 26—prescribing basic discovery rules—was significantly revised in 1970 and amended again in 1980, 1983, 1993, 2000, and 2006.¹⁷¹ Rule 16, the pretrial conference rule, was revised in 1983 to officially recognize judicial involvement in settlement and again in 1993 to expand judicial settlement powers.¹⁷² And Rule 11, the sanctioning rule, was made stricter in 1983 and then relaxed in 1993.¹⁷³

In addition, courts have been very active in shaping civil procedure over the past thirty years. For example, the Supreme Court expanded the availability of summary judgment in 1986 by reinterpreting Rule 56 to alter doctrines reasonably well accepted at the time.¹⁷⁴ And in two recent decisions, the Supreme Court modified Rule 8(a)(2)'s liberal notice-pleading standard that had been in place for fifty years, ushering in the era of plausibility pleading.¹⁷⁵ This dynamic approach to the Federal Rules hardly seems consistent with a firm conviction that the existing rules are optimal.

Moreover, parties are sometimes (perhaps frequently) in a better position than rulemakers to design procedures for their own cases. Rulemakers necessarily aim their general rules at the typical or average case.

171. FED. R. CIV. P. 26 advisory committee's notes.

172. FED. R. CIV. P. 16 advisory committee's notes.

173. FED. R. CIV. P. 11 advisory committee's notes. The 1983 amendments to Rule 11 were intended to "reduce the reluctance of courts to impose sanctions . . . by emphasizing the responsibilities of the attorney and reinforcing those obligations by the imposition of sanctions." *Id.* The 1993 amendments "place[d] greater constraints on the imposition of sanctions [in order to] reduce the number of motions for sanctions presented to the court." *Id.*

174. *Compare, e.g., Anderson v. Liberty Lobby, Inc.*, 477 U.S. 242, 255–56 (1986) (importing the directed-verdict test into the summary judgment standard, including reference to the trial burden of persuasion), *Celotex Corp. v. Catrett*, 477 U.S. 317, 325 (1986) (rejecting the lower-court approach based on the holding in *Adickes v. S.H. Kress & Co.*, 398 U.S. 144 (1970), that placed "the burden . . . on the party moving for summary judgment to produce evidence showing the absence of a genuine issue of material fact"), and *Matsushita Elec. Indus. Co. v. Zenith Radio Corp.*, 475 U.S. 574, 596–98 (1986) (requiring that evidence presented "ten[d] to exclude the possibility" that an action was taken lawfully to defeat summary judgment (alteration in original) (quoting *Monsanto Co. v. Spray-Rite Serv. Corp.*, 465 U.S. 752, 764 (1984))), with *id.* at 598–600 (White, J., dissenting) (decrying the Court's "remarkable . . . departure from traditional summary judgment doctrine" by adopting a standard allowing judges to "decide for [themselves] whether the weight of the evidence favors the plaintiff").

175. *See Ashcroft v. Iqbal*, 129 S. Ct. 1937, 1949 (2009) (declaring that a properly pleaded complaint must plausibly suggest that the complaining party is entitled to relief); *Bell Atl. Corp. v. Twombly*, 550 U.S. 544, 557 (2007) (same). Congress has gotten involved in pleading reform as well. *See, e.g., Private Securities Litigation Reform Act of 1995*, Pub. L. No. 104-67, sec. 101(b), § 21D(b)(1)–(2), 109 Stat. 737, 746–47 (codified as amended at 15 U.S.C. § 78u-4(b)(1)–(2) (2006 & Supp. IV 2011)) (requiring a securities fraud complaint to state certain facts with particularity); *see also Tellabs, Inc. v. Makor Issues & Rights, Ltd.*, 551 U.S. 308, 324 (2007) (holding that § 78u-4(b)(2)'s pleading requirement is satisfied "only if a reasonable person would deem the inference of scienter cogent and at least as compelling as any opposing inference one could draw from the facts alleged").

Although rational parties also use averages when they contract *ex ante*, their average is taken over a much smaller domain since it focuses on those cases that might arise between the two of them. For this reason, the rules that parties choose in advance are likely to reflect superior information about their own future disputes.

This does not mean that party-selected rules are necessarily optimal from a social point of view. That depends, among other things, on the external effects of those rules. But it does mean that we should not blithely assume that the current system of procedural rules is superior to those that parties select.

The second assumption focuses on the superiority of the process for making procedural rules rather than on the superiority of the rules themselves. It is certainly possible to authorize party rulemaking through the formal rulemaking process, but this critique runs deeper. If party rulemaking conflicted with important values that should be honored by *all* rulemaking processes, even private bargaining in a Type III model, there would be reason to be concerned about allowing parties to make their own rules. The problem with this critique, however, lies in its assumption that formal rulemaking embodies values with the requisite force and reach.

The two most promising candidates for such values are uniformity and publicity.¹⁷⁶ Perhaps party rulemaking undermines the uniformity goal of the Federal Rules by creating different rules for different cases and threatens publicity and transparency values by delegating rulemaking to private bargaining.¹⁷⁷ But neither claim survives close scrutiny. The current Federal Rules are uniform only at a very high level of generality. They do not in fact create uniform procedure in particular cases; instead they delegate broad discretion to trial judges to tailor procedures to case-specific circumstances.¹⁷⁸

176. See Taylor & Cliffe, *supra* note 22, at 1100, 1103 (stressing the values of public debate and procedural uniformity and warning that private prelitigation agreements undermine these values).

177. See *id.* at 1100–04 (arguing that by enforcing a private prelitigation agreement that deviates from the Federal Rules, courts subvert the wisdom that was gained from public debate over the Rules, diverge from the uniformity the Rules were intended to provide, and potentially impart litigation advantages that contravene a “framework for dispute resolution that is fundamentally fair and evenhanded”); cf. Davis & Hershkoff, *supra* note 22, at 550–54 (arguing that a cost of “contract procedure” is that it allows parties to “end-run the public rulemaking process” and change the rules for private gain without sufficient public debate about the consequences).

178. See Bone, *Procedural Discretion*, *supra* note 7, at 1967–70 (giving examples of how the Federal Rules grant case-specific discretion). There are good reasons to be skeptical about a trial judge’s ability to use this discretion effectively. See *id.* at 1986–2001 (citing bounded-rationality constraints, information-access obstacles, and strategic-interaction effects). One of the problems—information-access obstacles—involves the difficulty of obtaining reliable, case-specific information from parties who, in the midst of adversarial battle, have strong incentives to be strategic about disclosure. See *id.* at 1996–2000 (discussing these incentives). Procedural contracting might be superior in this regard if it is able to harness the parties’ private information.

If the value of uniformity is consistent with judicial tailoring, it is not clear why it is not also consistent with party tailoring.¹⁷⁹

Moreover, the publicity value does not rule out all party rulemaking. In fact, party choices in the adjudicative setting become public and subject to public criticism when they are implemented.¹⁸⁰ If there are systemic problems, rulemakers can respond by amending the formal rules to regulate those problems. One might worry that the privacy of the bargaining process conceals bargaining defects. But the results of those defects are not concealed; they are visible when a judge applies the parties' chosen rule to pending litigation.¹⁸¹

Thus, one should not reject party rulemaking just because party-selected rules might diverge from the existing system or because the bargaining process lacks features characteristic of formal rulemaking. Whether and when it is a good idea to enforce contracts of this sort, especially those that implement Type III rulemaking, should depend on the costs and benefits.

3. *The Costs of Party Rulemaking.*—In this section, I discuss the most significant costs identified by critics of party rulemaking and conclude that none present a strong case to disallow it, except in a few specific circumstances.

a. *Third-Party Costs.*—One risk of allowing party rulemaking is that a party-selected rule might harm third parties not privy to the contract. These harms might include delay costs if party-made rules prolong litigation and increase case congestion, or they might include error costs if those rules affect the ability of third parties to recover on their claims.¹⁸² Although these are legitimate concerns, the problem with evaluating third-party effects lies in measuring the seriousness of the harms. Strategic behavior in ordinary litigation already harms third parties in multiple ways. When the existing parties prolong discovery or file summary judgment motions to delay

179. Uniformity with respect to substantive stakes—so-called transsubstantivity—is not necessarily a virtue. See Robert G. Bone, *Securing the Normative Foundations of Litigation Reform*, 86 B.U. L. REV. 1155, 1156, 1160–63 (2006) [hereinafter Bone, *Litigation Reform*] (criticizing the transsubstantivity principle).

180. In this regard, party rulemaking in adjudication differs from arbitration, which can remain secret throughout the entire process. See Brunet, *supra* note 117, at 8 (“The desire for secrecy can be a prime determinant in selecting arbitration.”).

181. For a discussion of bargaining inequality, see *supra* section III(B)(1). It is important to bear in mind, however, that settlement can make it difficult for judges to monitor bargaining defects if the settlement takes place before the chosen rule is implemented. Still, in many cases, at least some rules dealing with pretrial matters, such as pleading rules, are likely to be implemented before settlement. Moreover, there is no need for every case to go through trial for a pattern of abuse to become visible. Finally, the lawyer for a party disadvantaged by a sharply one-sided contractual rule might ask the court for relief from the rule and do so early in the case to avoid any adverse effect on settlement.

182. See, e.g., Kapeliuk & Klement, *Contractualizing Procedure*, *supra* note 21, at 38–44 (describing negative externalities of private, contractualized procedure).

litigation, for example, their choices are likely to impose delay costs on third parties in other suits. When a plaintiff exercises her right not to include others as co-plaintiffs in her suit, the result can adversely affect those not joined by delaying their suits or creating stare decisis effects.¹⁸³ The important question therefore is not whether party-chosen rules might harm third parties but instead whether, and by how much, those rules are likely to exacerbate the harmful effects that already exist.

The answer is unclear. In fact, some forms of party rulemaking might mitigate third-party harms. Parties usually have incentives to choose rules that reduce the duration or cost of their future suits, and in many cases, these rules will also reduce delay costs for other litigants. For example, parties might impose limits on discovery in order to save costs and control abuse.¹⁸⁴ If those limits reduce the time to disposition, they will also marginally reduce the delay costs in other suits. Sometimes party choices might have the opposite effect, such as when limiting discovery unexpectedly reduces settlement incentives and increases the risk of a time-intensive trial.¹⁸⁵ But identifying the problem cases is bound to be very difficult.

To be sure, there are cases where party rulemaking is much more certain to adversely affect third parties. To illustrate, suppose *A* and *B* agree not to complicate future litigation between them by joining additional parties.¹⁸⁶ Suppose that a suit arises in which the only source of recovery is *B*'s insurance policy, which is limited in amount. Another person, *C*, has also suffered injuries as a result of *B*'s actions, and the policy proceeds are insufficient to cover all of *B*'s potential liability to *A* and *C*. This is a classic limited-fund situation.¹⁸⁷ If the court enforces the pre-suit agreement between *A* and *B* and allows *A* and *B* to litigate without joining *C*, then *A* might exhaust the entire limited fund and leave *C* out in the cold.

183. Potential stare decisis effects usually do not trigger compulsory joinder under Rule 19 or support intervention as of right under Rule 24. *See, e.g., Janney Montgomery Scott, Inc. v. Shepard Niles, Inc.*, 11 F.3d 399, 407 (3d Cir. 1993) (holding that creation of the “persuasive effect of a district court decision” on an absent party’s rights was insufficient to make joinder compulsory under Rule 19); 4 JAMES WM. MOORE ET AL., *MOORE’S FEDERAL PRACTICE* § 19.03(3)(e) (3d ed. 2011) (noting that stare decisis can be enough to compel joinder under Rule 19, but only if the issue is truly difficult and likely to be reviewed on appeal); 7C CHARLES ALAN WRIGHT, ARTHUR R. MILLER & MARY KAY KANE, *FEDERAL PRACTICE AND PROCEDURE* § 1908.2, at 369–74 (3d ed. 2007) (showing that intervention under Rule 24 is not available in every case where the decision may have stare decisis effect).

184. *See supra* notes 109–11 and accompanying text.

185. *See supra* note 115 and accompanying text.

186. They might do this for various reasons, such as to limit the costs of a future lawsuit, simplify settlement bargaining, or reduce the time to trial. Joinder agreements of this sort also might have adverse effects on delay costs by increasing the number of separate suits and thus case backlog, but predicting these effects would be very difficult.

187. *See* FLEMING JAMES, JR. ET AL., *CIVIL PROCEDURE* § 10.12, at 612 (5th ed. 2001) (noting that the “necessary party rule” can apply to a suit seeking a judgment distributing a fund when the claim of a nonparty on the fund “may as a practical matter be worthless”); STEPHEN C. YEAZELL, *CIVIL PROCEDURE* 771 (6th ed. 2004) (listing the limited-fund scenario as one of the common situations for application of Rule 19).

In this hypothetical, the contract between *A* and *B* imposes a negative externality on *C*. But this problem is hardly unique to procedural contracting. If it is in the interests of *A* and *B* to litigate without *C*, they can do so under the current rules simply by choosing not to join *C* (as long as these strategies support an equilibrium).¹⁸⁸ In fact, *A* would have an incentive to pay *B* a portion of the surplus created by not joining *C* if that was necessary to induce *B*'s cooperation. The point is that it is unclear whether party rulemaking, especially rulemaking before a dispute arises, adds significantly to the external costs that would otherwise exist.

In addition, under the current mandatory-joinder rule applicable in federal court, the plaintiff has a duty to identify all persons who are required to be joined if feasible¹⁸⁹ to avoid externalities. And the trial judge has power to require joinder of any such person *sua sponte*, even when the parties prefer to litigate without the person.¹⁹⁰ The same solution to the problem can be used for party rulemaking. Whenever a party-made procedural rule threatens serious harm to identifiable third parties, the judge could require joinder of the third party or refuse to enforce the parties' agreement. Indeed, as we saw in Part II above, judges today have power to review procedural agreements to make sure that they are not seriously unfair.¹⁹¹ It might be difficult to identify everyone who is seriously affected, but this same problem exists under the current system.

b. Public Litigation Costs.—Another problem with party rulemaking stems from the fact that private litigation incentives are not necessarily socially optimal because private parties do not internalize the public costs of the court system.¹⁹² Thus, the parties might agree to rules that reduce their private costs but increase public costs. The problem with evaluating this argument is the same as the problem with evaluating the argument from third-party effects. It is extremely difficult to identify cases where party rulemaking generates costs substantially in excess of those already created by the current system.¹⁹³

Indeed, parties have such wide latitude to engage in costly strategic behavior currently that it is not evident why adding another strategic option—*ex ante* contracting—will increase costs. For example, *A* and *B*

188. The result could be achieved noncooperatively in this case if both parties had an interest in cutting costs, facilitating settlement, or expediting trial. In fact, it is difficult to see why *B* would raise the joinder objection in this scenario except perhaps as a threat to force settlement (in which case *B* is not likely to carry through on the threat if it benefits from a simpler case as well).

189. FED. R. CIV. P. 19(a), (c).

190. See FED. R. CIV. P. 19(a)(2) ("If a person has not been joined as required, the court must order that the person be made a party.").

191. See *supra* notes 94–95 and accompanying text.

192. See Kapeliuk & Klement, Contractualizing Procedure, *supra* note 21, at 39–42 (discussing the inefficient overuse of public judicial resources by litigants).

193. See *id.* at 41 (describing the difficulty of identifying the inefficiencies created by contractualized procedures).

might agree to rules that signal type and deter frivolous filings but cost more to administer. One might expect those rules to increase public litigation costs, but they could reduce public (and private) costs by discouraging frivolous filings or facilitating settlements in meritorious suits.¹⁹⁴ Of course, the opposite effect is possible too. For example, parties might agree in advance to limit discovery in order to save costs and prevent abuse. But lower anticipated discovery costs and limited access to information could end up reducing the likelihood of settlement and increasing the risk of trial, thereby adding public costs.¹⁹⁵ The important general point here is that the complexity of intense strategic interaction makes it difficult in many cases to determine whether party rulemaking will increase or reduce public litigation costs relative to the ordinary litigation baseline.

One might worry about the added administrative costs when judges have to interpret contracts to determine what procedural rules parties have selected.¹⁹⁶ But administrative costs are already high in the current system, and they might actually decline with party rulemaking. District judges now spend considerable time and energy crafting procedures for individual cases, and the Federal Rules of Civil Procedure depend on this case-specific discretion.¹⁹⁷ If party agreements replace open-ended Federal Rules with clearer alternatives—as one might expect, given party incentives to reduce process cost and risk—these administrative costs could be reduced as well.

c. Outcome-Quality Costs.—Many commentators worry about the impact of party rulemaking on the quality of outcomes.¹⁹⁸ If parties contract for procedural rules that increase or skew the risk of error—such as rules constraining discovery, restricting witness testimony, or limiting appeals—those rules can adversely affect the outcomes in both the parties' suit and in future suits that rely on the earlier suit as a baseline for settlement

194. See *supra* note 114 and accompanying text.

195. See *supra* notes 115–16 and accompanying text. Moreover, if one party seeks broader discovery *ex post*, the court will have to determine whether the agreement was breached, and this will add public as well as private litigation costs. Of course, each party should also take account of the additional expected costs of trial. However, parties only consider the private costs of trial and not the public costs. Moreover, the parties usually bear most of the discovery costs, except when motions to compel are frequent. Therefore, limited discovery might be optimal for the parties when it is not optimal from a social perspective.

196. See Moffitt, *supra* note 21, at 514–15 (addressing this objection).

197. See Bone, *Procedural Discretion*, *supra* note 7, at 1967–75 (describing the range of procedural discretion and the history of heavy reliance on it).

198. See, e.g., Taylor & Cliffe, *supra* note 22, at 1103–04 (noting that prelitigation agreements can be used to gain an unfair strategic advantage); Thornburg, *supra* note 22, at 209 (contending that while it might make sense for courts to enforce procedural contracts when they are sure the provisions were genuinely bargained for, it is too difficult to be sure and as a result these contracts should not necessarily be entitled to specific performance); Kapeliuk & Klement, *Contractualizing Procedure*, *supra* note 21, at 44–46 (explaining the potential fairness and efficiency concerns when parties enter into predispute procedural arrangements unaware of the likely consequences).

valuation.¹⁹⁹ For example, one party might be more powerful and better informed than the other and thus able to impose a contract with procedural rules that skew the error risk in its favor. Or the two parties might have equal power and choose rules that reduce costs but increase the risk of error. If there are enough such contracts and the chosen procedures produce erroneous outcomes, the compensation and deterrence goals of the substantive law might be impaired.

This is a legitimate concern, but one has to be careful about evaluating its significance. The current litigation system also produces bad outcomes. Parties do not always get access to critical information through discovery, hire competent experts at trial, and appeal. Given the risks of the current system, it is hard to tell how much of an additional risk party rulemaking would add.

Moreover, when parties have roughly equal bargaining power, neither can insist on a rule that favors one party over the other. Therefore, they are likely to choose rules that affect the variance of the error-risk distribution symmetrically and avoid skewing it to the disadvantage of one side. A symmetric increase in error risk will not affect expected recovery and thus should leave the incentives of risk-neutral actors relatively undisturbed. This is so because risk-neutral actors average over past cases when they predict the potential liability from engaging in a particular course of action or estimate the value of a future case for purposes of settlement. Averaging cancels errors on the high side of the mean against errors on the low side, leaving the mean undisturbed.²⁰⁰

The situation is different when one party has much more bargaining power than the other. The former can insist on a contract with procedural rules that skew the error risk in its favor. For example, a large corporation might use a form contract to foist unfavorable procedural rules onto a consumer. If litigation ensues, the unfavorable rules might produce an outcome improperly skewed in the corporation's favor. This is a serious concern, but it too should not be exaggerated. As section III(B)(1) explained, the contract might actually make the consumer better off if the corporation passes along some of its savings in the form of a lower price. Moreover, the market can deter some of the more abusive practices through competition and reputation effects.

In addition, it is important to bear in mind that parties with more power already enjoy substantial litigating advantages that skew outcomes improperly. To be sure, a weaker party is represented by a lawyer during

199. The latter effect is particularly important because most cases end in settlement. See Marc Galanter, *The Vanishing Trial: An Examination of Trials and Related Matters in Federal and State Courts*, 1 J. EMPIRICAL LEGAL STUD. 459, 459 (2004) (reporting data showing that in 2002 the portion of federal civil cases resolved by trial was only about 2%).

200. See BONE, CIVIL PROCEDURE, *supra* note 44, at 20–29 (explaining expected value and the rational-choice model). However, a mean-preserving increase in the error risk might have an adverse impact on risk-averse parties.

litigation but probably is not when a contract is formed prior to suit. However, one should not exaggerate the benefits of legal representation. In the cases of greatest concern, the weaker party is likely to be the plaintiff and represented by a lawyer hired on contingency or by a class action attorney. In either situation, there are serious risks of high attorney–client agency costs, which can undermine the attorney’s effectiveness in assuring a good outcome.²⁰¹ Finally, as I explain more fully below, the importance one assigns to the risk of skewed outcomes depends on whether one takes a utilitarian or a rights-based approach.²⁰²

Another way that party rulemaking can have an adverse impact on future suits is by affecting the quality of legal precedent.²⁰³ Any evaluation of these effects, however, depends on one’s theory of adjudication and, in particular, on what constitutes a good decision as well as a good decision-making process.²⁰⁴ On the one hand, for example, one might view decisions on legal issues as an abstract exercise of reasoning from extant legal materials without regard to case-specific facts. From this perspective, party-chosen rules that limit discovery or restrict access to experts should have little impact. Constraints on appellate review might have an adverse effect over the long run, but the same is true under the current system, given that most cases settle and therefore are not subject to review.²⁰⁵ On the other hand, one might view judicial decision making as intimately connected in some way to the facts of each particular case. On this view, rules that limit or distort the case-specific information available to the judge could have an adverse effect

201. See, e.g., Macey & Miller, *supra* note 144, at 19–27 (discussing the serious agency problems in small-claim class actions). See generally Geoffrey P. Miller, *Some Agency Problems in Settlement*, 16 J. LEGAL STUD. 189 (1987) (discussing agency problems created by claim-sharing arrangements in the settlement context).

202. See *infra* subpart III(C).

203. Some critics of party rulemaking mention this concern. See, e.g., Thornburg, *supra* note 22, at 209 (arguing that a court’s decision whether to enforce party-made rules “might be an easy ‘yes’ if we could be sure that . . . the outcome of the dispute would only affect those parties”). Moreover, it figures prominently in the literature critical of the judicial focus on settlement, of ADR, and of the Supreme Court’s strongly favorable attitude toward arbitration. See, e.g., Owen M. Fiss, *Against Settlement*, 93 YALE L.J. 1073, 1089 (1984) (noting that settlement is a “poor substitute” for judgment, which is fundamental to “bring[ing] a recalcitrant reality closer to our ideals”); David Luban, *Settlements and the Erosion of the Public Realm*, 83 GEO. L.J. 2619, 2641 (1995) (describing a lack of legal precedents as an unfortunate consequence of a world dominated by settlement). My discussion here focuses on the *quality* of precedents. There is no reason to believe that party rulemaking will have a seriously adverse effect on the *quantity* since it is not obvious that party-made procedure would generate more settlements than the current procedural system does. In any event, it is not clear how much precedent is optimal, and without that baseline, it is impossible to tell whether the actual quantity is deficient. For a critical take on the precedent argument in the context of arbitration, see Hylton, *supra* note 149, at 243–47.

204. For example, some commentators praise international commercial arbitration for producing good precedents notwithstanding its procedural limitations. E.g., Rogers, *supra* note 154, at 370–71.

205. See *supra* note 199.

on the quality of judicial decisions over the long run, although it is unclear how serious it would be. I shall return to this problem in Part IV below.

d. Judicial Legitimacy.—It is quite common for critics of party rulemaking to cite the potential risks to judicial legitimacy.²⁰⁶ One common example imagines parties agreeing that their future cases should be decided by a judicial flip of a coin.²⁰⁷ Unfortunately, those making this argument do not carefully define what they mean by legitimacy. They might believe that a process is legitimate if it produces a reasonably good outcome subject to cost and fairness constraints. However, this view of legitimacy imposes no additional constraints beyond those already discussed. In particular, it cannot rule out the use of a coin flip in all cases. If the parties genuinely consent to coin flipping and its attendant risks, and if using a coin flip does not impose serious externalities on third parties (or at least none more serious than already exist under current procedures), then it is difficult to see how someone holding this view of legitimacy could object.

There is, however, a stronger version of the argument from legitimacy. The stronger version supposes that certain procedures are off-limits even when parties consent to use them and no third-party litigants are harmed by them. There are two versions of the stronger legitimacy critique. One version focuses on perceived legitimacy and the other on normative legitimacy.²⁰⁸ Perceived legitimacy is concerned with whether the public perceives the court system as legitimate; normative legitimacy is concerned

206. See, e.g., Davis & Hershkoff, *supra* note 22, at 547–51 (discussing risks to perceived legitimacy as well as other legitimacy-related concerns); Taylor & Cliffe, *supra* note 22, at 1090 (assuming that legitimacy would be impaired if the judge had to honor the parties' agreement to flip a coin); Thornburg, *supra* note 22, at 207–09 (lumping a number of different arguments under the theme of preserving the legitimacy and integrity of the courts); see also Dodge, *supra* note 22, at 764–70 (discussing concerns of legitimacy and judicial integrity); Kapeliuk & Klement, *Contractualizing Procedure*, *supra* note 21, at 42–44 (discussing potential adverse effects on a court's reputation capital as a constraint on party rulemaking).

207. See, e.g., Sarah Rudolph Cole, *Managerial Litigants? The Overlooked Problem of Party Autonomy in Dispute Resolution*, 51 HASTINGS L.J. 1199, 1225 (2000) (arguing that an agreement to decide a dispute by flipping a coin would “undermine the integrity of the court as an institution by making it appear that courts exist to serve the whims of litigants”); Taylor & Cliffe, *supra* note 22, at 1090 (noting that “a court would not enforce an agreement to resolve a dispute by judicial coin toss” because it is “simply too ridiculous and makes a mockery of why the court is there”); see also *LaPine Tech. Corp. v. Kyocera Corp.*, 130 F.3d 884, 891 (9th Cir. 1997) (Kozinski, J., concurring) (noting that he would refuse to enforce an arbitration clause that “provided that the district judge would review the [arbitration] award by flipping a coin or studying the entrails of a dead fowl”), *overruled on other grounds sub nom. Kyocera Corp. v. Prudential-Bache Trade Servs., Inc.*, 341 F.3d 987, 1000 (9th Cir. 2003) (en banc); cf. *United States v. Josefik*, 753 F.2d 585, 588 (7th Cir. 1985) (noting that an agreement to trial by twelve orangutans would be invalid). One might wonder why parties would not just flip the coin themselves, but maybe they do not trust one another to do it fairly.

208. See generally Bone, *Rethinking*, *supra* note 123, at 233–36 & n.155 (distinguishing perceived from normative legitimacy and noting the prevalence of perceived legitimacy arguments in the nonparty-preclusion setting).

with whether a procedure is actually legitimate on normative grounds independent of public perception or belief.²⁰⁹

The critique regarding perceived legitimacy is the more common of the two among those who discuss party rulemaking. They assume that the public will perceive courts as acting illegitimately and lose faith in adjudication if judges are forced to abide by agreements that require them to take bizarre actions.²¹⁰ Sometimes the argument is presented in terms of potential costs to judicial “reputation,” but it amounts to the same thing. Public perception is what matters and party rulemaking can lead to bad perceptions.²¹¹

The argument is weak when framed in this way. The first problem is that perceptions are malleable. Take the example of the coin flip. Much of its force depends on an implicit assumption that deciding cases by flipping a coin would be purely arbitrary, and as a result, the public would perceive coin flipping as offensive to a system of adjudication based on reasoned deliberation.²¹² But the assumption is flawed. Coin flipping would not be an arbitrary choice if it were the choice that satisfied the parties’ preferences.²¹³ This could matter to the public’s reaction. A member of the public might respond differently if she knew that it was the parties who chose the coin flip and understood the reasons why they did. In other words, information about reasons might make the method seem less arbitrary and elicit a more positive response.

Second, public perceptions are circular. People come to believe that a procedure is inappropriate for adjudication because it conflicts with what

209. *Id.* at 233, 236.

210. *See, e.g.,* Cole, *supra* note 207, at 1225 (“If courts begin to make decisions in an arbitrary manner, . . . the respect the public currently has for the judiciary as a decision-maker will be dissipated.”); Moffitt, *supra* note 21, at 509–11 (“Society has an important interest in preserving the public perception that courts are legitimate.”). Some commentators do not clearly state whether they are concerned about perceived or normative legitimacy, but insofar as they focus on the way individuals would react to a procedure upon seeing it, they seem to be concerned with perceived legitimacy. *See, e.g.,* Noyes, *supra* note 21, at 625 (noting that bizarre procedures would “make the court look silly or incompetent”); Taylor & Cliffe, *supra* note 22, at 1090 (worrying that enforcing unusual procedural agreements could make a “mockery” of the court); Thornburg, *supra* note 22, at 207–08 (arguing that court “ritual[s] legitimize[] legal authority” and cashing this out in terms of symbolic and educational benefits).

211. *See, e.g.,* Kapeliuk & Klement, *Contractualizing Procedure*, *supra* note 21, at 42–44 (discussing ways in which enforcement of procedural contracts might affect a court’s “reputational capital”).

212. *See* JON ELSTER, *SOLOMONIC JUDGEMENTS: STUDIES IN THE LIMITATIONS OF RATIONALITY* 102 (1989) (noting that randomness in legal decisions is often associated with arbitrariness or whimsy); Shay Lavie, *Reverse Sampling: Holding Lotteries to Allocate the Proceeds of Small-Claims Class Actions*, 79 *GEO. WASH. L. REV.* 1065, 1082 & n.72 (2011) (noting that people oppose lotteries because they substitute luck for reason).

213. Indeed, flipping a coin can be justified on moral grounds when it is impossible to tell which party is correct and both have equally strong substantive entitlements. *See* Lewis A. Kornhauser & Lawrence G. Sager, *Just Lotteries*, 27 *SOC. SCI. INFO.* 483, 495–503 (1988) (discussing equal-entitlement and scarcity conditions for using the lottery as an exclusive or nonexclusive method of allocation and noting that using the lottery under these conditions is supported by reasons).

they believe courts *ought* to do. Thus, the answer to the normative question of what courts should and should not do affects the answer to the empirical question of how the public will react when courts do something they should not. If judges routinely flipped coins, for example, public opinion might well shift toward accepting coin flipping as a proper decisional method. But the shift in public opinion by itself would not make coin flipping normatively legitimate.²¹⁴

In fact, complaints about perceived legitimacy are often disguised complaints about normative legitimacy. The critic starts with an intuitive belief that the particular procedure is normatively illegitimate or unfair, imputes that belief to the public, and then concludes that the public will perceive the courts as illegitimate if the procedure is implemented. By framing a normative intuition as an empirical claim about public perceptions, the critic never has to confront the theoretical basis for the underlying intuition.

This last point is crucial. The legitimacy that matters is not perceived legitimacy but normative legitimacy. Admittedly, there would be reason to worry if party rulemaking produced weird rules so frequently that the public lost faith in the court system. But this is not a realistic risk. Parties are not likely to choose extremely bizarre rules; there is no reason for them to do so. Moreover, it is simply implausible to think that the public would give up on the courts if judges occasionally decided cases according to unusual procedures, especially if they had good reasons for doing so.²¹⁵

The conclusion is clear. If it is normative legitimacy that matters, the critics of party rulemaking must ground their arguments in a theory of adjudicative legitimacy. Part IV takes up that challenge.

C. *Balancing the Competing Considerations*

Any decision whether to enforce procedural agreements should depend on the balance of benefits and costs. The previous analysis, however, highlights a serious problem with striking this balance. The strategic environment of litigation is far too complex to make confident predictions. An agreement to restrict procedural options might reduce the public costs of the court system, but it also might increase them. An agreement to expand procedural options might increase public costs, but it also might reduce them.

214. Even a moral conventionalist, who believes that moral principles are those that the society in general accepts as moral, is likely to focus on well-considered beliefs about legitimacy, not raw perceptions. See generally THE ENCYCLOPEDIA OF PHILOSOPHY: SUPPLEMENT 103-04, 361-65 (Donald M. Borchert ed., 1996) (defining *conventionalism* as the view that truth is a matter of social convention and describing three competing metaethical theories of morality, all of which focus on careful consideration rather than crude perception).

215. For example, a judge might flip a coin in cases of genuine indeterminacy or party agreement. See *supra* note 213 and surrounding text; see also Davis & Hershkoff, *supra* note 22, at 547-48 (recognizing that arguments based on perceived legitimacy are "necessarily speculative" and depend on the circumstances).

It is possible that private procedural choices might significantly affect outcome quality in a way that harms third parties or creates other negative externalities, but strategic litigation without modified procedures produces similar results, especially in a world of settlement.

The indeterminacy of the cost-benefit balance has important implications for party rulemaking. These implications vary to some extent depending on whether one applies a utilitarian or rights-based analysis.²¹⁶ In a utilitarian approach, at least one with a law and economics bent, all the benefits and costs, private and public, are balanced. The goal is to minimize social costs (or maximize social welfare) in the aggregate.²¹⁷ Some commentators who follow this approach assume that judges can strike the social cost-benefit balance in individual cases and make sound case-specific decisions about whether to enforce the parties' agreement in the face of *ex post* opposition.²¹⁸ However, this proposal is impractical. A judge in an individual case lacks the information and expertise to make highly complex predictions about case-specific benefits and costs.

The analysis proceeds a bit differently under a rights-based approach. The goal in rights-based theory is not to maximize social welfare but rather to furnish each individual what her right guarantees, even when doing so reduces overall welfare.²¹⁹ More precisely, a right constrains, checks, or trumps justifications for limiting what the right guarantees when those justifications rely on increasing social welfare in the aggregate.²²⁰

Balancing is still appropriate in a rights-based theory, but social costs and benefits play a much more limited role.²²¹ How this balance works for

216. See Robert G. Bone, Twombly, *Pleading Rules, and the Regulation of Court Access*, 94 IOWA L. REV. 873, 910–15 (2009) [hereinafter Bone, *Regulation of Court Access*] (describing the difference between a utilitarian and a rights-based theory focused on outcome quality); see also Bone, *Rethinking*, *supra* note 123, at 237–64 (applying utilitarian and rights-based theories to evaluate nonparty preclusion rules).

217. Bone, *Regulation of Court Access*, *supra* note 216, at 910.

218. See, e.g., Kapeliuk & Klement, *Contractualizing Procedure*, *supra* note 21, at 42 (“[W]e suggest that in evaluating whether to enforce pre-dispute procedural arrangements that may impact judicial time and costs, courts should be aware of the possible tradeoff between pre-dispute and post-dispute costs and benefits of such arrangements.”); *id.* at 49 (“[C]ourts should be careful not to overturn efficient modifications.”).

219. See Robert G. Bone, *Procedure, Participation, Rights*, 90 B.U. L. REV. 1011, 1012 (2010) [hereinafter Bone, *Procedure, Participation, Rights*] (describing this conception of a right); Bone, *Regulation of Court Access*, *supra* note 216, at 912 (same).

220. Here, I gloss over the distinctions between two different versions of rights-based theory: outcome-oriented and process-oriented. An outcome-oriented theory values procedure instrumentally, as a means to the end of producing judgments and settlements that conform to the parties' substantive rights. A process-oriented theory values procedure intrinsically, as a way to respect the autonomy and dignity of individuals by guaranteeing them an opportunity to participate in litigation that affects them personally. See Bone, *Agreeing to Fair Process*, *supra* note 43, at 508–16 (describing the difference between outcome-based and process-based rights theories).

221. See Bone, *Procedure, Participation, Rights*, *supra* note 219, at 1013 (“Because they are rights, they must resist arguments for limiting procedure based on the high social costs of litigation,

procedural rights is an extremely complex matter and beyond the scope of this Article.²²² One thing is clear, however. Consent does more work in a rights-based theory because valid consent is the ultimate justification for enforcing party-rulemaking agreements *ex post*. By contrast, in a utilitarian theory, the ultimate justification for enforcement remains the social cost-benefit balance, and consent is simply a proxy for potential Pareto improvement. Moreover, insofar as costs and benefits matter to a rights-based theory, *private* costs and benefits to the right holder should receive much greater weight than public costs and benefits because of the focus on respecting the individual's right. There is an exception, however, when party choice seriously risks impairing the procedural rights of third parties. In such a case, the conflicting rights must be accommodated.

Unless consent is problematic, a rights-based approach should raise no obstacle to enforcing an agreement.²²³ This does not necessarily mean, however, that all agreements should be enforced. Because procedural rights impose minimum requirements, they have nothing to say one way or the other about party choices that are consistent with the minimum but go further.²²⁴ It follows that the decision whether to implement rules created by party rulemaking depends first on whether consent is valid, then on whether third-party rights are threatened, and finally on whether the social costs and benefits justify enforcing the parties' choice.

With this background in place, it is easy to see that evaluation and prediction are problems in a rights-based theory as well. Public costs and benefits are still relevant to the analysis when the parties agree to rules that extend beyond the rights-based floor, and as we have seen, these factors are very difficult to predict. The same is true for the likely impact on third parties.

Still, there are three relatively clear situations that should trigger concern under both a utilitarian and a rights-based approach.²²⁵ First, when

but to fit prevailing intuitions of procedural fairness, they must also yield to social cost arguments, at least to some significant degree.”).

222. I have argued elsewhere that social costs should have more weight in limiting the scope of procedural rights than the scope of other kinds of rights. Bone, *Agreeing to Fair Process*, *supra* note 43, at 513–17; Bone, *Procedure, Participation, Rights*, *supra* note 219, at 1015–18.

223. This assumes that parties are free to trade their procedural rights. Obviously, they can do so during litigation. After all, parties trade procedural rights when they settle, and most lawsuits end in settlement. *See supra* note 199. Moreover, there is no apparent reason to treat the matter differently when parties bargain before a dispute arises. Settlement involves *ex ante* trading, just as pre-suit agreements do. When a party settles a pending suit, she agrees to trade her procedural rights in advance of the time she would otherwise exercise them. One might argue that parties have better information as well as attorney representation after a suit is filed. But information varies continuously over time and representation is not necessarily confined to litigation, so there is no reason to draw a sharp line at the point a dispute arises or a suit is filed.

224. This assumes that the baseline itself satisfies rights.

225. I do not mean to suggest that these are the only situations that should be considered. They are, however, the clearest cases. For example, one might also object to courts enforcing an agreement as part of a settlement that requires the parties to maintain the confidentiality of

the parties expressly agree to exclude an identifiable third party with legal rights that might be seriously affected, the judge should refuse to enforce the agreement and join the third party if the judge decides that the mandatory-joinder rule requires it. Second, the judge should refuse enforcement of a sharply one-sided agreement in a market with little competition or under circumstances where it is clear that one party would not have accepted the onerous terms if given a chance to bargain, all on the theory that party consent and private benefits are questionable. Third, the judge should refuse enforcement if she is convinced that the agreed-upon procedures are inadequate for proper consideration of statutory or constitutional civil rights claims, such as in the employment context, or if she is convinced that those procedures seriously interfere with private enforcement of public law, such as when parties agree to exclude class actions in cases that are likely to involve small stakes and a substantive claim that Congress has assigned to private enforcement.

Beyond these three situations, however, it is not clear what to do in light of the problems with prediction and evaluation. One might ban all party rulemaking on the conservative theory that no deviations from the baseline system should be allowed without clear justification. However, this seems excessively restrictive. In fact, it makes sense to give broad scope to party rulemaking as long as defective consent is not a serious impediment. Party welfare is improved (at least *ex ante*), and the public cost-benefit balance, given its indeterminacy, does not clearly weigh against—and might even weigh in favor of—enforcement.

The case for broad enforcement of party-made rules is particularly strong for Type II rulemaking.²²⁶ The fact that the parties can take the same action during litigation means that their choice is not likely to increase public costs significantly, compared to the default system, and also not likely to impair adjudicative legitimacy. Therefore, as long as consent is not seriously suspect, *ex post* enforcement of *ex ante* commitments should enhance welfare and respect rights. The case for enforcement is not as clear for Type III rulemaking. Nevertheless, given that the parties usually have better information about their cases than the court, allowing parties to customize rules could reap substantial benefits.²²⁷

potentially damaging discovery materials and asks the judge to seal any court records that might reveal those materials. See, e.g., Richard L. Marcus, *The Discovery Confidentiality Controversy*, 1991 U. ILL. L. REV. 457, 502–05 (discussing the objections and concluding that they do not support a general rule against enforcement). Agreements of this sort can be treated as a form of procedural rulemaking insofar as they alter a default rule of public access. It is far from clear, however, that a general rule barring enforcement in all cases is justified. *Id.*

226. See Dodge, *supra* note 22, at 783–85 (making the same point).

227. Other commentators have reached similar conclusions, but mostly on the basis of a utilitarian approach. See, e.g., Hylton, *supra* note 149, at 212–13 (“Although I note important exceptions along the way, the principal argument of this paper is that waiver and arbitration agreements should be enforced . . . as long as they have been entered into knowingly and voluntarily.”); Shavell, *supra* note 109, at 3 (“Because *ex ante* ADR agreements made by

Moreover, there is an alternative to case-specific evaluation with all its prediction problems. The formal rulemaking process can be used to create general rules regulating party rulemaking. This formal process has advantages over case-specific assessment.²²⁸ Rulemaking committees are able to evaluate effects for general categories of cases; they have access to empirical information not available to judges, and they can take a global perspective necessary to evaluate complicated intrasystem effects. The general rules would have to rely on criteria that judges can assess relatively easily in specific cases, but there is no a priori reason to believe that this cannot be done. Prediction, though easier, will still be difficult, and general rules are under- and over-inclusive. But the same is true for the current Federal Rules of Civil Procedure.

This analysis suggests that the arguments against party rulemaking, even Type III rulemaking, are relatively weak on cost-benefit and rights-based grounds, at least beyond a small set of clearly problematic cases. One might conclude from this that parties should be given wide latitude to design their own procedures and rules. But this conclusion grates against some strong intuitions. The critics have a point when they say that there are procedures, like a coin flip, that seem sharply at odds with what courts are supposed to do even when those procedures serve the parties' private interests, reduce litigation costs, and produce no serious externalities to identifiable third persons. Efforts to capture this intuition in terms of external costs to the court system, whether in the form of harm to reputation or negative public perception, miss the point. The objection is not that there are other costs to include in the cost-benefit balance. The objection is that some party choices should be categorically forbidden because what is chosen is not appropriate for adjudication.

If this is correct, it means that the case against party rulemaking cannot be mainly about cost-benefit balancing or bargaining-power inequities. Instead it must be based on an independent argument from adjudicative legitimacy. The following discussion explores that argument.

IV. The Legitimacy Critique

As we saw in the previous part, any serious legitimacy problem with party rulemaking must be based on normative legitimacy, and normative legitimacy does not depend on public perceptions or judicial reputation. If some exercises of party rulemaking threaten legitimacy, it must be because the particular rules that the parties have chosen alter core elements of adjudication that are essential to its institutional legitimacy. The challenge is to

knowledgeable parties raise their well-being, the agreements raise social welfare (in the absence of external effects). Thus, it is suggested that ex ante ADR agreements should ordinarily be enforced by the legal system.”).

228. For an account of the relative advantages of the formal rulemaking process, see Bone, *The Process of Making Process*, *supra* note 13, at 920–26.

identify these core elements, and to do that, one needs a theory of adjudication. This Article is not the place to elaborate a comprehensive theory—an undertaking that would, of course, be extremely complex. The following discussion focuses on the element of adjudication that I believe is most central to understanding the appropriate limits on party rulemaking: the mode of judicial reasoning. Party rulemaking is most problematic when it generates rules that interfere with this reasoning method.

Some readers might object to the argument developed in this part on the ground that it assumes some natural essence to adjudication. This is incorrect. The argument is not essentialist or conceptualistic. It focuses on institutions as they actually operate in the American legal system. Although complex institutions have multiple functions and those functions overlap to some degree, a major institution like adjudication has a core function (or functions) that distinguishes it from other institutions, such as legislation or arbitration. Moreover, it has certain core features that are critical to its specialized function, and those features ground its normative legitimacy as an institution. One can formulate this same point in terms of integrity rather than legitimacy. In this alternative formulation, the core features of adjudication are those essential to its institutional integrity and thus important to preserve. Whether formulated in terms of legitimacy or integrity, the question is the same: What features of adjudication define its institutional core?

A. *Identifying Adjudication's Core*

One approach to teasing out the core elements of adjudication is to compare it with arbitration. Since parties are free to structure their own procedures in arbitration (within very generous limits),²²⁹ it is worth asking why party choice is more troubling in adjudication. One problem with answering this question is that arbitration takes so many different forms today that it is hard to distill a paradigmatic core. For example, there are areas such as sports and labor arbitration in which many arbitrators feel obliged to follow precedent and give reasons for their decisions.²³⁰ There are other types of arbitration, such as international commercial arbitration, in which arbitrators write opinions but do not feel obliged to follow

229. Contractual freedom to shape procedure lies at the core of arbitration. Brunet, *supra* note 117, at 3–4. There are some limits, however. For example, the parties cannot contractually provide for broader grounds of judicial vacatur and modification of arbitration awards than the Federal Arbitration Act (FAA) allows. See *Hall St. Assocs., L.L.C. v. Mattel, Inc.*, 128 S. Ct. 1396, 1406 (2008) (holding that contracting parties cannot expand the scope of review beyond the strict limits of the FAA).

230. See Christopher J. Bruce, *The Adjudication of Labor Disputes as a Private Good*, 8 INT'L REV. L. & ECON. 3, 9 (1988) (reporting significant evidence of precedent at work in labor arbitrations); Gabrielle Kaufmann-Kohler, *Arbitral Precedent: Dream, Necessity or Excuse?*, 23 ARB. INT'L 357, 365–66 (2007) (reporting an increase in arbitrators' reliance on precedent in the sports field).

precedent.²³¹ And there are still others, such as domestic commercial arbitration, in which arbitrators feel obliged neither to write opinions nor to give reasons unless the parties direct them to do so.²³² Moreover, arbitral procedures in some areas have moved rather close to the litigation model, incorporating such litigation procedures as broad discovery and strict evidence rules.²³³ Finally, since arbitral procedure depends on party choice, one can find different procedural packages for different sets of party preferences.

Nevertheless, it is possible to piece together a baseline for arbitration that differs from adjudication. Arbitration, at its core, focuses on resolving the dispute between the particular parties to the arbitration agreement.²³⁴ This is, after all, why parties have such broad freedom to shape the arbitral process as well as the arbitrator's decision-making protocol. They can choose to have their dispute decided under the laws of a preferred legal system, or they can instruct the arbitrator to decide according to customary norms or even the arbitrator's sense of fairness.²³⁵ Moreover, arbitration is not thought at its core to be about reasoning to a decision from general principles with an obligation to respect precedent. To be sure, arbitrators are expected to do this sometimes, but when they are, it is usually explained by features of the particular context rather than properties of arbitration in general.²³⁶ As far as arbitral procedure is concerned, the move toward a

231. See Kaufmann-Kohler, *supra* note 230, at 362–65 (reporting data showing that arbitrators do not tend to follow precedent in international commercial arbitration). But see Christopher R. Drahozal, *Is Arbitration Lawless?*, 40 LOY. L.A. L. REV. 187, 192–94, 212 (2006) (agreeing with other commentators that arbitrators ordinarily follow the law, despite incentives not to do so, but conceding that no empirical evidence exists “on the extent to which arbitration awards are reasoned or what proportion of reasoned awards are published”).

232. See Alan Scott Rau, *The Culture of American Arbitration and the Lessons of ADR*, 40 TEX. INT'L L.J. 449, 512 (2005) (noting that there is no obligation to write a reasoned opinion and that the AAA discourages doing so).

233. See Thomas J. Stipanowich, *Arbitration: The “New Litigation,”* 2010 U. ILL. L. REV. 1, 11–15, 35–46 (describing this trend). Some but not all of these developments have taken place in the fields of consumer and employment arbitration, partly in response to concerns about the unfairness of adhesion contracts and bargaining inequality. Cf. *id.* at 37, 40 (noting that some courts and state legislatures have responded to the expansion of consumer arbitration by allowing process defenses to arbitration agreements or by imposing new procedural requirements on arbitrations).

234. This point should not be exaggerated. In some areas, such as investor–state arbitration, arbitration is an important source of norm creation. See Kaufmann-Kohler, *supra* note 230, at 376 (arguing that investment law has “a strong need for consistent rule creation”). Also, labor arbitrators, when they engage in interest arbitration, decide matters not strictly covered by the terms of a collective bargaining agreement and in so doing help to shape the framework for ongoing workplace relationships. See RAU ET AL., *supra* note 155, at 18–22 (describing interest arbitration in collective bargaining agreements).

235. 1 THOMAS H. OEHMKE, [2011] OEHMKE COMMERCIAL ARBITRATION § 11.1 (West 3d ed.) (noting that parties can contractually choose to have their dispute decided by the laws of a preferred legal system); Rau, *supra* note 232, at 514–18 (discussing the arbitrator's power to depart from formal legal rules and do equity in the individual case by relying on such things as “commercial understanding, good business practice and notions of honorable behavior”).

236. See Kaufmann-Kohler, *supra* note 230, at 375–78 (offering a context-specific explanation for differing attitudes toward precedent in different arbitration settings); see also Rau, *supra* note

litigation model is a relatively recent phenomenon, in large measure responsive to the Supreme Court's decisions expanding the scope of arbitrable subject matter.²³⁷ And some commentators find this development a bit troubling insofar as it imposes mandatory procedures at odds with arbitration's core commitment to flexibility and party autonomy.²³⁸

Adjudication is different. Judges do not simply decide disputes between the parties. They enforce the substantive law. Moreover, they are expected, if not obliged, to give reasons for their decisions and those reasons are supposed to attend to principle and usually to precedent. In addition, judges are expected to be independent of the parties and (at least in federal court) insulated from political and third-party influence.²³⁹ Independence is meant, among other things, to ensure the sort of detachment essential to principled reasoning. The same is not true for arbitrators. To illustrate, consider the tripartite arbitration panel. It is quite common in these cases for each party to choose one arbitrator and the two arbitrators to choose the third.²⁴⁰ This is not just a convenient method of choosing a panel; rather, it is in many cases a device for constituting a particular decision-making process. In these cases, the party-appointed arbitrators are supposed to be sympathetic to the interests of the party who appointed them.²⁴¹ The result is a panel that engages in a mix of deliberation and straightforward interest bargaining, probably with greater emphasis on the latter.²⁴² Of course, many political scientists describe multimember courts in much the same way,²⁴³ but the characterization elicits concern and controversy when it is applied to adjudication

232, at 510 (“We do not in any event expect that an arbitrator will decide a case the way a judge does.”).

237. See Stipanowich, *supra* note 233, at 9–11 (discussing the Supreme Court's shift in attitude toward arbitration, which led to increased responsibility for arbitrators and necessitated an expansion of arbitration procedure and practice).

238. See *id.* at 50–52 (arguing that unless parties are free to choose the type of arbitration that suits their needs, they will be stuck with “arbitration [with] one-size-fits-all procedures” that mimics litigation and will ultimately be frustrating and unsatisfactory).

239. However, precisely how independent and insulated is a matter of some controversy and uncertainty. See generally John Leubsdorf, *Theories of Judging and Judge Disqualification*, 62 N.Y.U. L. REV. 237 (1987) (discussing the law of judicial disqualification and impartiality).

240. Alan Scott Rau, *Integrity in Private Judging*, 38 S. TEX. L. REV. 485, 497–98 (1997).

241. See *id.* at 498 (discussing the usual assumption that “the ‘tripartite’ form was chosen precisely so that each party can have a ‘friendly’ representative on the arbitration panel who can make sure its ‘side’ is taken seriously”).

242. See *id.* at 501 (explaining how the need to obtain a majority of party-appointed arbitrators “often leads to a process of negotiation and compromise”). Until 2003, the presumption in United States commercial arbitration was that party-appointed arbitrators should be sympathetic to the position of the parties appointing them, but the AAA changed this presumption “[a]t the urging of the international arbitration bar” so that now party-appointed arbitrators must be impartial and neutral in AAA commercial arbitrations unless the parties provide otherwise. RAU ET AL., *supra* note 155, at 263–65.

243. See, e.g., LEE EPSTEIN & JACK KNIGHT, *THE CHOICES JUSTICES MAKE* 56–79 (1998) (defending a strategic model of Supreme Court decision making and arguing that it manifests itself partly in the form of strategic bargaining among Justices).

as compared to relative complacency (if not enthusiastic endorsement) when it is applied to arbitration.²⁴⁴

This suggests that adjudication's core distinctiveness lies in its commitment to reasoning from general principle and doing so in a way that engages the facts of particular cases. Although respecting precedent does not follow inevitably from this commitment, it is closely linked to it either pragmatically (e.g., following precedent limits cognitive error, saves decision costs, or protects reliance interests) or morally (e.g., following precedent is required by equal concern and respect or a norm of integrity, which also supports the core commitment to principled reasoning).²⁴⁵ I realize that this formulation is highly imprecise and that people can hold different views about what principled reasoning entails in adjudication. Even so, the formulation as it stands points us in a productive direction for thinking about party rulemaking. If parties choose procedural rules that undermine the capacity of judges, and perhaps even juries, to engage in principled reasoning of the right sort, then perhaps their choices should not be honored. This is just a beginning, however, for we must explain how procedure is connected to principled reasoning and why parties to a particular case should be constrained if they bear the risks and costs of their own choices.

Before proceeding, it is useful to approach the question of limits from a different direction. Many commentators who reach the question of legitimacy rely on a distinction between two models of adjudication: the dispute resolution model, and the norm-creation or public law model.²⁴⁶ This is a

244. Compare Aaron H. Caplan, *Malthus and the Court of Appeals: Another Former Clerk Looks at the Proposed Ninth Circuit Split*, 73 WASH. L. REV. 957, 981 (1998) (describing judges reasoning with each other as desirable and judges bargaining with each other as undesirable), with Rau, *supra* note 240, at 501–02 (supporting negotiation among party-appointed arbitrators).

245. See, e.g., LARRY ALEXANDER & EMILY SHERWIN, *THE RULE OF RULES: MORALITY, RULES, AND THE DILEMMAS OF LAW* 136–56 (2001) (describing and evaluating three models of precedent with attention to the reasons each can give for following precedent); RONALD DWORKIN, *LAW'S EMPIRE* 225–75 (1986) [hereinafter DWORKIN, *LAW'S EMPIRE*] (explaining how the norm of integrity justifies the practice of following precedent); RONALD DWORKIN, *TAKING RIGHTS SERIOUSLY* 110–23, 340–41 (1978) [hereinafter DWORKIN, *TAKING RIGHTS SERIOUSLY*] (explaining a theory of judicial decision making based on principle and combining fit and justification, all required by a background principle of equal concern and respect); Jonathan R. Macey, *The Internal and External Costs and Benefits of Stare Decisis*, 65 CHI.-KENT L. REV. 93, 95 (1989) (arguing that “stare decisis is an enormously efficient mechanism for conveying information” and enables judges to leverage “the ability to tell when like cases are alike . . . into a facility for deciding a wide variety of cases that involve substantive legal issues about which the judges may know next to nothing”).

246. See, e.g., Moffitt, *supra* note 21, at 505–07 (separating the private-dispute-resolution function from the public-goods-creation function in analyzing the limits of party rulemaking); Thornburg, *supra* note 22, at 206–08 (distinguishing between dispute resolution and public functions of courts). This dichotomy also plays a role in alternative-dispute-resolution scholarship, including evaluations of arbitration. See, e.g., Harry T. Edwards, *Commentary, Alternative Dispute Resolution: Panacea or Anathema?*, 99 HARV. L. REV. 668, 671–72 (1986) (distinguishing between private dispute resolution and public law disputes in the context of alternative dispute resolution); Leo Kanowitz, *Alternative Dispute Resolution and the Public Interest: The Arbitration Experience*, 38 HASTINGS L.J. 239, 241–42 (1987) (noting the same distinction).

very common way to look at adjudication.²⁴⁷ Courts perform two separate functions on this view: they resolve disputes and they make the law. At first glance, this dichotomy might seem tailor-made for evaluating party rulemaking: on this view, party rulemaking should be given broad scope in those cases that are mostly about dispute resolution, such as private contract and tort suits, and very narrow scope in those cases that are mostly about public law or norm creation, such as civil rights suits.

This dichotomy is attractive because it offers a relatively simple way to handle the legitimacy issue, but it fails on two counts. First, it fails to adequately explain why procedures or procedural rules chosen by the parties necessarily threaten the norm- or law-creation function of the court, at least any more so than party choices within the current procedural system. Second and more importantly, the argument grossly oversimplifies adjudication—and in a way that misleads more than helps.²⁴⁸ The fact is that courts do not, and are not supposed to, merely resolve disputes. That is one of the reasons why adjudication and arbitration are different; they perform different functions. Courts resolve disputes *according to the substantive law*. The italicized phrase makes all the difference. It means that we have to pay attention to the conditions for optimal law enforcement.

The point seems so obvious when stated, but the dichotomy between the dispute resolution and public law models ignores its implications. As every first-year law student learns, the substantive law must be interpreted. Interpretation is necessary when a rule is vague or ambiguous, and it is also necessary whenever a rule is applied to the facts of a specific case. There are relatively clear rules, to be sure, but a judge must determine whether the rule's directive is clear when read against the facts. Moreover, when the judge applies a common law rule, the judge *always* must decide whether the case at hand is one warranting further development of the rule. Indeed, the common law does not present itself as a set of canonical rules, although here I realize that I am making a controversial assertion. There are rules, to be sure, but they are linked to and conditioned on general principles and the facts of previous cases in a way that gives them a more-or-less flexible

247. See, e.g., Abram Chayes, *The Role of the Judge in Public Law Litigation*, 89 HARV. L. REV. 1281, 1283–84 (1976) (contrasting the “traditional model” with the “public law model”); Meir Dan-Cohen, *Bureaucratic Organizations and the Theory of Adjudication*, 85 COLUM. L. REV. 1, 1–7 (1985) (highlighting differences between the “arbitration model” and the “regulation model”); Owen M. Fiss, *The Social and Political Foundations of Adjudication*, 6 LAW & HUM. BEHAV. 121, 122–25 (1982) (parsing the “dispute resolution” model and the “structural reform” model); Kenneth E. Scott, *Two Models of the Civil Process*, 27 STAN. L. REV. 937, 937–39 (1975) (distinguishing the “Conflict Resolution Model” and the “Behavior Modification Model”).

248. For a lengthier discussion of this criticism, see Robert G. Bone, *Lon Fuller's Theory of Adjudication and the False Dichotomy Between Dispute Resolution and Public Law Models of Litigation*, 75 B.U. L. REV. 1273, 1279–82 (1995) [hereinafter Bone, *Lon Fuller's Theory of Adjudication*].

character.²⁴⁹ Thus, the judge in every case must engage in an interpretive exercise to determine the legal norm that should decide the case. This is not to say that judges never apply common law rules as such; it means that the judge must decide whether applying a relatively clearly established rule in a strict way is appropriate given rule-of-law values such as predictability and consistency, the magnitude of the case, and the costs of a deeper analysis.²⁵⁰

The reasoning process I have just described is not limited to some special category of public law cases. It, or something roughly like it, is how all judges should go about deciding any case according to the substantive law. They interpret the law, including precedent, in light of the facts of the dispute in order to enforce the parties' substantive rights and decide the case accordingly. Thus, resolving disputes and making law or norms are not two distinct functions; instead, they are two integrated aspects of a single function. Judges neither resolve disputes just for the sake of resolving them nor make decisions just for the sake of guiding future conduct. Adjudication is about both together—and this is so for each and every case. This means that any theory of adjudication must account for the blended character of the decision-making method, and no dichotomy between the dispute resolution model and the public law model can possibly do that.²⁵¹

This approach to the problem of justifying limits on party rulemaking takes us to a similar place as does the arbitration comparison. The core element of adjudication is its distinctive mode of principled reasoning. Judges must interpret the law as they enforce it, and they interpret the law by placing existing legal norms alongside the facts of the particular case. One way to describe this process is in terms of reflective equilibrium. The judge moves back and forth between her best understanding of the law and whatever

249. See DWORKIN, *TAKING RIGHTS SERIOUSLY*, *supra* note 245, at 71–80 (discussing the importance of principles in common law reasoning).

250. Scholars disagree about when judges should ask the rule-of-law questions instead of just apply the rule strictly. Compare, e.g., DWORKIN, *LAW'S EMPIRE*, *supra* note 245, at 264–66 (noting that judges sometimes should apply rules strictly but always subject to the demands of principled integrity), with FREDERICK SCHAUER, *PLAYING BY THE RULES: A PHILOSOPHICAL EXAMINATION OF RULE-BASED DECISION-MAKING IN LAW AND IN LIFE* 196–206 (1991) (arguing for a strong presumption in favor of strict application).

251. Some commentators posit that these two sides of adjudication are “inexorably linked in every institutionalized mode of adjudication” and always in tension with one another. Dan-Cohen, *supra* note 247, at 5. While this view comes closer to the mark in recognizing the intimate connection between dispute resolution and norm creation, it falls short in conceiving of the two as distinct and in tension. I am not convinced that it makes sense to frame adjudication in terms of polar “models.” It is too easy and gives up too soon. Before settling for an account in terms of polar models, one should first make an effort to construct a more general theory aimed at a coherent account of the institution. To be sure, any such theory is likely to identify conflicting values, but it should also provide a more general framework to understand the nature of the conflict and assist in accommodating it.

moral or practical intuitions are generated by the facts, adjusting law and intuition until they fit together in a reflective equilibrium.²⁵²

This is just the first step of the analysis, however. The next step is to explain how and when party choice of procedure interferes with this core element.

B. Party Choice and Principled Reason

My claim is straightforward: because the reasoning process is central to adjudication, we should focus on those procedural rules that have a strong effect on how that process is conducted. This claim is likely to elicit two sorts of objections at the outset. One objection is that parties bear all the risks of the procedures they adopt so there is no reason for special concern about procedures that frame or guide the reasoning process in their case. This is a serious objection and I shall take it up in the following subpart.

The second objection is that there is no way to identify particular procedures that strongly affect the reasoning process because all procedures have the capacity to do so, and one cannot determine in general which procedures are likely to have stronger effects than others. It is correct that many procedures affect the judge's process of deliberation. For example, a party-made rule that restricts discovery to save costs could end up producing only limited evidence that makes the case seem much easier than it really is, and the judge as a result might invest less time deliberating about the law and the facts. But the limited discovery in this example does not necessarily alter the basic process of principled reasoning through reflective equilibrium; it affects the intensity with which the judge engages in it. Intense deliberation is not an essential feature of adjudication. For example, judges today vary in how intensely they deliberate over hard issues. Those who shirk can be criticized for adjudicating poorly, but within outside limits, they cannot be accused of not adjudicating at all.

Compare this example to a situation where the parties agree that their case should be decided by a panel of experts along with the judge and that all these decision makers should give great weight to their own sense of fairness as well as the substantive law.²⁵³ I am much more confident that this agreement alters the reasoning process. In any event, the following discussion assumes that those procedures that frame or guide the decision-making process are likely to have a particularly strong effect on how that process is conducted.

Which procedures and procedural rules fall into this category? This is a difficult question to answer. It might be helpful to consider the views of Lon

252. See DWORKIN, *LAW'S EMPIRE*, *supra* note 245, at 225–27 (describing a process of legal reasoning grounded in an effort to form “the best constructive interpretation of the community’s legal practice”). See *generally* RAWLS, *supra* note 126, at 17–19 (describing the process of seeking a reflective equilibrium).

253. This is perfectly possible in arbitration. See *supra* notes 235–42 and accompanying text.

Fuller, the well-known legal-process theorist who thought a great deal about adjudication in just these terms. Fuller believed that all the main features of adversarial process and party participation were core elements of adjudication and essential to the type of reasoning method that judges should employ.²⁵⁴ Fuller described that method as a form of moral reasoning roughly similar to the reflective-equilibrium idea summarized above.²⁵⁵ He believed that the institution of adjudication had evolved a set of processes that made this method of reasoning work well and that the reasoning method itself was essential to adjudication's ability to assure reasonably good rules and principles over the long run.²⁵⁶

Thus, Fuller isolated a set of procedural elements that he believed were essential to the core function of adjudication and thus to its legitimacy. Fuller did not discuss party rulemaking in particular, but we can guess at his views. He probably would not have been concerned about parties limiting discovery, shortening the statute of limitations, or allowing fee shifting as long as the changes were reasonable, and he probably would not have worried too much about agreements to alter pleading standards or bar joinder. But given his focus on party participation through reasoned argument, he probably would have frowned on agreements that instruct the judge to decide the case solely on the basis of written submissions without any oral argument and perhaps also on agreements that give the judge broad power to limit and control adversarial confrontation during oral argument or trial. For Fuller, the interplay between judge and counsel at oral argument was important to the judge's ability to engage with the case both sympathetically and objectively.²⁵⁷ And Fuller would certainly be concerned about the

254. See Lon L. Fuller, *The Forms and Limits of Adjudication*, 92 HARV. L. REV. 353, 364, 366 (1978) (noting that “[a]djudication is . . . a device which gives formal and institutional expression to the influence of reasoned argument in human affairs” and that “the distinguishing characteristic of adjudication lies in the fact that it confers on the affected party a peculiar form of participation . . . that of presenting proofs and reasoned arguments for a decision in his favor”); see also Bone, *Lon Fuller's Theory of Adjudication*, *supra* note 248, at 1308 (“According to Fuller, adjudication simply could not function properly without party participation by proof and reasoned argument in an adversarial format.”).

255. See Bone, *Lon Fuller's Theory of Adjudication*, *supra* note 248, at 1303–06 (describing how Fuller modeled legal reasoning as a type of moral reasoning employing an approach akin to reflective equilibrium).

256. See generally *id.* (assessing Fuller's views on institutions, the law, and adjudication). Fuller was a pragmatist with a view of society as a field of complicated interaction. He believed that individual decisions produced complex indirect effects and triggered interactions that altered social practice and legal rules dynamically. *Id.* at 1286–87. Thus, the only way to assure good outcomes in the long run was to assure good process. *Id.* at 1287.

257. *Id.* at 1306. Fuller believed that individual participation through lawyers in an adversarial setting allowed the judge to be objectively detached and sympathetically engaged with the case at the same time. *Id.* at 1307. He also believed that detachment and engagement were essential to the form of moral reasoning that adjudication required. *Id.* at 1309. The idea was that party presentation drew the judge into a sympathetic engagement with the facts (mediated by lawyers who could frame the facts in helpful legal terms), while adversarial interaction kept the judge at a distance and prevented too close an alignment with one side. *Id.* at 1308–09. The process of

hypothetical agreement above that placed a panel of experts alongside the judge and dictated a particular decision method.²⁵⁸

There is much in Fuller with which I disagree. I am not a philosophical pragmatist like Fuller, and I do not believe that social institutions are structured around natural-ordering principles.²⁵⁹ Nor do I believe that individual participation and adversarial process are as tightly bound up with adjudicative legitimacy as Fuller insisted they were. For Fuller, the natural-ordering principle of adjudication was the guarantee of individual participation through evidence and reasoned argument in an adversarial setting, and it was this principle that supported an individual right to participate and demanded adversarial procedures.²⁶⁰ But I do agree with Fuller that it makes sense to focus on the reasoning process and to be concerned about procedural choices that might interfere with that process.

It is no doubt impossible to catalog all the procedural rules that belong in this special category. Pleading rules and related rules about motions to dismiss do not belong, nor do most discovery rules, joinder rules, summary judgment rules, evidence rules, and so on. However, the category does include rules defining the decision-making body and the decision protocol, including not only those directed to the judge but also to the jury.²⁶¹ To be sure, we treat the jury process more or less as a black box, but that does not mean that we are indifferent to how the jury conducts its deliberations. We expect a reasoned decision and one that takes account of relevant law and principle.²⁶² It is more difficult to classify rules that aim to shape reasoning incentives indirectly, but some of these might have a strong enough connection to the quality of deliberation to be included, such as rules dealing with appeal rights.

Again, my purpose is not to provide an exhaustive list but rather to illustrate what I have in mind and to highlight the kind of inquiry that classification entails. It might be useful, for example, to distinguish between rules that aim to regulate the conduct of parties and rules that aim to regulate the conduct or decision-making process of the judge. Although this is hardly a clean division, and the precise status of any rule will depend on its particu-

engaging the facts of the case deeply while reflecting from a distance about the implications at the level of general principle—and moving back and forth between fact and principle—was what Fuller believed judges should do. *Id.* at 1309–10.

258. For example, Fuller had problems with tripartite arbitration panels, at least when arbitrators acted in an adjudicative-type capacity. *See* Fuller, *supra* note 254, at 397 (“Where there is from the beginning no real hope of a unanimous decision, this arrangement [tripartite arbitration panels] comes close to being little more than a contractual legitimation of the practice of holding posthearing conferences . . .”).

259. For Fuller’s views on natural-ordering principles, see Bone, *Lon Fuller’s Theory of Adjudication*, *supra* note 248, at 1288–92.

260. *Id.* at 1303.

261. This special category also includes rules regarding judicial impartiality.

262. For example, most people would condemn a jury that simply flipped a coin. *See supra* notes 206–07 and accompanying text.

lar content and effect, it seems reasonable to suppose, as a rough first cut, that the latter set of rules is more likely than the former to be problematic on legitimacy grounds. In the end, not everyone will agree about which party-made rules impair legitimacy, but at least the debate will focus on the right questions.

C. *Justifying the Limits*

This leaves the last and most difficult question: Why limit the parties' freedom to tinker with rules that frame the reasoning process if the parties are free to tinker with everything else? One possible answer is that when parties change these core rules, they make the judge oversee a process that is not "adjudication." But this answer merely begs the question. Even if the resulting process is not properly called "adjudication," the question still remains: Why is it undesirable for judges to engage occasionally in a process that is not adjudication, especially when the parties prefer it? The answer cannot be simply a matter of proper labeling.

We could justify imposing restraints if parties routinely made procedural changes that altered the decision-making process in ways that produced bad law or bad precedents. The concern here is not an undersupply of good precedent but rather an oversupply of bad precedent. The problem, however, lies in the premise that parties would routinely make radical procedural alterations if allowed to do so. This is ultimately an empirical question, but I am highly skeptical.

Still, it might be possible to justify restraints even if radical departures happen only occasionally and not routinely. The concern on this view is that asymmetric information will generate a "lemons" problem.²⁶³ The argument starts from the premise that it will be very difficult for lawyers and judges in later cases to distinguish between those precedents that are tainted by parties' tinkering with core elements and those that are not.²⁶⁴ If lawyers and judges cannot sort the good precedents from the bad, they are likely to discount the value of all precedents. However, one should not put too much weight on this argument. The quality of legal rulings and precedents created by the current system varies a great deal already, yet the system of precedent still seems to function. Given this, it is hard to tell whether party rulemaking will have a significant negative impact.

263. See generally George A. Akerlof, *The Market For "Lemons": Quality Uncertainty and the Market Mechanism*, 84 Q.J. ECON. 488 (1970) (describing the lemons model, which is based on informational asymmetry). For example, suppose the market for used cars includes sellers who market high-quality cars and sellers who market low-quality cars. If consumers are unable to verify quality before purchase, they will pay only the average of the value over high- and low-quality cars, which will not be enough for the sellers of high-quality cars. As a result, only low-quality cars will remain. *Id.* at 489-90.

264. This assumes that judges, when they create precedents, do not routinely describe the procedures the parties have chosen.

We might worry about outcomes in a slightly different way. The core elements of adjudicative process have developed over centuries. A Burkean conservatism might counsel against letting parties tinker with these elements because of deep uncertainty about the consequences.²⁶⁵ This argument has some merit, I believe, but it also has the potential to sweep too broadly. Nothing about it focuses necessarily on judicial reasoning. Any longstanding feature of procedure can fall within its scope.

This leaves two other ways to justify limits on party rulemaking. One focuses on the interests of the parties themselves. The other focuses on long-term effects not already discussed.

First consider party interests. One might argue that even sophisticated parties are in a poor position to make predictions about how procedural rules that alter core elements of adjudication will affect their cases. Adjudication is such a complex institution that predicting outcome effects is bound to be extremely difficult, if not impossible. If parties are likely to overestimate their ability to evaluate these risks due to bounded-rationality constraints, then preventing them from tinkering with core elements of adjudication might be justified as serving their best interests.

The problem with this argument is that it proves too much. It is very hard to predict the effects of any procedural change in the highly strategic environment of litigation. Unless one is prepared to ban all party rulemaking, which seems unjustified as a way to serve party interests, it is not clear how to draw lines.²⁶⁶ In fact, this same argument could justify banning arbitration or at least requiring arbitration to be modeled on the core elements of adjudication.

Finally, consider long-term effects. One such effect that concerns some scholars analyzing party rulemaking has to do with harm to the symbolic value of adjudication.²⁶⁷ In this connection, one might argue that adjudication has social value as a symbol of our collective commitment to principled reason in government decisions and actions, and that this commitment is reflected in certain invariant elements of procedure closely tied to principled decision making.²⁶⁸ Giving parties wide latitude to shape these elements

265. See generally Ernest Young, *Rediscovering Conservatism: Burkean Political Theory and Constitutional Interpretation*, 72 N.C. L. REV. 619, 648–50 (1994) (describing Edmund Burke's philosophy, which counseled reliance on wisdom inherent in existing institutions because "[e]ven the most intelligent individual is limited not only by his own understanding but also by the relative deficiency of his own experience").

266. Moreover, it is excessively paternalistic to assume that the parties cannot figure out that certain procedures might serve their interests better than others. For example, there are some fairly predictable advantages associated with having a panel of experts decide a case rather than a nonexpert judge. To be sure, there are also likely to be unanticipated consequences, but it is not clear why sophisticated parties should not be allowed to take the risk and strike the balance for themselves.

267. See *supra* note 210 and accompanying text.

268. There are two different ways in which adjudication might be a symbol. Adjudication might be thought to be an intrinsic symbol of principled reason in government, entirely apart from

threatens the value of the symbol in two ways. First, it sends the message that adjudication is simply a device for serving party interests. Second, it places the invariant elements up for grabs, which introduces variability and distorts the symbol's message.

I agree that adjudication has symbolic value, but I do not believe that symbolic value can carry the full weight of justifying a party-rulemaking ban. The conclusion that party rulemaking endangers symbolic value depends on empirical assumptions about public perceptions. I find it rather far-fetched that people would lose the meaning of the symbol if judges sometimes followed procedures that parties dictate. The default system of procedure would be as it is today, and there would still be decisions and judgments. Perhaps there would be fewer trials, which might reduce the most dramatic way that the symbol's meaning is communicated. But that is far from certain, and anyway, there are very few trials today.²⁶⁹

There is, however, a more serious long-term effect that supports a stronger argument for limiting party rulemaking. This effect has to do with the likely response of judges to a system in which parties can direct what judges do and how they do it. If a judge believes that all judges should think hard about law and fact in an effort to reach a principled decision, the judge is likely to invest the effort necessary to do so, and to do so very well. Procedures and decision rules designed to frame, guide, and encourage the operation of principled reason are constant reminders to the judge of its importance. Also, these procedures and rules provide a foundation for the development of a shared norm among judges in favor of principled decision making. If each judge knows that other judges are subject to the same norms and thus similarly incentivized, all judges are more likely to invest in the process. Doing it well will confer prestige and other reputation benefits. Furthermore, with a collective practice of principled reason, the norm is more likely to become internalized and serve as a reason in itself for action.

This beneficial equilibrium is put at risk by allowing parties to alter the procedures and rules that frame and incentivize the proper reasoning process. In such a world, judges would likely revise their perception of the ideal judicial role, perhaps modeling it more on the role of an arbitrator. Also, the altered perceptions would destroy or seriously undermine the benefits of a reputation market and impede the internalization of the proper norm. Moreover, these consequences could happen even if parties only occasionally employed radically altered procedures. To be sure, there would have to be enough cases to make the threat of party rulemaking sufficiently credible to undermine the equilibrium, but the critical number need not be very large.

how people understand its meaning. This makes little sense, however. If a symbol's value is the message it symbolizes, then it must matter how people perceive its meaning. The other way that adjudication might be a symbol is that people perceive it as conveying a symbolic message about the importance of principled reason. This is the way I treat it in the discussion in the text.

269. See *supra* note 199.

This is so because the consequences flow from inferences about judicial role based on the fact that parties are allowed to make their own rules and not on how many parties actually do it.²⁷⁰

I believe that this might well be the most compelling defense of party-rulemaking limits on legitimacy grounds. Of course, it depends on empirical assumptions about judicial psychology. But if those assumptions are correct, then it might be wise to restrain parties from making radical changes to procedure that impair the uniformity of core elements. Moreover, it might be sensible to allow judges to review permissible party agreements for reasonableness subject to a presumption in favor of enforcement, just as they do for most agreements today. This would instill a feeling in the judge that she has control and is not just a tool for the parties to manipulate. Moreover, doing so enables judges to prevent changes that they believe risk the core elements of adjudicative reasoning in individual cases.

To recap, the key to the legitimacy critique of party rulemaking is to identify a core element of adjudication. I argued that the core is a commitment to a mode of reasoning that engages general principle and case-specific facts in an effort to reach reflective equilibrium. It follows that party rulemaking is most problematic when it alters procedures and rules designed to frame, guide, or incentivize this reasoning process. If doing so is likely to increase bad law or bad precedents significantly, then there is reason to ban party rulemaking when it affects these procedures and rules. But if party rulemaking is not likely to produce bad law or bad precedent on a sufficiently large scale, then the most compelling reason to ban it is to prevent long-term adverse consequences to judicial role-perceptions and, as a result, to the quality of judicial decisions.

V. Conclusion

Our journey has been a lengthy one, but this is because the problem we set out to solve is a complicated one, more complicated than scholars have realized. Whether to allow party rulemaking, especially Type III rulemaking, requires the answer to a deep question about the source of adjudication's normative legitimacy. Before we reached that question, however, we explored other arguments that have been made against party rulemaking and concluded that they justify imposing restraints in three limited situations: when parties expressly agree to exclude a third party with legal rights that might be seriously affected, when the agreement is sharply one-sided and the

270. The same effect can be created in a slightly different way. When judges expect that parties will sometimes choose procedures that undermine the reasoning process, they are not likely to invest as much in developing the requisite reasoning skills (assuming that it is costly for them to do so), and as a result, the quality of all judicial decision making suffers. In other words, party rulemaking that tinkers with the reasoning process generates externalities that can adversely affect the reasoning process in all other cases. The way to avoid this result is to make the core procedures closely linked to optimal reasoning mandatory and ban party rulemaking that tinkers with them. I thank my colleague Abe Wickelgren for suggesting this alternative formulation.

market has little competition, and when the agreement adopts procedures that seriously disable private enforcement of the substantive law or impair the proper consideration of civil rights claims. And in the end, our analysis led us to conclude that if parties are to be barred in other situations, it must be because those alterations impair the normative legitimacy of the institution.

We then explored the core of adjudication's legitimacy and found it in a distinctive mode of reasoning. Even a thin account of this reasoning process was enough to predicate important insights and recommendations. In particular, it provided a basis for banning party rulemaking that tinkers with procedures and rules closely tied to the judge's or the jury's reasoning process, and it also supported giving judges power to review rulemaking agreements for reasonableness.

Party rulemaking is here to stay and it has the potential to transform the way we think about procedural law. It also creates risks to the institution of adjudication itself. We must think harder about these consequences and their policy implications. The future of adjudication may depend upon it.

Injunctions as More (or Less) than “Off Switches”: Patent-Infringement Injunctions’ Scope

John M. Golden*

Injunctions have often been viewed as mere “off switches” that prevent future violations of rights protected by so-called property rules. But injunctions in fact come in a variety of forms having different objects, scopes, and degrees of effectiveness. In practical situations, an injunction might amount to little more than a threat of higher-than-normal monetary sanctions delivered at substantially higher-than-normal speed.

This Article builds on these insights by investigating the potential and actual scopes of injunctions against patent infringement. An economic model for infringer incentives shows how concerns of injunction scope are substantially analogous to widely examined concerns of patent scope. A new taxonomy provides named classifications for different forms of injunctions. A systematic study of patent-infringement injunctions issued by U.S. district courts in 2010 indicates how often these different forms appear in practice. Startlingly, this study suggests that the majority of such patent-infringement injunctions take an “obey the law” form that violates the Federal Rules of Civil Procedure, at least as the U.S. Court of Appeals for the Federal Circuit has traditionally understood those rules. In another indication of patent law’s technology specificity, only 12% of the injunctions directed to biomedical-substance technology feature such apparent error. Meanwhile, courts frequently issue specially tailored injunctions that protect patent rights more or less than a conventional “do-not-infringe” order would. Prophylactic injunctions and other specially tailored injunctions should be recognized as legitimate forms of relief that can enable better balancing of concerns of notice, rights protection, rights limitation, and administrability.

* Professor, The University of Texas School of Law. For helpful comments, I thank David Adelman, Ronen Avraham, Robert Bartlett, Bernie Black, Oren Bracha, Aaron Edlin, Jeannie Fromer, Mira Ganor, Mark Gergen, Keith Hylton, Louis Kaplow, Stefanie Lindquist, Lee Petherbridge, Doug Rendleman, Ben Roin, Charles Silver, Henry Smith, Holger Spamann, Matt Spitzer, Eric Talley, Tracy Thomas, Marketa Trimble, Abe Wickelgren, and the editors of the *Texas Law Review*. I also thank participants in or reviewers for the following: the 2011 Conference on Empirical Legal Studies; Harvard Law School’s Law and Economics Seminar; the Fordham Law School’s Intellectual Property Colloquium; the Fordham Intellectual Property Law Institute’s 19th Annual Intellectual Property Law and Policy Conference; the 2011 Intellectual Property Scholars Conference; the 2011 Samsung–Stanford Conference on Patent Remedies; the University of California, Berkeley, School of Law’s Law and Economics Workshop; and The University of Texas School of Law’s Law, Business, and Economics Workshop.

I.	Introduction.....	1400
II.	Comparability of Concerns with Patent Scope and Injunction Scope.....	1405
	A. Design-Arounds and Persistently Ambiguous Patent Rights.....	1406
	B. Injunctions as Gateways to Potential Contempt.....	1409
	C. Modeling Infringer Incentives	1415
III.	U.S. Law and a Taxonomy of Injunction Types	1420
	A. Type-0 Through Type-2 Do-Not-Infringe Injunctions	1420
	B. Purely Reparative Injunctions.....	1424
	C. Specially Tailored Injunctions	1426
	1. <i>Characteristics of Specially Tailored Injunctions</i>	1426
	2. <i>Legal Status of Specially Tailored Injunctions</i>	1428
IV.	District Court Practice in Issuing Patent-Infringement Injunctions	1433
	A. Systematic Violation of Federal Rule of Civil Procedure 65(d).....	1435
	1. <i>Consented-To, Uncontested, and Actively Opposed Orders</i>	1435
	2. <i>Preliminary Versus Permanent Injunctions</i>	1439
	3. <i>Subject-Matter Specificity of Error Rates and Earlier Error Rates Revisited</i>	1440
	4. <i>Geographic Ubiquity of Error</i>	1443
	5. <i>Potential Explanations for High Rule 65(d) Error Rates</i>	1444
	B. Purely Reparative Injunctions in U.S. Practice.....	1449
	C. Specially Tailored Injunctions in U.S. Practice	1449
	1. <i>Correlated-Activity Injunctions</i>	1450
	2. <i>Destruction, Disablement, or Delivery Injunctions</i>	1451
	3. <i>Reformulated-Bounds Injunctions</i>	1452
	4. <i>Moderated Injunctions</i>	1455
V.	Crafting the Optimal Injunction.....	1456
VI.	Conclusion	1470

I. Introduction

Much recent debate on patent-infringement remedies has focused on two issues: when injunctive relief should be available¹ and how damages

1. See, e.g., John M. Golden, *Principles for Patent Remedies*, 88 TEXAS L. REV. 505, 506–07 & n.5 (2010) (noting the existence of “conflicting answers from lower courts and academic commentators regarding how to decide when injunctions should issue” (footnotes omitted)); Peter Lee, *The Accession Insight and Patent Infringement Remedies*, 110 MICH. L. REV. 175, 240 (2011) (proposing to “apply[] accession doctrine—with some modification—to deny injunctive relief when

should be calculated.² This Article addresses a comparatively neglected question: what the scope of a patent-infringement injunction should be when it is granted. Neglect of this issue might help explain a startling fact that this Article's empirical study reveals: the majority of patent-infringement injunctions issued by U.S. district courts in 2010 appear to violate the Federal Rules of Civil Procedure as traditionally interpreted by the U.S. Court of Appeals for the Federal Circuit.³

The fact that many injunctions have taken an apparently improper form highlights the underlying capacity of injunctions to take many different forms. Although much commentary treats injunctions as mere "off switches" that enforce property rules,⁴ injunctions can take any of a number of different shapes having differing degrees of effectiveness.⁵ Even if there is no debate over the timing and duration of an injunction,⁶ there can be debate over an injunction's scope—i.e., over the extent and nature of the matter and

an infringer substantially improves on an underlying patented invention"). Denials of patent-infringement injunctions have raised questions of when and how a court should award an "ongoing royalty" to compensate for expected activity that the court does not enjoin. See *Paice LLC v. Toyota Motor Corp.*, 504 F.3d 1293, 1314 (Fed. Cir. 2007) ("Under some circumstances, awarding an ongoing royalty for patent infringement in lieu of an injunction may be appropriate."); H. Tomás Gómez-Arostegui, *Prospective Compensation in Lieu of a Final Injunction in Patent and Copyright Cases*, 78 *FORDHAM L. REV.* 1661, 1664–65 (2010) (seeking "to demonstrate that federal courts have no authority to award compulsory prospective compensation . . . for postjudgment copyright and patent infringements"). Exploration of the proper limits and form of ongoing-royalty orders is outside the scope of this Article.

2. See, e.g., Golden, *supra* note 1, at 582–86 (examining reform proposals relating to the calculation of damages); Mark A. Lemley & Carl Shapiro, *Patent Holdup and Royalty Stacking*, 85 *TEXAS L. REV.* 1991, 2017–35 (2007) (discussing legal standards and practices regarding reasonable-royalty damages).

3. See *infra* text accompanying note 162.

4. See Mark A. Lemley & Philip J. Weiser, *Should Property or Liability Rules Govern Information?*, 85 *TEXAS L. REV.* 783, 783 (2007) ("[L]aw and economics scholars call injunctive relief a 'property rule.'"); see also Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 *HARV. L. REV.* 1089, 1092, 1115–16 (1972) (characterizing enforcement of a legal entitlement by an injunction as a form of "protect[ion] by a property rule"). But see Joshua P. Davis, *Taking Uncertainty Seriously: Revising Injunction Doctrine*, 34 *RUTGERS L.J.* 363, 397 (2003) ("The choice a judge makes regarding preliminary and final injunctive relief need not be dichotomous.").

5. See Doug Rendleman, *The Trial Judge's Equitable Discretion Following eBay v. MercExchange*, 27 *REV. LITIG.* 63, 74 (2007) ("The judge's decisions in drafting an injunction are contextual and discretionary: these are the details of what to forbid or require and the timing of whether or not to give the defendant a period to adjust and, if so, how long."); cf. OWEN M. FISS, *THE CIVIL RIGHTS INJUNCTION* 7 (1978) (describing the law as having "long embraced a pluralism with regard to injunctions" and proposing a new scheme for classifying injunctions as "preventive," "reparative," or "structural").

6. Cf. Orit Fischman Afori, *Flexible Remedies as a Means to Counteract Failures in Copyright Law*, 29 *CARDOZO ARTS & ENT. L.J.* 1, 29–30 (2011) (suggesting, in the copyright context, that "an injunction could be granted for a limited period of time only" in order to balance copyright-owner and free speech interests); John M. Golden, Commentary, *"Patent Trolls" and Patent Remedies*, 85 *TEXAS L. REV.* 2111, 2148 n.136 (2007) (discussing considerations that might favor or disfavor staying an injunction).

activities that an injunction forbids or requires.⁷ Further, issuance of an injunction does not necessarily halt potentially infringing activity. In practical situations, an injunction might amount to little more than a threat of higher-than-normal monetary sanctions delivered at higher-than-normal speed. As the recent en banc case of *TiVo Inc. v. EchoStar Corp.*⁸ suggests, a rational enjoined party might choose to take deliberate action that risks a later holding of contempt.⁹ Consequently, an injunction can be fundamentally ineffective at enforcing a property rule—at least to the extent such a rule is understood to involve a state “guarantee[] [of] property right assignments against infringement.”¹⁰

The potential ineffectiveness of an injunction in enforcing a property rule calls into question the generality and depth of a commonly invoked dichotomy between legal regimes enforced through injunctions and those enforced through monetary relief alone. As Louis Kaplow and Steven Shavell have previously pointed out, this dichotomy dissolves once monetary relief can come in “any quantum”: in their words, “a liability rule with very high damages is equivalent to property rule protection of victims.”¹¹ This Article emphasizes the further point that the dichotomy can likewise dissolve where, as seems probable in much civil litigation, an injunction fundamentally acts as a mere gateway to a risk of monetary sanctions that an enjoined party can rationally choose to bear.¹²

In any event, the likely limitations and potential leakiness of injunctive relief highlight questions about how a patent-infringement injunction can best be crafted to balance protection of patentee rights with protection of

7. See, e.g., Supplemental Brief of Defendants-Appellants on Rehearing En Banc at 34, *TiVo Inc. v. EchoStar Corp.*, 646 F.3d 869 (Fed. Cir. 2011) (en banc) (No. 2009-1374) (contending that, “at least for a first-time infringer, [an] injunction may not legally go further” than “prevent[ing] the violation of any right secured by patent” (quoting 35 U.S.C. § 283 (2006))). For purposes of this Article, “injunction scope” is not considered to be concerned with the nature and extent of the individuals or entities whose compliance with the injunction is ordered.

8. 646 F.3d 869 (Fed. Cir. 2011) (en banc).

9. *TiVo Inc. v. EchoStar Corp.*, No. 2009-1374, 2010 U.S. App. LEXIS 4543, at *39 (Fed. Cir. Mar. 4, 2010) (“Given EchoStar’s refusal to disable the DVR functionality in its existing devices and the fact that its original attempts to design around TiVo’s patent were wholly unsuccessful, the district court had ample justification for . . . determin[ing] that court pre-approval of any new design-around effort was necessary to prevent future infringing activity.”), *vacated*, 376 F. App’x 21 (Fed. Cir. 2010).

10. Louis Kaplow & Steven Shavell, *Property Rules Versus Liability Rules: An Economic Analysis*, 109 HARV. L. REV. 713, 715 (1996) (contrasting such a state guarantee under a property rule with “liability rules, under which [the state] merely discourages violations by requiring transgressors to pay victims for harms suffered” (emphasis omitted)); see also *id.* at 723 (defining a property rule as involving “absolute protection of [an] entitlement”).

11. *Id.* at 724; cf. *id.* at 756 (discussing the possibility of “[v]iewing property rules and the conventional liability rule as members of a continuum of liability rules”).

12. Cf. *id.* at 757 (“When we consider how property and liability rules are actually applied, we also see that the view that they lie on a continuum is descriptively helpful, because the rules often turn out to be different from both true property rules and the liability rule with damages equal to harm.”).

legitimate public and infringer interests.¹³ This Article consequently studies issues of injunction scope theoretically, empirically, and normatively. Part II takes a theoretical approach. After discussing how injunctions are enforced through contempt proceedings and sanctions, Part II presents an economic model for an infringer's incentives that illustrates how, practically speaking, the frequently drawn distinction between injunctive and monetary relief can be substantially illusory. The model further illustrates how concerns about injunction scope parallel concerns about the scope of underlying rights—at least when the scope of those rights, like patent rights, tends to remain uncertain even after a prior judgment of infringement.¹⁴ Part III follows with a taxonomy of injunction types and a primer on U.S. law regulating patent-infringement injunctions' scope. Part IV uses Part III's taxonomy to present and analyze results from a systematic study of patent-infringement injunctions issued by U.S. district courts in 2010. Part V then provides normative recommendations for how courts should craft injunctions in light of concerns of administrability, notice, patent-right protection, and avoidance of undue chilling of potentially legitimate activity.

Part III's taxonomy and Part IV's empirical results warrant a fuller preview. Part III's taxonomy features five categories that reflect the express language of court orders. These taxa can be described as follows:

- Measure-zero “do-not-infringe” injunctions (also characterized as Type-0 injunctions) explicitly forbid only future infringement that involves the exact products or processes already adjudged to infringe (e.g., “Do not make, use, offer to sell, sell, or import into the United States the SuperHypo widget held to infringe claims 1 through 5 of the '777 patent.”).¹⁵

13. With respect to concerns of potential patent “holdup” or “holdout,” crafting of injunctive language might be particularly effective in addressing concerns about the chilling of design-around activity from overly broad or vague injunctions. On the other hand, as long as injunctive language requires an infringer to abandon its existing course of activity, careful tailoring of injunctive language might do little to address holdup or holdout concerns that result from a lock-in effect under which any significant change in the infringer's course will impose large costs that have essentially nothing to do with the merits of the patented invention. Consequently, consideration of problems of injunction scope might help disentangle holdout concerns that arise from two different causes: (1) an injunction so broadly written that all feasible design-arounds risk a holding of contempt, and (2) lock-in effects that mean that any design-around, no matter how legally permissible, entails high cost.

14. Cf. Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839, 916 (1990) (“[A] ‘strengthening’ of property rights will not always increase incentives to invent When a broad patent is granted or expanded via the doctrine of equivalents, its scope diminishes incentives for others to stay in the invention game”).

15. In mathematics, a measure “is a rule that assigns a non-negative number (or $+\infty$) . . . to each set . . . in [a] collection” of sets. 2 PAUL BAMBERG & SHLOMO STERNBERG, *A COURSE IN MATHEMATICS FOR STUDENTS OF PHYSICS* 801 (1990). The assigned number is generally representative of the size of the set with which it is associated. Cf. *id.* (observing that one axiom of measure theory is that the measure of a “countable union of disjoint sets” equals the sum of the measures of the individual disjoint sets included in the union). Hence, according to one standard

- Colorable-differences do-not-infringe injunctions (also characterized as Type-1 injunctions) add to a Type-0, measure-zero injunction an explicit prohibition of infringement that involves only relatively insignificant variations of the products or processes specified by accompanying Type-0 language (e.g., “Do not make, use, offer to sell, sell, or import into the United States the SuperHypo widget or any widget no more than colorably different from the SuperHypo widget.”).¹⁶
- “Obey-the-law” injunctions (also characterized as Type-2 injunctions) use language that, at least on its face, generally prohibits infringement of a patent or patent claim without tying the scope of the prohibition to products or processes already adjudged to infringe (e.g., “Do not infringe the ’777 patent”; alternatively, “Do not infringe claim 1 of the ’777 patent.”).
- Purely reparative injunctions have as their sole direct purpose the correction of harm caused by past infringement (e.g., “Destroy all SuperHypo widgets manufactured in the United States that are now located abroad, even though there is no expectation that they will return to the United States.”¹⁷).
- Specially tailored injunctions do at least one of the following: (i) prohibit at least some activity that might be noninfringing (e.g., “Do not display SuperHypo widgets on a website.”); (ii) require action, such as destruction of infringing devices, that might not be absolutely required to prevent future infringement (e.g., “Destroy all SuperHypo widgets that are in the United States and under your control.”); (iii) define their scope without reference either to underlying patent rights or to matter already adjudged to infringe (e.g., “Do not make, use, offer to sell, sell, or import into the United States

convention, the measure of an interval from a to b along the line of real numbers (a straight line stretching from $-\infty$ to $+\infty$) equals the magnitude of the difference between a and b . *Id.* Thus, the measure of the interval from 5 to 9 is 4. On the other hand, the measure of a single real number such as 5 equals the magnitude of the difference between 5 and 5—namely, zero. Under this convention, the number 5 is a set of measure zero along the line of real numbers.

16. Because a Type-1, colorable-differences injunction generally includes a foundational Type-0, measure-zero injunction, this Article will commonly refer to an order that has Type-0 and Type-1 aspects simply as a “Type-1 injunction.”

17. Because United States patent law generally does not forbid the use abroad of a product or process whose use in the United States would infringe a U.S. patent, a product or process located abroad would, generally speaking, not be expected to factor into future U.S. patent infringement unless there were reason to expect that the product or something generated by the process would be brought to the United States. *See* 35 U.S.C. § 271 (2006) (describing forms of patent infringement); *see also* Microsoft Corp. v. AT&T Corp., 550 U.S. 437, 441 (2007) (“It is the general rule under United States patent law that no infringement occurs when a patented product is made and sold in another country.”).

any widgets.”); or (iv) include an explicit carve out for infringing (or likely infringing) behavior (e.g., “Do not use, offer to sell, sell, or import any SuperHypo widget, although you are not prohibited by this order from making one or more SuperHypo widgets.”).

These five categories of injunctions—Type-0, Type-1, Type-2, purely reparative, and specially tailored—are nonexclusive. A single court order can feature injunctive language from multiple taxa. Indeed, as defined in this Article, a Type-1, colorable-differences do-not-infringe injunction essentially requires the simultaneous existence of a Type-0, measure-zero injunction. In practice, such a Type-0/Type-1 combination is commonly supplemented by a Type-2, obey-the-law injunction or a specially tailored injunction.

Nonetheless, the five injunction types are not equal under U.S. law. The Federal Circuit has held that the U.S. Patent Act does not authorize courts to grant purely reparative injunctions.¹⁸ Similarly, the Federal Circuit has held that the Federal Rules of Civil Procedure generally forbid Type-2, obey-the-law injunctions.¹⁹

Strikingly, Part IV's empirical analysis reveals vastly different rates of compliance with these bans. In a dataset of 143 patent-infringement injunctions issued in 2010, only one contains purely reparative language. In contrast, nearly 60% contain Type-2, obey-the-law language. Although active contestation of injunctions or their bases by an adversarial party seems to correlate with a reduced percentage of injunctions containing disfavored Type-2 language, the error rate for such actively contested injunctions is still about 44%. Part IV provides potential explanations for the high error rates.

Part IV also discusses empirical results indicating that U.S. district courts issue specially tailored injunctions in nontrivial numbers and in a number of subtypes. With these subtypes in mind, Part V suggests that, although Type-1, colorable-differences injunctions might provide an appropriate general default, specially tailored injunctions might often enable a better balancing and promotion of patentee, infringer, and societal interests.

II. Comparability of Concerns with Patent Scope and Injunction Scope

Problems of patent scope have figured much in the minds of patent commentators and practitioners.²⁰ Problems of injunction scope have been comparatively neglected despite involving substantially overlapping normative concerns. Most centrally, patent and injunction scope both implicate a fundamental interest in balancing the goals of rewarding innovative patentees

18. See *infra* text accompanying notes 109–14.

19. See *infra* text accompanying notes 94–97.

20. See, e.g., John M. Golden, *Construing Patent Claims According to Their “Interpretive Community”*: A Call for an Attorney-Plus-Artisan Perspective, 21 HARV. J.L. & TECH. 321, 322 (2008) (“Determination of the scope of a patented invention is one of the most contentious and difficult tasks of modern patent law.”).

and ensuring that patent rights do not excessively impede the use or further development of innovations.²¹ In the context of an existing injunction, the difficulty of balancing comes into sharp focus as courts and parties try to determine the proper legal status of an alleged “design-around”—a term used here to refer to a functional substitute for an infringing product or process that an adjudged infringer posits to be at least potentially noninfringing or injunction-compliant. Despite the potential inefficiency of design-around activity, U.S. patent law has tended to take the position that “legitimate design-around efforts should always be encouraged.”²² Of course, U.S. patent law also looks to protect the competing interests of patent holders, the result being that courts must seek to distinguish between legitimate design-around activity and illegitimate attempts at circumvention.²³

A. *Design-Arounds and Persistently Ambiguous Patent Rights*

To fully understand the difficult line-drawing problems associated with enforcing patent rights, one needs to appreciate how the malleability of technology can complicate the task of defining the proper scope of such rights. Consider, for example, a selection of the language from claim 1 of U.S. Patent No. 4,963,736 (the ’736 patent):

A mass spectrometer . . . comprising:

(a) first and second vacuum chambers separated by a wall, . . .

(c) a first rod set in said first vacuum chamber . . . and a second rod set in said second vacuum chamber²⁴

For patent-claim language, this selection might be thought remarkably clear. But suppose an adjudged infringer of claim 1 redesigns its mass spectrometer so that it has an additional vacuum chamber that does not contain any rods and that is located between the only two vacuum chambers containing rod sets.²⁵ What chamber should now be considered to be the “second vacuum chamber”? If the second vacuum chamber is the newly added middle chamber, claim 1 is no longer literally infringed because there is no rod set in this chamber. If, instead, one of the outside vacuum chambers is considered to be the second vacuum chamber and the other outside vacuum chamber is considered to be the first vacuum chamber, then there

21. See *Merges & Nelson*, *supra* note 14, at 843 (“Without extensively reducing the pioneer’s incentives, the law should attempt at the margin to favor a competitive environment for improvements . . .”).

22. *TiVo Inc. v. EchoStar Corp.*, 646 F.3d 869, 883 (Fed. Cir. 2011) (en banc).

23. See *id.* (“But an assertion that one has permissibly designed around a patent should not be used to mask continued infringement.”).

24. U.S. Patent No. 4,963,736 col. 14 II. 24–35 (filed Nov. 15, 1989).

25. See *Applera Corp. v. Micromass UK Ltd.*, 186 F. Supp. 2d 487, 501, 504 (D. Del. 2002) (observing that a device accused of infringing the ’736 patent included “an empty vacuum chamber before the hexapole ion bridge chamber alleged by [the plaintiffs] to be the ‘first vacuum chamber’ in the ’736 patent”).

will be an argument that the “first and second vacuum chambers” are no longer “separated by a wall.”²⁶ Instead, they are separated by an entire additional vacuum chamber. A frustrated patentee might respond that the claim language does not say “separated *only* by a wall” and thus that the inclusion of an additional separating wall and some space between the two separating walls does not take the redesigned device outside the scope of the claim.²⁷ The adjudged infringer might counter that the patentee’s response renders the “separated by a wall” language superfluous: even without this language, we would know that the first and second vacuum chambers are separated by at least one wall or wall equivalent; otherwise, they could not be distinct first and second chambers.

Moreover, the “first,” “second,” and “separated by a wall” language is not the only potential basis for dispute. What, for example, is meant by the term *rod*? The term might have been thought so clear in an initial round of litigation that it was not even construed; that appears to have been the case in an actual dispute that involved the ’736 patent.²⁸ Faced with the problem of construing *rod* for the first time in contempt proceedings, a court might think that it was defining the term reasonably and even quite broadly by viewing *rod* as signifying, for example, a continuous solid structure having a length along a substantially straight longitudinal axis that is significantly greater than its maximum width transverse to that axis.²⁹ But this definition’s use of the adjective *continuous* could leave a significant technological loophole: it would appear to exclude from literal infringement a structure that consists of a series of cubes that are arrayed in a straight line with small spaces between them. Should the use of such an array of cubes make impossible a holding of contempt? Regardless of how one answers this question, the main point is the potential existence of significant, latent ambiguity even in patent claims whose meaning has already been litigated.

Such ambiguity is potentially significant and problematic because of competing interests in protecting already-infringed rights against further violation and, at the same time, allowing the continuation of legitimate, competitive activity involving potentially quite similar but nonetheless noninfringing technology—perhaps the canonical form of design-around activity. Such activity is commonly considered to be socially beneficial because it can contribute to a number of apparently desirable ends:

26. See *id.* at 509 (noting that the defendants “propose[d] a construction of ‘separated by a wall’ and ‘interchamber orifice’ that would require the wall and interchamber orifice to join or link the two vacuum chambers and spaces”).

27. See *id.* at 510 (noting the plaintiffs’ argument that “‘separated by a wall’ should be construed to mean only that ‘there is at least a wall between the first and second vacuum chambers’”).

28. See *id.* at 508 (finding “the proper construction of *rod* to be self-evident” and that the term does not need construction “because ‘a rod is a rod’”).

29. Cf. MERRIAM-WEBSTER’S COLLEGIATE DICTIONARY 1079 (11th ed. 2003) (defining *rod* as “a slender bar (as of wood or metal)”).

maintaining a competitive technological marketplace, ensuring exploration of a variety of technological options before large sectors of society become locked into potentially suboptimal standards, and curbing the potential for patentee holdout or other causes of patent-rights overvaluation.³⁰ In the face of a holdout or another cause of exorbitant patentee demands, a competitor or other potential user of technology can investigate the possibility of a design-around—a technological replacement or alternative to what is covered by a rights holder's patents—and can use the expected cost of this design-around to set a cap on what the potential user is willing to pay for a patent license. The cap might still be high enough that the patentee can exercise considerable holdout power, but the patentee might not know this.³¹ The threat of even an imperfectly specified design-around might commonly cause patentees to exercise restraint in their demands, and a plausible design-around will at least limit the holdout power that a patentee possesses.

Indeed, commentators sometimes argue that a positive justification for a patent system is “the incentive to design around” that patent rights provide.³² The Federal Circuit itself has declared that “[o]ne of the benefits of a patent system is its so-called ‘negative incentive’ to ‘design around’ a competitor’s products, . . . thus bringing a steady flow of innovations to the marketplace.”³³ Mindful of Fritz Machlup’s counterarguments that design-around activity tends to be socially wasteful,³⁴ as well as Edmund Kitch’s related suggestion that patents can help limit duplicative work on follow-on

30. See, e.g., SUBCOMM. ON PATENTS, TRADEMARKS, & COPYRIGHTS, S. COMM. ON THE JUDICIARY, 85TH CONG., AN ECONOMIC REVIEW OF THE PATENT SYSTEM 50 (Comm. Print 1958) (prepared by Fritz Machlup) [hereinafter ECONOMIC REVIEW OF THE PATENT SYSTEM] (describing how “‘invent[ing] around’ [a] competitor’s patent” can enable a rival to compete with a patent holder); F. SCOTT KIEFF ET AL., PRINCIPLES OF PATENT LAW 71 (4th ed. 2008) (arguing that the “secondary inventive activity” of designing around patents is generally desirable because it helps ensure fuller exploration of technological alternatives); Golden, *supra* note 6, at 2130 (noting that under a specified model of party incentives where the expected cost of a design-around is less than the expected cost of simply litigating and potentially being subjected to court-ordered remedies, “the possibility of a design-around improves the potential infringer’s position . . . [and] should reduce the amount for which the potential infringer is willing to settle”).

31. Cf. Golden, *supra* note 6, at 2132 (“[A] patent holder will likely approach negotiations at a significant informational disadvantage with respect to the potential infringer’s expected costs.”).

32. KIEFF ET AL., *supra* note 30, at 70; see *id.* at 71 (arguing that “[w]hile at first blush [designing around] may seem wasteful, as redundant, it becomes immediately apparent that such secondary inventive activity is usually a very good thing” because “[o]ften, a second-generation product is better than the first”); ROGER E. SCHECHTER & JOHN R. THOMAS, INTELLECTUAL PROPERTY: THE LAW OF COPYRIGHTS, PATENTS AND TRADEMARKS § 13.4.1, at 289 (2003) (“The Patent Act is also thought to stimulate technological advancement by inducing individuals to ‘invent around’ patented technology.”); see also ECONOMIC REVIEW OF THE PATENT SYSTEM, *supra* note 30, at 51 (“[F]rom merely defending the need of ‘inventing around a patent’ as a minor item of waste, the discussion has recently proceeded to eulogize it as one of the advantages of the system, indeed as one of its ‘justifications.’” (footnotes omitted)).

33. State Indus., Inc. v. A.O. Smith Corp., 751 F.2d 1226, 1236 (Fed. Cir. 1985).

34. See ECONOMIC REVIEW OF THE PATENT SYSTEM, *supra* note 30, at 51 (“The production of the knowledge of how to do in a somewhat different way what we have already learned to do in a satisfactory way would hardly be given highest priority in a rational allocation of resources.”).

innovation,³⁵ I view design-around activity primarily as a means to mitigate the patent system's costs. If one can tolerate an analogy between patent law and roadwork, one might similarly recognize that roadwork often produces temporary blockages or slowdowns that can cause drivers to experiment with new ways to travel.³⁶ The likelihood of causing such experimentation is unlikely to be considered a positive justification to undertake the roadwork in the first instance, but such design-around activity can reduce the social costs of the work, therefore helping to make those costs tolerable.

B. Injunctions as Gateways to Potential Contempt

If one accepts at least a weak-form argument for the desirability of design-around activity, one is then likely to face a special concern with injunctions. An injunction of unclear scope can generate a special risk that socially desirable design-around activity will be chilled. This is largely because of the stronger and quicker sanctions that proceedings to enforce an injunction can make available. Generally speaking, injunctions are in personam orders that are enforced through comparatively summary proceedings invoking a court's contempt power.³⁷ Such proceedings can be criminal or civil in nature.³⁸ If an enjoined party is found guilty of criminal contempt, a court may order determinate sanctions, such as an unconditional fine or jail term, to punish the contemnor and "vindicate[] the court's authority."³⁹ Unless contempt is committed directly in the presence of the court,⁴⁰ however, criminal contempt proceedings need to comply with rules

35. Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265, 276 (1977) (describing how patent law "puts the patent owner in a position to coordinate the search for technological and market enhancement of the patent's value so that duplicative investments are not made and so that information is exchanged among the searchers").

36. Cf. ECONOMIC REVIEW OF THE PATENT SYSTEM, *supra* note 30, at 52 ("The continental blockade in the Napoleonic War led to the development of beet sugar; the blockade in World War I led to the process of obtaining nitrogen from air; the U-boat blockade in World War II led to the invention of atabrine as a substitute for quinine; etc., etc. Does it follow that it would be a good idea to institute more blockades?").

37. See 1 DAN B. DOBBS, *DOBBS LAW OF REMEDIES* § 2.8(1), at 186 (2d ed. 1993) ("Because [equitable decrees] are personal orders, they are often enforced coercively, through the contempt power.").

38. *Id.* at 187 (discussing the possibility of civil or criminal contempt proceedings); see also *KSM Fastening Sys., Inc. v. H.A. Jones Co.*, 776 F.2d 1522, 1524 & n.2 (Fed. Cir. 1985) (noting the possibility of "civil and criminal contempt proceedings"), *overruled on other grounds*, *TiVo Inc. v. EchoStar Corp.*, 646 F.3d 869 (Fed. Cir. 2011) (en banc); James C. Nemmers, *Enforcement of Injunctive Orders and Decrees in Patent Cases*, 7 IND. L. REV. 287, 291 (1973) (discussing the use of civil and criminal contempt in patent cases).

39. 1 DOBBS, *supra* note 37, § 2.8(3), at 196-97; see also *Spindelfabrik Suessen-Schurr v. Schubert & Salzer Maschinenfabrik Aktiengesellschaft*, 903 F.2d 1568, 1578 (Fed. Cir. 1990) ("A civil contempt sanction is remedial, and for the benefit of the complainant[,] while a criminal contempt sentence is punitive, to vindicate the authority of the court." (internal quotation marks omitted)).

40. FED. R. CRIM. P. 42(b) ("Notwithstanding any other provision of these rules, the court (other than a magistrate judge) may summarily punish a person who commits criminal contempt in

of criminal procedure,⁴¹ presumably including a requirement for proof beyond a reasonable doubt that the violation of the injunction was intentional.⁴² Perhaps in substantial part because of this, criminal contempt for violation of a patent-infringement injunction appears to be so rare in U.S. patent practice that its possibility is likely to be negligible for most practical purposes.⁴³

In contrast, civil contempt can loom as a very real threat for an adjudged infringer. Civil contempt sanctions may be coercive or compensatory—i.e., designed “coercively to induce compliance with the court’s decree or remedially to obtain for the plaintiff a substitute for the defendant’s compliance.”⁴⁴ The general legal rule is that compensatory civil contempt

its presence if the judge saw or heard the contemptuous conduct and so certifies; a magistrate judge may summarily punish a person as provided in 28 U.S.C. § 636(e).”)

41. 1 DOBBS, *supra* note 37, § 2.8(1), at 187 (observing that for purposes of imposing “a criminal sanction,” “the incidents of trial must comport with all the relevant rules of criminal procedure, including the constitutional protections afforded to those charged with crime” and possibly including a right to a jury trial and a requirement of “proof beyond a reasonable doubt”); *see also* FED. R. CRIM. P. 42(a) (providing notice and other procedural requirements for criminal contempt proceedings); *Victor Stanley, Inc. v. Creative Pipe, Inc.*, 269 F.R.D. 497, 537–38 (D. Md. 2010) (noting that in criminal contempt proceedings, “the court must refer the matter to the United States Attorney for prosecution,” “appoint a private prosecutor” if the U.S. Attorney declines “(a highly probable outcome in most instances),” demand proof “beyond a reasonable doubt,” and provide for “a jury trial if the sentence will be longer than six months” (citations omitted)).

42. *See TWM Mfg. Co. v. Dura Corp.*, 722 F.2d 1261, 1272 (6th Cir. 1983) (stating, in reviewing a holding of criminal contempt for violation of a patent-infringement injunction, that “[i]n criminal contempt, willful disobedience must be proved beyond a reasonable doubt” and that “[w]illfulness, for this purpose, implies a deliberate or intended violation, as distinguished from an accidental, inadvertent or negligent violation”); Nemmers, *supra* note 38, at 295 (“A criminal contempt proceeding is a crime ‘in the ordinary sense,’ and therefore the acts of the accused must be shown beyond a reasonable doubt to have been willful and deliberate.” (footnotes omitted)).

43. An electronic search of all federal court opinions in the Westlaw database since 1970 that used “criminal contempt” and “patent” in the same paragraph turned up no case in which a district court made a holding of criminal contempt that was not reversed on appeal, and two cases in which such holdings were reversed. *See Spindelfabrik*, 903 F.2d at 1580 (holding that a “\$2 million fine” “constituted punishment for criminal contempt, and cannot stand because it was imposed without following the requisite procedures for criminal contempt”); *TWM*, 722 F.2d at 1272 (holding that “the evidence [did] not support a finding of criminal contempt”). In 1973, James Nemmers reported that he was able to identify only two reported cases “in which criminal contempt was clearly charged for violation of an injunction in a patent case”—one from 1970 and another from 1911. Nemmers, *supra* note 38, at 291 n.20 (citing *United States ex rel. Shell Oil Co. v. Barco Corp.*, 430 F.2d 998 (8th Cir. 1970); *Kreplik v. Couch Patents Co.*, 190 F. 565 (1st Cir. 1911)). More generally, Nemmers reported that courts have historically tended to find criminal contempt principally in situations involving “‘direct’ contempt[.]”—i.e., contemptuous behavior “committed in the presence of the court.” *Id.* at 289, 291. *See generally* FED. R. CRIM. P. 42(b) (“Notwithstanding any other provision of these rules, the court (other than a magistrate judge) may summarily punish a person who commits criminal contempt in its presence if the judge saw or heard the contemptuous conduct and so certifies; a magistrate judge may summarily punish a person as provided in 28 U.S.C. § 636(e).”). Violation of patent-infringement injunctions can be expected rarely, if ever, to constitute such direct contempt. *See* Nemmers, *supra* note 38, at 289–90 (“Since violations of injunctive relief in patent cases do not occur in the presence of the court, contempts considered hereinafter will be in the category of ‘indirect’ contempts . . .”).

44. 1 DOBBS, *supra* note 37, § 2.8(3), at 197; *see also Spindelfabrik*, 903 F.2d at 1578 (“Judicial sanctions in civil contempt proceedings may, in a proper case, be employed for either or

sanctions are confined to the “complainant’s actual loss.”⁴⁵ Nonetheless, the Federal Circuit has held, perhaps questionably,⁴⁶ that in civil contempt proceedings, district courts may enhance awards for actual damages up to at least treble damages⁴⁷ just as they may enhance damages up to treble their actual size in ordinary infringement proceedings.⁴⁸ On the other hand, treble damages might well represent a practical limit on the sort of noncoercive civil sanctions the Federal Circuit will tolerate, and they appear likely to be reserved for particularly obnoxious—and thus readily avoidable—forms of contempt. In *Spindelfabrik Suessen-Schurr v. Schubert & Salzer Maschinenfabrik Aktiengesellschaft*,⁴⁹ the Federal Circuit upheld treble damages for civil contempt in a situation where the district court “justifiably characterized [the contemnor’s] actions as ‘flagrant contemptuous conduct.’”⁵⁰ Likewise, in *Stryker Corp. v. Davol Inc.*,⁵¹ the Federal Circuit upheld an award of treble damages for (presumably civil) contempt in light of “the conclusory nature of the opinion of counsel used . . . to justify selling [a] revised device and the minor changes made to the original [infringing] device.”⁵²

For civil contempt, courts may impose larger monetary sanctions in the form of fines, but these fines must generally be conditional on future noncompliance. Specifically, a coercive civil contempt decree may order imprisonment or payment of a fine unless or until the contemnor complies with the underlying injunction.⁵³

both of two purposes: to coerce the defendant into compliance with the court’s order, and to compensate the complainant for losses sustained.” (internal quotation marks omitted).

45. See *United States v. United Mine Workers*, 330 U.S. 258, 304 (1947) (declaring that a compensatory “fine must of course be based upon evidence of [the] complainant’s actual loss”).

46. See *Broadview Chem. Corp. v. Loctite Corp.*, 311 F. Supp. 447, 453 (D. Conn. 1970) (“[T]o the extent that double or treble damages serve a punitive purpose, they may not be awarded in a civil contempt proceeding.”); Nemmers, *supra* note 38, at 306 (“The increased damages provisions of [35 U.S.C. §] 284 for a deliberate infringement are punitive; and although the contemnor may lose his profits, he cannot be assessed punitive damages in a civil contempt proceeding.”).

47. *Spindelfabrik*, 903 F.2d at 1578 (holding that where a district court “justifiably characterized . . . actions as ‘flagrant contemptuous conduct[,]’” “the district court did not abuse its discretion in trebling the damages and awarding attorney fees” in a civil contempt proceeding, but that the district court erred in imposing a \$2 million fine without following criminal contempt procedures because the fine was not conditional on noncompliance and thus constituted “punishment for criminal contempt”).

48. See 35 U.S.C. § 284 (2006) (empowering district courts to “increase the damages up to three times the amount found or assessed”).

49. 903 F.2d 1568 (Fed. Cir. 1990).

50. *Id.* at 1578.

51. 234 F.3d 1252 (Fed. Cir. 2000).

52. *Id.* at 1260.

53. 1 DOBBS, *supra* note 37, § 2.8(3), at 197; see *Spindelfabrik*, 903 F.2d at 1578–79 (“The ability of the contemnor to avoid the sanction by complying with the court order is an important factor in determining whether a contempt adjudication is civil or criminal.”). See generally *United States v. United Mine Workers*, 330 U.S. 258, 304 (1947) (stating that, in civil contempt proceedings, a court seeking to use a fine “to make the defendant comply” with the court’s earlier order “must . . . consider the character and magnitude of the harm threatened by continued

Alternatively or additionally, courts have held that a civil contempt decree may order disgorgement of the contemnor's profits attributable to contempt⁵⁴ and payment of the rights holder's court costs and attorneys' fees.⁵⁵ A district court may also modify the initial injunctive decree, perhaps expanding its scope in hopes that the resulting amplified injunction will provide greater protection of the patentee's legitimate interests.⁵⁶ Such contempt sanctions can result from proceedings that "are generally summary in nature and may be decided by the court on affidavits and exhibits without the formalities of a full trial, although the movant bears the heavy burden of proving violation by clear and convincing evidence."⁵⁷ Further, "lack of intent to violate an injunction alone cannot save an infringer from a finding of [civil] contempt," although "diligence and good faith efforts . . . may be considered in assessing penalties."⁵⁸

Nonetheless, despite the potential procedural speed of contempt proceedings and the heavier sanctions that they can make available, there is a good chance that an adjudged infringer made subject to a patent-infringement injunction will find concerns with patent scope and injunction scope to be directly comparable. When any threat of being found in contempt is realistically limited to a threat of civil contempt and when, furthermore, any stigma associated with civil (as opposed to criminal) contempt can be assumed negligible,⁵⁹ a risk of being found in contempt can essentially

contumacy, and the probable effectiveness of any suggested sanction in bringing about the result desired").

54. See *Schaefer Fan Co. v. J&D Mfg.*, 265 F.3d 1282, 1290 (Fed. Cir. 2001) (upholding a district court's award of contempt sanctions equaling the contemnor's "total profit" from the sale of two types of fans); *Brine, Inc. v. STX, L.L.C.*, 367 F. Supp. 2d 61, 71 (D. Mass.) ("[A] sanction in the amount of gross profit from the sales of the X2+ provides a natural means of imposing a penalty that is proportionate to the severity of the contempt."), *aff'd*, 139 F. App'x 281 (Fed. Cir. 2005); Nemmers, *supra* note 38, at 305 ("It appears settled that the complainant is entitled to the contemnor's profits" "even though 'profits' of the infringer are not recoverable in the ordinary patent infringement action."). As with treble damages for contempt, disgorgement of the entirety of a contemnor's profits appears to be considered an extreme sanction reserved for egregious behavior. See *Schaefer*, 265 F.3d at 1290 (holding disgorgement of the contemnor's "total profit" to be justified where the contemnor had previously been found in contempt and had "acted willfully by failing to obtain any opinion of counsel on whether [certain products] would breach the agreement and the court's orders").

55. Nemmers, *supra* note 38, at 307 (noting a contempt complainant's ability "to recover his attorney's fees and costs and expenses incurred in conducting the civil contempt proceeding," subject to the trial court's discretion); see also *Stryker*, 234 F.3d at 1260 (holding that the district court "did not abuse its discretion in awarding treble the compensatory royalty damages, attorney fees, and costs").

56. See *Spindelfabrik*, 903 F.2d at 1577 (holding that in contempt proceedings, "[t]he district court did not abuse its discretion in broadening the injunction to cover 'any automated rotor spinning machine,' without the qualifying word 'infringing'").

57. *KSM Fastening Sys., Inc. v. H.A. Jones Co.*, 776 F.2d 1522, 1524 (Fed. Cir. 1985), *overruled on other grounds*, *TiVo Inc. v. EchoStar Corp.*, 646 F.3d 869 (Fed. Cir. 2011) (en banc).

58. *TiVo*, 646 F.3d at 880.

59. Concerns about serious stigma associated with contempt seem more strongly associated with criminal contempt than with civil contempt. See, e.g., Richard B. Kuhns, *The Summary*

amount to no more than a risk of being subjected to heightened but still limited monetary sanctions.⁶⁰ Moreover, these heightened monetary sanctions will only attach if the patentee succeeds in showing violation of an injunction under a heightened burden of proof—one requiring clear and convincing evidence.⁶¹ Under such circumstances, an injunction, rather than acting as a clear off switch for infringement, serves instead primarily as a gateway to a potential set of enhanced monetary remedies.⁶²

Contempt Power: A Critique and a New Perspective, 88 YALE L.J. 39, 79 n.235 (1978) (expressing concern for “the stigma of a conviction” for criminal contempt); F. Joseph Warin & Michael D. Bopp, *Corporations, Criminal Contempt and the Constitution: Do Corporations Have a Sixth Amendment Right to Trial by Jury in Criminal Contempt Actions and, if So, Under What Circumstances?*, 1997 COLUM. BUS. L. REV. 1, 40 n.161 (“The collateral effects of a criminal contempt finding include the stigma that attaches in such circumstances.”); Douglas C. Berman, Note, *Coercive Contempt and the Federal Grand Jury*, 79 COLUM. L. REV. 735, 758 (1979) (“[C]oercive contempt incarceration does not entail the limitations of civil rights (such as disenfranchisement) that normally attend a felony conviction, and . . . the stigma of being a convicted criminal does not attach.”); Kathleen A. Burdette, Comment, *Making Parents Pay: Interstate Child Support Enforcement After United States v. Lopez*, 144 U. PA. L. REV. 1469, 1527 (1996) (“[C]riminal convictions carry a much more severe social stigma than do civil contempt orders.”). At least one commentator has suggested that even criminal contempt tends not to carry much stigma. See Eric L. Jensen, *The Waiver of Juveniles to Criminal Court: Policy Goals, Empirical Realities, and Suggestions for Change*, 31 IDAHO L. REV. 173, 174 n.4 (1994) (excluding “criminal contempt violations” from the scope of a discussion of the treatment of criminal activity by juveniles because of “the absence of strong criminal stigma attached” to such violations). But other commentators have suggested that even civil contempt can inflict significant stigma. See, e.g., Earl C. Dudley, Jr., *Getting Beyond the Civil/Criminal Distinction: A New Approach to the Regulation of Indirect Contempts*, 79 VA. L. REV. 1025, 1065 (1993) (“It may seriously be doubted whether any nonlawyers are sufficiently aware of—let alone appreciate the import of—the [civil-versus-criminal contempt] distinction for it to make any real difference in the opprobrium attaching to a contempt judgment.”); Recent Case, *International Bhd. of Teamsters v. NLRB (D.C. Cir. 1958)*, 72 HARV. L. REV. 1577, 1580 (1959) (expressing concern that even when “only civil penalties” are likely to be imposed, “the threat of a large fine and the stigma of citation for contempt may inhibit [labor] activities beyond those specifically restrained” (footnote omitted)); Note, *Procedures for Trying Contempts in the Federal Courts*, 73 HARV. L. REV. 353, 357 (1959) (suggesting “the possibility [that] stigma . . . may result from the imposition of sanctions similar in form to criminal punishment”).

60. An infringer might perceive the heightening of sanctions as especially sharp if the infringer were insured for ordinary and perhaps even willful infringement but not for contempt. But patent-litigation insurance of any stripe appears to be relatively rare. See CJA CONSULTANTS LTD., PATENT LITIGATION INSURANCE: A STUDY FOR THE EUROPEAN COMMISSION ON POSSIBLE INSURANCE SCHEMES AGAINST PATENT LITIGATION RISKS § 7.7.1 (2003) (“Contrary to received belief, the extent of Patent litigation insurance in the USA in relation to the extent of litigation appears to be small, and limited to defence, including damages.”); see also Colleen V. Chien, *Predicting Patent Litigation*, 90 TEXAS L. REV. 283, 295 (2011) (“[T]he market for patent insurance is extremely small and highly inefficient. Offerings are limited and expensive. Defensive policies . . . fail to cover many situations.” (footnotes omitted) (internal quotation marks omitted)). Thus, an adjudged infringer is not likely to perceive a special pinch from lack of insurance coverage for contempt sanctions.

61. *TiVo*, 646 F.3d at 883 (“The patentee bears the burden of proving violation of the injunction by clear and convincing evidence . . .”).

62. Because monetary sanctions for civil contempt are generally limited to quite finite values, contempt sanctions and the injunctions that they back up do not act as the sort of limiting endpoint for a liability-rule continuum that Kaplow and Shavell have associated with a true property rule. Kaplow & Shavell, *supra* note 10, at 756 (“[T]he property rule protecting victims mirrors a liability

Particularly in situations where the expected cost of complying with an injunction is great—or alternatively stated, where the expected value of continuing to violate patent rights is large—an adjudged infringer can face a set of options substantially analogous to those faced when mulling the risk of an infringement suit alone. A course of action that risks a finding of contempt might nonetheless be profitable; indeed, it might be an adjudged infringer's most profitable option. One might speculate that this was the situation in *TiVo Inc. v. EchoStar Corp.*, where, in determining contempt sanctions, the district court had settled on an award for EchoStar's contempt of approximately \$200 million⁶³ after refusing TiVo's request for disgorgement of approximately \$1 billion in profits that TiVo contended was attributable to EchoStar's contempt.⁶⁴

More generally, civil contempt sanctions—which judicial practice suggests are likely to amount to treble damages or less⁶⁵—will predictably act as only imperfect deterrents when the damages awarded in the original infringement proceedings constitute only a small percentage of the infringer's profits. Such a situation is likely to occur reasonably frequently because a patented invention often forms only one part of a much more intricate infringing product or process⁶⁶ and thus, particularly in light of a recent push to ensure proper proportionality of reasonable-royalty damages, will justify compensatory damages equaling only a fraction of the infringer's profits.⁶⁷ The resulting discrepancy between infringer profits and assessed patent value might make particularly probable a court's denial of an injunction on grounds that the discrepancy will give the patentee undue leverage in subsequent licensing negotiations. Such was the case in *Paice LLC v. Toyota Motor Corp.*,⁶⁸ where the jury award of \$25 per vehicle amounted to no more than about 0.1% of each infringing vehicle's typical total price.⁶⁹ But when

rule with extremely high, or infinite, damages.”). A patent-infringement injunction does not generally equate with an award of infinite damages for continued infringement.

63. *TiVo Inc. v. Dish Network Corp.*, 655 F. Supp. 2d 661, 666 (E.D. Tex. 2009) (deciding on an award of “approximately \$110 million in compensation based on the jury’s [prior damages] award and approximately \$90 million in sanctions,” plus “attorney’s fees and costs” for the contempt proceedings).

64. *Id.* at 665 (“TiVo’s suggested disgorgement of nearly \$1 billion is unreasonable under the circumstances of this case.”).

65. See *supra* note 54 and text accompanying notes 46–52.

66. Cf. Lemley & Shapiro, *supra* note 2, at 2009 (“A microprocessor may include 5,000 different inventions . . .”).

67. See Bo Zeng, Note, *Lucent v. Gateway: Putting the “Reasonable” Back into Reasonable Royalties*, 26 BERKELEY TECH. L.J. 329, 366 (2011) (observing that in various recent cases, “the Federal Circuit made a critically important effort to ensure that reasonable royalty damage awards are actually reasonable”).

68. No. 2:04-CV-211-DF, 2006 WL 2385139, at *6 (E.D. Tex. Aug. 16, 2006) (denying the patentee’s motion for a permanent injunction), *aff’d in part and vacated in part*, 504 F.3d 1293 (Fed. Cir. 2007).

69. *Id.* at *3 (quoting defendants’ observation that the jury’s award amounted to “1/8th of one percent of the \$20,000 price of a Prius and even less of a percentage of the price of the Highlander (\$33,000) and the RX400h (\$42,000)”).

an injunction issues in such situations, an infringer might nonetheless risk contempt with substantial confidence that as long as the infringer has taken reasonable steps to avoid a finding of flagrant contempt, any contempt sanctions will amount to no more than about three times a prior, small-percentage jury award.

In short, at least when sanctions for criminal contempt are no more than a remote possibility, injunctions' effects on the behavior of their targets can be substantially similar to those of injunctions' oft-supposed opposite—"liability rules, under which [the state] merely discourages violations by requiring transgressors to pay victims for harms suffered."⁷⁰ In at least superficial conformity with Louis Kaplow and Steven Shavell's suggestion that merely "probabilistic invasions of property interests" are frequently governed properly and "primarily by a liability rule,"⁷¹ the liability-rule-like nature of injunctions might extend to various situations where courts issue an injunction because there is a strong interest in providing extra deterrence or recompense for continued infringement⁷² but where the scope of underlying rights is uncertain and there is also a strong social interest in not discouraging, or even in affirmatively encouraging, activity that treads close to underlying rights' boundaries.

C. Modeling Infringer Incentives

Whatever the generality of the insight about injunctions' potential liability-rule-like nature, options available to a party enjoined for patent

70. Kaplow & Shavell, *supra* note 10, at 715 (emphasis omitted). Paul Heald has separately analogized injunctions to damages awards. Paul J. Heald, *Permanent Injunctions as Punitive Damages in Patent Infringement Cases* 3 (Ill. Pub. Law & Legal Theory Research Paper No. 10-38, 2011) (discussing the "analogy injunctions can bear to punitive damages"), available at <http://ssrn.com/abstract=1851681>. But, at least in contexts of high design-around costs, his analogy is to punitive damages, *id.*, which commentators have commonly associated with property rules rather than liability rules. See, e.g., Richard Craswell, *Property Rules and Liability Rules in Unconscionability and Related Doctrines*, 60 U. CHI. L. REV. 1, 4 (1993) (observing that in contract law, "punitive damages" operate to protect a "promise with a property rule"); William S. Dodge, *The Case for Punitive Damages in Contracts*, 48 DUKE L.J. 629, 634 (1999) (contending that "contractual entitlements should be protected with property rules, including punitive damages"); Henry E. Smith, *Mind the Gap: The Indirect Relation Between Ends and Means in American Property Law*, 94 CORNELL L. REV. 959, 966 (2009) (describing "property rules" as "embodied in injunctions and punitive damages").

71. Kaplow & Shavell, *supra* note 10, at 716.

72. Deterrence or recompense made available through an injunction can be viewed as "extra" to the extent net contempt sanctions or awards to the patentee are expected to exceed those available through ordinary infringement proceedings, whether because contempt proceedings are expected to yield higher monetary awards or because enforcement through contempt is more likely to occur due to contempt proceedings' greater speed and presumably lower cost, etc. Of course, there is a background question of when such a higher level of effective remedies should be viewed as desirable at all. See *id.* at 773 (contending that in situations involving "harmful externalities" and an "absence of bargaining between victims and injurers, a liability rule with damages equal to estimated harm is unambiguously superior to property rules even though actual harm in a given case may be difficult to determine").

infringement commonly include (1) paying the patentee for a license, (2) halting all activity potentially barred by the injunction, and (3) redesigning the infringing product or process. Patent and injunction scope affect the relative desirability of these options. Greater patent scope increases the range of redesigns that are likely to infringe and thus tends to raise the cost of redesign either by forcing the infringer to choose a more radical redesign or by raising the risk that a given redesign will infringe. Likewise, greater injunction scope increases the range of redesigns that will put the infringer at risk of contempt, with similar implications for the likely cost of redesign. If the probability of a contempt holding or the penalty for contempt is sufficiently small, a rational infringer might choose to implement a redesign that risks contempt, just as the infringer might previously have chosen an original design that risked a holding of infringement.

A relatively simple mathematical model helps illustrate the nature of an adjudged infringer's potential decision-making calculus. Consider a situation where a product has been found to infringe, and a rational, profit-maximizing infringer has three basic alternatives:

- (1) pay for a license, which for simplicity we assume will require payment of a lump-sum licensing fee, F ;
- (2) implement an "ironclad" redesign that will cost D_1 to implement and that will be generally recognized, by the patentee as well as the courts, to be both injunction compliant and noninfringing;
- (3) implement a more doubtful redesign that will cost a lower amount, D_2 (i.e., $D_2 < D_1$), but that carries (i) a nonzero chance, θ_{con} , of resulting in a holding of contempt and (ii) a further nonzero chance, θ_{inf} , of resulting in a holding of infringement even when no contempt is found.

Because the adjudged infringer's willingness to pay for a license will depend on the expected profitability of redesign options (2) and (3),⁷³ I focus on these options below.

To model the profitability of options (2) and (3), I use a variant of a model that Carl Shapiro developed with an eye to infringement concerns only.⁷⁴ In particular, I suppose that in the absence of additional threats of litigation, the product resulting from either redesign will generate a constant

73. The adjudged infringer will rationally only pay for a license if the profitability of that course is at least as great as the more profitable of options (2) and (3). In terms of the model that follows, the rational infringer's willingness to pay for a license is thus capped by a licensing cost equaling the lesser of D_1 and D_2 , where the value of D_2 is indicated by Equation 4, *infra* at text accompanying notes 77–78. Transaction costs of negotiating and complying with a license will likely mean that the rational infringer's maximum licensing fee is lower than this maximum acceptable licensing cost.

74. See Carl Shapiro, *Injunctions, Hold-Up, and Patent Royalties*, 12 AM. L. & ECON. REV. 280, 286–87 (2010) (developing a model to determine potential outcomes of royalty negotiations between a "downstream firm" and a "patent holder"); see also Lemley & Shapiro, *supra* note 2, at 1995–97 (using Shapiro's model).

profit margin M per unit sold, with N units expected to be sold, and $MN \geq D_1$.⁷⁵ I also assume that the patentee will not sue in response to the ironclad redesign in option (2) but that, if the adjudged infringer implements the more doubtful redesign in option (3), the patentee will file a motion for contempt. By assumption, the parties will not settle, and the probability that the patentee's motion will yield a holding of contempt is θ_{con} , where $0 \leq \theta_{\text{con}} \leq 1$. If contempt is not found, the patentee will file a new suit alleging patent infringement. Once again, the parties will not settle. The patentee's probability of success in the new suit will be θ_{inf} , where $0 \leq \theta_{\text{inf}} \leq 1$.

Significantly, the probabilities θ_{con} and θ_{inf} should correlate with injunction scope and patent scope, respectively. Invalidity and enforceability of underlying patent claims are assumed not to be at issue in the proceedings contemplated, whether due to the law of the case for contempt proceedings⁷⁶ or estoppel in a subsequent suit for infringement.⁷⁷ Thus, a broader injunction will likely generate a greater probability of patentee success in a contempt proceeding, θ_{con} , and a broader patent will likely generate a greater probability of patentee success in an infringement suit, θ_{inf} . A model incorporating the values θ_{con} and θ_{inf} can therefore suggest how injunction and patent scope affect the decision-making calculus of an adjudged infringer.

But there are additional variables that enter this calculus. Suppose that by the time contempt proceedings are concluded, the adjudged infringer will have made and sold n_{con} units of the redesigned product and will also have incurred L_{con} in contempt-proceeding litigation costs. If the infringer is held not to be in contempt, the infringer will then sell an additional n_{inf} units and incur an additional L_{inf} in litigation costs before the conclusion of the infringement proceedings. On the other hand, if the infringer is held in contempt, the infringer will have to pay the equivalent of a monetary penalty

75. For simplicity, I neglect time discounting for products sold after a period of time has elapsed.

76. See *KSM Fastening Sys., Inc. v. H.A. Jones Co.*, 776 F.2d 1522, 1529 (Fed. Cir. 1985) ("The validity of the patent is the law of the case in [contempt] proceedings."), *overruled on other grounds*, *TiVo Inc. v. EchoStar Corp.*, 646 F.3d 869 (Fed. Cir. 2011) (en banc).

77. To the extent defenses of invalidity and unenforceability were available, the adjudged infringer will likely have challenged validity and enforceability in the original litigation. See Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 NW. U. L. REV. 1495, 1502 (2001) ("Virtually every patent infringement lawsuit includes a claim that the patent is either invalid or unenforceable . . . (or commonly both)."). Thus, the infringer might be precluded from making such challenges in later litigation. See, e.g., *Roche Palo Alto LLC v. Apotex, Inc.*, 531 F.3d 1372, 1381 (Fed. Cir. 2008) (affirming a district court's holding that "validity challenges . . . were barred by the doctrine of claim preclusion"); *Power Integrations, Inc. v. Fairchild Semiconductor Int'l, Inc.*, 763 F. Supp. 2d 671, 678–79 (D. Del. 2010) (holding that various arguments for invalidity were barred by issue preclusion); cf. *Foster v. Hallco Mfg. Co.*, 947 F.2d 469, 480–81 (Fed. Cir. 1991) ("[I]f a consent judgment, by its terms, indicates that the parties thereto intend to preclude any challenge to the validity of a particular patent, even in subsequent litigation involving a new cause of action, then that issue can be precluded.").

P and will have its postcontempt expected profits reduced by δ_{con} from $M(N - n_{\text{con}})$ to π_{con} . Finally, if the infringer prevails in the contempt proceeding but the redesigned product is later held to infringe, the infringer will pay a damages award A and have its postjudgment expected profits reduced by δ_{inf} from $M(N - n_{\text{con}} - n_{\text{inf}})$ to π_{inf} .

Under these circumstances, the infringer's expected gains from the two redesigns are as follows:

- (1) expected gain from pursuing the ironclad redesign:

$$G_1 = MN - D_1; \quad (\text{Eq. 1})$$

- (2) expected gain from pursuing the more doubtful redesign:

$$G_2 = Mn_{\text{con}} - D_2 - L_{\text{con}} + \theta_{\text{con}}(\pi_{\text{con}} - P) \\ + (1 - \theta_{\text{con}})\{Mn_{\text{inf}} - L_{\text{inf}} + \theta_{\text{inf}}(\pi_{\text{inf}} - A) + (1 - \theta_{\text{inf}})M(N - n_{\text{con}} - n_{\text{inf}})\}. \quad (\text{Eq. 2})$$

The formula for G_2 looks complicated but can be significantly simplified by (1) recognizing that $\delta_{\text{con}} = M(N - n_{\text{con}}) - \pi_{\text{con}}$ and $\delta_{\text{inf}} = M(N - n_{\text{con}} - n_{\text{inf}}) - \pi_{\text{inf}}$, and (2) defining new variables $\Delta_{\text{con}} = P + \delta_{\text{con}}$ and $\Delta_{\text{inf}} = A + \delta_{\text{inf}}$. These last variables, Δ_{con} and Δ_{inf} , represent sums of (a) the cost of a court-imposed penalty or damages award (P or A) and (b) the decrease in postjudgment profits expected to result from an adverse court decision (δ_{con} or δ_{inf}). Hence, in this relatively simple model, Δ_{con} and Δ_{inf} equal the total expected costs to the adjudged infringer of adverse judgments in contempt and infringement proceedings, respectively.

Use of the new Δ variables leads to the following equation for G_2 :

$$G_2 = MN - D_2 \quad (\text{Eq. 3})$$

where the total effective cost of the more questionable redesign D_2 satisfies the formula

$$D_2 = D_2 + L_{\text{inf}} + \theta_{\text{inf}} \Delta_{\text{inf}} + L_{\text{con}} + \theta_{\text{con}} \Delta_{\text{con}} \quad (\text{Eq. 4})$$

with

$$\Delta_{\text{con}} = \Delta_{\text{con}} - L_{\text{inf}} - \theta_{\text{inf}} \Delta_{\text{inf}}. \quad (\text{Eq. 5})$$

Subtraction of $L_{\text{inf}} + \theta_{\text{inf}} \Delta_{\text{inf}}$ from Δ_{con} to give the value of Δ_{con} reflects the fact that if an adjudged infringer is found in contempt, it at least "saves" on the expected costs of facing subsequent infringement litigation.

Equations 3 through 5 neatly illustrate the comparability of questions of patent scope and injunction scope. Further, they put mathematical emphasis on a point made at the very beginning of this Article: injunctions are not mere off switches for infringement. An injunction's relative capacity to bring a halt to infringement is influenced by the scope of the injunction, the cost of compliance with the injunction, and the backstop penalties or remedies for a finding of contempt. Despite the issuance of an injunction, an adjudged infringer can continue to have a multiplicity of plausible options, perhaps particularly when violation of an injunction is unlikely to lead to

proceedings for criminal contempt.⁷⁸ As Doug Rendleman has observed, instead of complying straightforwardly with an injunction, an adjudged infringer “may dissemble, may claim that the injunction is vague and impossible or difficult to understand, may seek delay, may search for loopholes, and may change as little as possible to ‘obey.’”⁷⁹

In particular, under the model presented here, a rational economic actor will favor the ironclad redesign over the questionable redesign only when $G_1 \geq G_2$ or, equivalently, $D_1 \leq D_2$. Other than adding another source of foreseeable litigation costs, the prospect of contempt proceedings deters choice of the more questionable redesign by adding to D_2 a quantity that equals the product of the likelihood of a contempt finding, θ_{con} , and the total effective cost, A_{con} , of a finding of contempt.⁸⁰ The prospect of an infringement suit similarly deters choice of the more questionable redesign by adding to D_2 a quantity that equals the product of the likelihood of an infringement finding, θ_{inf} , and the total effective cost, A_{inf} , of a finding of infringement.⁸¹ Moreover, in situations where the rational infringer prefers the questionable redesign to the ironclad redesign (i.e., where $D_2 < D_1$) and where $A_{\text{con}} > 0$, higher values of θ_{con} and θ_{inf} correspond to higher values for the maximum licensing fee, $F_{\text{max}} \leq D_2 < D_1$, that the rational infringer is willing to pay.⁸²

Consequently, to the extent θ_{con} and θ_{inf} serve as proxies for injunction and patent scope, Equations 3 through 5 show injunction and patent scope to play fundamentally analogous roles in protecting patent value through deterrence and likely enhancement of licensing fees. Conversely, limitations on injunction and patent scope facilitate various kinds of postinjunction activity by discounting costs of potential infringement or contempt by the less-than-100% probabilities that infringement or contempt will actually be found. Patent law that seeks to optimize social welfare must try to tune injunction and patent scope—and thus quantities such as θ_{con} and θ_{inf} —to strike a proper

78. DOUGLAS LAYCOCK, MODERN AMERICAN REMEDIES: CASES AND MATERIALS 766–69 (4th ed. 2010) (discussing the categories of “criminal contempt, coercive civil contempt, and compensatory civil contempt” and describing a potential three-step process in which a court first “issues the injunction,” second, “adjudicates the first violations and threatens specific fines for further violations,” and third, “adjudicates further violations and collects the fines”).

79. DOUG RENDLEMAN, COMPLEX LITIGATION: INJUNCTIONS, STRUCTURAL REMEDIES, AND CONTEMPT 425 (2010).

80. See *id.* (stating that an adjudicated infringer’s “intuitive calculation considers his gain from [an] activity, reduced by the probability and severity of a sanction”).

81. Facial lack of parallelism between A_{inf} and A_{con} can be explained as follows. The relevant effective cost of a contempt holding equals the quantity $A_{\text{con}} = A_{\text{con}} - L_{\text{inf}} - \theta_{\text{inf}} A_{\text{inf}}$, rather than A_{con} , because in the absence of a contempt finding, the infringer will face infringement proceedings having an expected cost, $L_{\text{inf}} + \theta_{\text{inf}} A_{\text{inf}}$. The quantity A_{con} equals the extent to which A_{con} exceeds this alternative expected cost. As there is, by assumption, no alternative remedy for the patentee once contempt and infringement proceedings fail, A_{inf} by itself equals the extent to which A_{inf} exceeds the residual alternative expected cost—zero—of such nonexistent alternative remedies.

82. Regrouping terms on the right-hand side of Equation 4 yields

$$D_2 = D_2 + L_{\text{con}} + \theta_{\text{con}} A_{\text{con}} + (1 - \theta_{\text{con}})(L_{\text{inf}} + \theta_{\text{inf}} A_{\text{inf}}).$$

The conclusion in the text follows from noting that $(1 - \theta_{\text{con}}) \geq 0$ and, by assumption, $A_{\text{inf}} \geq 0$.

balance between competing concerns of protecting patent value and optimizing productive activity by others.⁸³

III. U.S. Law and a Taxonomy of Injunction Types

To a large extent, existing U.S. law on patent-infringement injunctions suggests a taxonomy that informs much of the discussion in this Article. This taxonomy consists of the five injunction types—Type-0, Type-1, Type-2, purely reparative, and specially tailored—that Part I described. In this part, I first discuss the taxonomy's three types of do-not-infringe injunctions and then discuss purely reparative and specially tailored injunctions.⁸⁴

A. Type-0 Through Type-2 Do-Not-Infringe Injunctions

The simplest of the taxonomy's do-not-infringe injunctions is what I term a Type-0, measure-zero injunction. Such an injunction explicitly prohibits only infringement that involves the specific devices or processes already adjudged to infringe. For example, if manufacture or sale of Energizer Holdings' Schick Quattro razor were held to infringe patent rights associated with the blade arrangement in Gillette's Mach3 razor,⁸⁵ an injunction permanently enjoining Energizer Holdings "from making, using, selling, offering to sell, or importing Schick Quattro razors" would be a Type-0, measure-zero injunction.

From the patentee's standpoint, potential problems with such an order are obvious. What if Energizer takes its Schick Quattro, alters the design of the handle so that it is arguably more ergonomic, and markets the resulting "new and improved" product as the "Schick Quarto"? If minor or irrelevant design-arounds can avoid the force of a Type-0 injunction, that injunction will often have little real-world significance. In mathematical terms, the Type-0 injunction will effectively be of "measure zero," a practical nullity in any effort to vindicate a patentee's substantive rights.⁸⁶

83. In reality, the proper balance will likely reflect a variety of other concerns as well. See, e.g., Golden, *supra* note 1, at 509–11 (discussing various goals and behaviors that patent law might seek to promote or affect).

84. The three "supercategories" of (a) do-not-infringe injunctions, (b) purely reparative injunctions, and (c) prophylactic injunctions might be viewed as at least somewhat parallel to the three classifications proposed by Owen Fiss in 1978: (a) "the preventive injunction, which seeks to prohibit some discrete act or series of acts from occurring in the future"; (b) "the reparative injunction, which compels the defendant to engage in a course of action that seeks to correct the effects of a past wrong"; and (c) "the structural injunction, which seeks to effectuate the reorganization of an ongoing social institution." FISS, *supra* note 5, at 7.

85. Cf. *Gillette Co. v. Energizer Holdings, Inc.*, 405 F.3d 1367, 1368 (Fed. Cir. 2005) ("Gillette sued Energizer Holdings, Inc. . . . alleging Energizer's QUATTRO[®], a four-bladed wet-shave safety razor, infringes certain claims of the '777 patent.").

86. See *supra* note 15.

Courts have taken two primary approaches to dealing with this Type-0 problem. First, they commonly issue Type-1, colorable-differences injunctions that explicitly prohibit infringement not only via the precise products or processes already adjudged to infringe, but also via products or processes “no more than colorably different” from them.⁸⁷ Second, courts generally recognize that despite Type-0 injunctions’ measure-zero language, Type-0 injunctions should be understood to have Type-1 effect. The Federal Circuit has explained:

[W]here an injunction is written narrowly against a particular infringing device, contempt may, nevertheless, be found on the basis of a modified infringing device. An enjoined party under a narrow decree will not be permitted to escape [its force] on a purely “in rem” theory that only a particular device is prohibited, where it is evident that the modifications do not avoid infringement and were made for the purpose of evasion of the court’s order. Again, the standard is whether the differences between the two devices are merely colorable.⁸⁸

At least in principle, the third category of do-not-infringe injunctions—Type-2, obey-the-law injunctions—offers another way to protect against the facial narrowness of a Type-0 order. A Type-2 injunction generally prohibits continued infringement of a particular patent or claim.⁸⁹ In contempt proceedings, however, Type-2 language will not be read to have such broad, obey-the-law effect. The requirement of “clear and convincing” evidence for a holding of contempt,⁹⁰ as opposed to the normal preponderance-of-the-evidence standard for proof of infringement,⁹¹ provides one reason why such language generally does not make all forms of subsequent infringement sub-

87. See, e.g., *Ariba, Inc. v. Emptoris, Inc.*, No. 9:07-CV-90, slip op. at 1–2 (E.D. Tex. Jan. 22, 2009) (prohibiting activities involving “1. the Emptoris software (versions 5.2, 6.0, 6.1 and 7.0) heretofore marketed by Emptoris; and 2. all other software not more than colorably different therefrom”), *aff’d per curiam*, No. 2009-1230, 2010 WL 55625 (Fed. Cir. Jan. 8, 2010); *Callaway Golf Co. v. Acushnet Co.*, No. 06-091-SLR, slip op. at 2 (D. Del. Nov. 10, 2008) (prohibiting activity involving “any of the Pro V1® line of golf balls . . . or any variations thereof not more than colorably different”), *vacated*, 576 F.3d 1331 (Fed. Cir. 2009).

88. *KSM Fastening Sys., Inc. v. H.A. Jones Co.*, 776 F.2d 1522, 1526 (Fed. Cir. 1985), *overruled on other grounds*, *TiVo Inc. v. EchoStar Corp.*, 646 F.3d 869 (Fed. Cir. 2011) (en banc).

89. See, e.g., *Retractable Techs. Inc. v. Becton, Dickinson & Co.*, No. 2:07-cv-00250-DF-CMC, slip op. at 2 (E.D. Tex. May 19, 2010) (prohibiting activity involving products found specifically to infringe or products “only colorably different therefrom” and further prohibiting “otherwise infringing or inducing others to infringe the Infringed Claims”), *aff’d in part, rev’d in part*, 653 F.3d 1296 (Fed. Cir. 2011); *Nomadix, Inc. v. Second Rule LLC*, No. CV07-1946 DDP (VBKx), slip op. at 2 (C.D. Cal. Mar. 31, 2009) (prohibiting “directly or indirectly infringing any of [five] U.S. Patents . . . in any way”).

90. *Abbott Labs. v. TorPharm, Inc.*, 503 F.3d 1372, 1382 (Fed. Cir. 2007) (“There must be clear and convincing evidence of patent infringement to support a district court’s finding of contempt.”).

91. *Creative Compounds, LLC v. Starmark Labs.*, 651 F.3d 1303, 1314 (Fed. Cir. 2011) (“The patentee bears the burden of proving infringement by a preponderance of the evidence.” (quoting *SRI Int’l v. Matsushita Elec. Corp. of Am.*, 775 F.2d 1107, 1123 (Fed. Cir. 1985))).

ject to contempt. More to the point, however, the Federal Circuit has held that in contempt proceedings, a Type-2, obey-the-law injunction should be narrowly construed to apply only to products or processes “previously admitted or adjudged to infringe, and to other devices which are no more than colorably different therefrom and which clearly are infringements.”⁹² In other words, when presiding over contempt proceedings, a judge should effectively reform a Type-2, obey-the-law injunction so that its effective scope is no greater than that of a Type-1, colorable-differences order.

Authority for such judicial reformation derives at least partially from the Federal Circuit’s separate conclusion that Type-2, obey-the-law injunctions are technically prohibited and thus subject to vacatur on direct appeal. The U.S. Patent Act gives district courts power to “grant injunctions . . . on such terms as [they] deem[] reasonable.”⁹³ But consistent with due process concerns of notice,⁹⁴ Federal Rule of Civil Procedure 65(d) demands that any such injunction “state its terms specifically” and “describe in reasonable detail—and not by referring to the complaint or other document—the act or acts restrained or required.”⁹⁵ Pursuant to this demand, the Federal Circuit has repeatedly “rejected as overly broad . . . permanent injunction[s] that simply prohibit[] future infringement of a patent.”⁹⁶ The circuit has indicated

92. *KSM*, 776 F.2d at 1526. The *KSM* court observed:

The unreasonableness of a decree incorporating a vague or broad prohibition against “infringement” of a “patent” is alleviated because of the universal rule . . . that contempt proceedings, civil or criminal, are available only with respect to devices previously admitted or adjudged to infringe, and to other devices which are no more than colorably different therefrom and which clearly are infringements of the patent.

Id.; cf. 11A CHARLES ALAN WRIGHT, ARTHUR R. MILLER & MARY KAY KANE, FEDERAL PRACTICE AND PROCEDURE § 2955, at 311 (2d ed. 1995) (“A court’s failure to comply with the prerequisites in Rule 65(d) as to the proper scope or form of an injunction or restraining order does not deprive it of jurisdiction or render its order void.” (footnotes omitted)).

93. 35 U.S.C. § 283 (2006).

94. *H.K. Porter Co. v. Nat’l Friction Prods. Corp.*, 568 F.2d 24, 27 (7th Cir. 1978) (“Because of the risks of contempt proceedings . . . interests of liberty and due process make it indispensable for the chancellor or his surrogate to speak clearly, explicitly, and specifically if violation of his direction is to subject a litigant . . . to coercive or penal measures [and] to payment of damages.”).

95. FED. R. CIV. P. 65(d)(1); see also 11A WRIGHT, MILLER & KANE, *supra* note 92, § 2955, at 308–09 (“The drafting standard established by Rule 65(d) is that an ordinary person reading the court’s order should be able to ascertain from the document itself exactly what conduct is proscribed.”).

96. *E.g.*, *Int’l Rectifier Corp. v. IXYS Corp.*, 383 F.3d 1312, 1316 (Fed. Cir. 2004) (vacating an injunction that “by its terms . . . applies to ‘any device’ made or sold by IXYS that is within the scope of the patent claims”); see also *Forest Labs., Inc. v. Ivax Pharm., Inc.*, 501 F.3d 1263, 1272 (Fed. Cir. 2007) (narrowing an injunction by “delet[ing] the language ‘any products that infringe the ’712 patent, including’”); Marketa Trimble, *Cross-Border Injunctions in U.S. Patent Cases and Their Enforcement Abroad*, 13 MARQ. INTELL. PROP. L. REV. 331, 340 (2009) (noting the Federal Circuit’s indication that injunctions broadly prohibiting infringement of a patent violate Rule 65(d)); cf. PETER S. MENELL ET AL., PATENT CASE MANAGEMENT JUDICIAL GUIDE § 3.8.1, at 3-25 (2009) (stating that an injunction “must . . . specifically describe the infringing actions enjoined, with reference to particular products”). An injunction simply stating that further infringement of a specific patent or patent claim is prohibited would appear to require “reference [to] materials in other documents” and thus, on that ground alone, would be at least technically contrary to the plain

that a do-not-infringe injunction generally needs to “limit its prohibition to the manufacture, use, or sale of the specific infringing device, or to infringing devices no more than colorably different from the infringing device.”⁹⁷ In other words, a do-not-infringe injunction generally must have a Type-0, measure-zero or a Type-1, colorable-differences form.

On the other hand, in at least two instances, the Federal Circuit has tolerated violation of the general prohibition against Type-2, obey-the-law injunctions. In 1999, a Federal Circuit panel upheld a permanent injunction forbidding “any further infringement of the ’522 patent.”⁹⁸ The panel essentially found that, under the circumstances, use of Type-2 language constituted only harmless error. According to the panel, “any danger of unwarranted contempt actions [was] minimal, if not completely non-existent, because of the detailed record on which this injunction was entered.”⁹⁹ More than a decade later, in *Streck, Inc. v. Research & Diagnostic Systems, Inc.*,¹⁰⁰ a 2012 panel of the Federal Circuit followed this 1999 decision, using it to justify upholding an injunction that included Type-0 and Type-1 language but that also included Type-2 language forbidding “otherwise infringing the asserted claims [of the patents in suit].”¹⁰¹

At least for the moment, these 1999 and 2012 panel decisions appear to be substantially anomalous. In both 2004 and 2007, the Federal Circuit reaffirmed its more traditional interpretation of Rule 65(d), holding that “the only acts the injunction may prohibit are infringement of the patent by the adjudicated devices and infringement by devices not more than colorably different from the adjudicated devices. In order to comply with Rule 65(d), the injunction should explicitly proscribe only those specific acts.”¹⁰² In 2004, the circuit further suggested that the 1999 decision should be viewed either as correct but limited to a relatively exceptional set of facts, or as incorrect and not controlling “to the extent [it] is inconsistent with the rule [previously] pronounced.”¹⁰³ The Federal Circuit’s 2012 decision upholding

language of Rule 65(d). See *Dupuy v. Samuels*, 465 F.3d 757, 758 (7th Cir. 2006) (arguing for and applying a relatively strict, “literal interpretation” of Rule 65(d)’s requirement “that an injunction be a self-contained document rather than [one] incorporat[ing] by reference materials in other documents”); cf. *H.K. Porter*, 568 F.2d at 27 (“It is beyond cavil that when it merely incorporated by reference the Settlement Agreement, the April 15, 1968 order ignored that rule’s mandatory requirement that an injunction ‘shall describe in reasonable detail, and not by reference to the complaint or other document, the act or acts sought to be restrained.’”). But see *Landmark Legal Found. v. EPA*, 272 F. Supp. 2d 70, 74 (D.D.C. 2003) (“Courts are split on whether Rule 65(d) requires a strict interpretation.”).

97. *Additive Controls & Measurement Sys., Inc. v. Flowdata, Inc.*, 986 F.2d 476, 479–80 (Fed. Cir. 1993).

98. *Signtech USA, Ltd. v. Vutek, Inc.*, 174 F.3d 1352, 1359 (Fed. Cir. 1999).

99. *Id.*

100. 665 F.3d 1269 (Fed. Cir. 2012).

101. *Id.* at 1293 (internal quotation marks omitted).

102. *Int’l Rectifier Corp. v. IXYS Corp.*, 383 F.3d 1312, 1316 (Fed. Cir. 2004); accord *Forest Labs., Inc. v. Ivax Pharm., Inc.*, 501 F.3d 1263, 1271 (Fed. Cir. 2007) (quoting *Int’l Rectifier*).

103. *Int’l Rectifier*, 383 F.3d at 1317.

a Type-2 injunction did not cite, distinguish, or otherwise engage with the court's at least potentially contrary 2004 and 2007 precedents.¹⁰⁴ Even if the 2012 decision had, it would have been powerless to overrule them.¹⁰⁵

Thus, under currently controlling precedent, Type-2, obey-the-law injunctions appear to remain a generally prohibited form of injunctive relief. On the other hand, as appears commonly to be the case with violations of Rule 65(d) in U.S. law,¹⁰⁶ the practical significance of this prohibition is limited because the Federal Circuit has instructed that if a forbidden Type-2 injunction is not challenged on direct appeal, the injunction should not be treated as void in contempt proceedings but should instead be effectively "reformed"¹⁰⁷ by being understood to be restricted to the scope of a Type-1, colorable-differences order.¹⁰⁸

B. Purely Reparative Injunctions

Type-2, obey-the-law injunctions are not the only kind of patent-infringement orders that existing law forbids. The Federal Circuit has held that district courts lack authority to issue purely reparative injunctions¹⁰⁹ that appear to be directly concerned only with correcting for harm caused by past infringement.

An example of a purely reparative order would be an order that Energizer destroy all Schick Quattros manufactured in the United States that are now in Argentina, even though the court knows that the Argentinian Quattros will never make their way back to the United States.¹¹⁰ Such an order works to limit the harm to Gillette—as well as the gain to Energizer—

104. *Streck*, 665 F.3d at 1293 (citing only *Signtech* and Federal Rule of Civil Procedure 65(d) in reasoning about the propriety of Type-2 language).

105. See *Newell Cos. v. Kenney Mfg. Co.*, 864 F.2d 757, 765 (Fed. Cir. 1988) ("Where there is direct conflict [between prior decisions of Federal Circuit panels], the precedential decision is the first.").

106. See 11A WRIGHT, MILLER & KANE, *supra* note 92, § 2955, at 311 ("A court's failure to comply with the prerequisites in Rule 65(d) as to the proper scope or form of an injunction or restraining order does not deprive it of jurisdiction or render its order void." (footnotes omitted)).

107. BLACK'S LAW DICTIONARY 1394 (9th ed. 2009) (defining *reformation* as "[a]n equitable remedy by which a court will modify a written agreement to reflect the actual intent of the parties").

108. See *KSM Fastening Sys., Inc. v. H.A. Jones Co.*, 776 F.2d 1522, 1526 (Fed. Cir. 1985) (noting that, even when a decree "incorporat[es] a vague or broad prohibition against 'infringement' of a 'patent[.]' . . . contempt proceedings, civil or criminal, are available only with respect to devices previously admitted or adjudged to infringe, and to other devices which are no more than colorably different therefrom and which clearly are infringements of the patent"), *overruled on other grounds*, *TiVo Inc. v. EchoStar Corp.*, 646 F.3d 869 (Fed. Cir. 2011) (en banc).

109. See Tracy A. Thomas, *The Continued Vitality of Prophylactic Relief*, 27 REV. LITIG. 99, 102 (2007) ("Reparative injunctions repair the ongoing consequences of the past harm, and might order the reinstatement of an employee fired because of discrimination.").

110. See *Spine Solutions, Inc. v. Medtronic Sofamor Danek USA, Inc.*, 620 F.3d 1305, 1320 (Fed. Cir. 2010) (holding that because overseas sales "cannot infringe any U.S. patent, and there is little risk that the infringing devices will be imported," the district court "abused its discretion in imposing . . . extraterritorial restraints").

from Energizer's past infringing manufacture. But the order does nothing directly to help prevent infringement.

Although reparative injunctions are commonly available in other areas of U.S. law,¹¹¹ the Federal Circuit has held that they are not available under the U.S. Patent Act. The basis for this holding is § 283 of the Act, which states, "The several courts having jurisdiction of cases under this title may grant injunctions in accordance with the principles of equity to prevent the violation of any right secured by patent, on such terms as the court deems reasonable."¹¹²

The Federal Circuit has held that § 283's authorization of injunctions to "prevent the violation" of patent rights establishes an exclusive purpose for the permitted exercise of injunctive power. In the circuit's words, "An injunction is only proper to prevent future infringement of a patent, not to remedy past infringement."¹¹³ Although the Federal Circuit's conclusion that § 283 forbids purely reparative injunctions might be contestable,¹¹⁴ this Article will leave to another day potential debate over this issue. For present purposes, what matters is that purely reparative injunctions are—like Type-2, obey-the-law injunctions—generally forbidden under existing law. District courts' rates of compliance with the ban on purely reparative injunctions can therefore provide a reference point for study of district courts' rates of compliance with the ban on Type-2 injunctions.

111. 1 DOBBS, *supra* note 37, § 2.9, at 225 ("The reparative injunction goes when the evidence shows that an existing right has been violated but can be repaired or restored effectively.")

112. 35 U.S.C. § 283 (2006).

113. *Spine Solutions*, 620 F.3d at 1320; *see also* *Johns Hopkins Univ. v. CellPro, Inc.*, 152 F.3d 1342, 1365 (Fed. Cir. 1998) ("In accordance with the clear wording of [§ 283], an injunction is only proper to the extent it is to prevent the violation of any right secured by patent." (internal quotation marks omitted)).

114. One might argue that the prospect of a reparative order can have a deterrent effect that can help prevent infringement in general, even if it is too late to prevent infringement in the case at hand. Further, one might note that the case to which the Federal Circuit's ban on reparative orders tends to be traced involved a situation in which matter abroad had been produced in the U.S. prior to issuance of the relevant patent and thus was never directly involved in infringement. *See Johns Hopkins*, 152 F.3d at 1366 ("An injunction requiring return of [an] exported machine, which was never made, used, or sold during the term of the patent in the United States, is beyond the scope of Section 283 and hence an abuse of discretion."). Only later did the Federal Circuit apply language from that case to justify forbidding destruction of matter abroad that was directly involved in an infringement of an issued U.S. patent. *See Spine Solutions*, 620 F.3d at 1320 ("[T]he extraterritorial portion of the injunction appears to be premised solely on Medtronic's past infringement, not on the prevention of future infringement."); *see also* Non-confidential Brief for Defendants–Appellants Medtronic Sofamor Danek USA, Inc. & Medtronic Sofamor Danek, Inc. at 64, *Spine Solutions*, 620 F.3d 1305 (No. 2009-1538) ("The district court stated that . . . it is quite possible that some of Medtronic's exported devices were manufactured in violation of [the relevant] patent." (internal quotation marks omitted)).

C. Specially Tailored Injunctions

1. *Characteristics of Specially Tailored Injunctions.*—The en banc case of *TiVo Inc. v. EchoStar Corp.* featured an additional type of patent-infringement injunction: the specially tailored injunction. Although directly concerned with preventing infringement,¹¹⁵ such injunctions either (a) explicitly feature prohibitions or requirements that extend beyond what is formally necessary to prevent infringement, (b) define their bounds without reference either to patent rights or to matter already adjudged to infringe, or (c) include an explicit carve out for infringing (or likely infringing) behavior. A hypothetical injunction of subtype (a) might require the destruction of all Schick Quattro products in Energizer's possession.¹¹⁶ This injunction is formally extraproductive of Gillette's patent rights because there might be ways for Energizer to modify or otherwise to continue possessing an already manufactured Quattro without infringing Gillette's patent for a three-blade razor. Likewise, a subtype-(b) injunction that forbids Energizer Holdings from further activity involving multiblade razors would be formally extraproductive in that it would forbid activity involving two-blade razors that do not seem even arguably covered by Gillette's patent on razors having at least three (and perhaps exactly three) blades.¹¹⁷

The injunction in *TiVo* was a "partial disablement" variant of the hypothetical subtype-(a) order requiring destruction of Schick Quattros. In *TiVo*, the district court issued an order for injunctive relief requiring that the adjudged infringer "disable the DVR functionality (i.e., disable all storage to and playback from a hard disk drive of television data) in all but 192,708 units of the Infringing Products."¹¹⁸ Like a destruction order, this disablement order was extraproductive to the extent it forbade implementation of a redesign that might have rendered a product noninfringing while maintaining the specified functionality.

Notably, however, a specially tailored injunction, at least as defined here, need not be extraproductive in the manner of a destruction order, which

115. The distinction between a purely reparative injunction and a prophylactic injunction can be subtle and, arguably, excessively formal. Many prophylactic injunctions—such as a hypothetical order to destroy all Schick Quattros in Energizer's possession in the United States—can be viewed as at least partly reparative. Destruction prevents further infringement with the destroyed Quattros but also helps nullify the effects of past infringing manufacture and is thereby reparative. For purposes of legal characterization in relation to existing law, however, the key point appears to be that, regardless of any additional reparative effect or purpose, a prophylactic injunction has a direct connection to the statutorily sanctioned goal of preventing future infringement.

116. See, e.g., *Proveris Scientific Corp. v. InnovaSystems Inc.*, No. 05-12424-WGY, slip op. at 3 (D. Mass. May 11, 2007) (requiring the defendant to "destroy all inventory of its OSA product").

117. Cf. *Gillette Co. v. Energizer Holdings, Inc.*, 405 F.3d 1367, 1374 (Fed. Cir. 2005) ("Based on the preliminary record before this court, the district court erred in limiting the claims of [Gillette's] patent to encompass safety razors with solely three blades").

118. *TiVo Inc. v. EchoStar Corp.*, 646 F.3d 869, 877 (Fed. Cir. 2011) (en banc) (quoting *TiVo Inc. v. Dish Network Corp.*, 640 F. Supp. 2d 853, 858 (E.D. Tex. 2009)).

might be viewed as a form of prophylactic relief.¹¹⁹ A specially tailored injunction can provide less protection, at least facially, than a conventional do-not-infringe injunction. A prophylactic injunction might, for example, fail to encompass all merely colorable variants of subject matter already adjudged to infringe. More particularly, whereas a Type-1, colorable-differences injunction might forbid use of “no more than colorable” variants of an adjudged-to-infringe purification process running at a pH of 5.0, a specially tailored injunction might only forbid use of purification processes running at a pH of between 4.6 and 5.4. The latter injunction might forbid a smaller range of activity than the former, colorable-differences injunction because a process running at, say, a pH of 4.5 would be outside the scope of the specially tailored injunction but might be a “no more than colorable” variant of the original infringing process.¹²⁰ If one accepts that a conventional, colorable-differences injunction provides a proper point of reference, one might then characterize the specially tailored injunction as subprotective, rather than extraprotective, of patentee interests.

Why might courts—or the parties who commonly draft injunction orders as a matter of actual fact—bother themselves with the trouble of crafting specially tailored relief? Why might a court issue extraprotective orders in some circumstances and subprotective orders in others? Recall that in issuing an injunction, a court needs to consider not only patentee interests but also legitimate interests of the adjudged infringer and society at large. A specially tailored injunction can be prophylactically extraprotective of patentee interests or, alternatively, subprotective of those interests in a way that might be viewed as prophylactically protective of infringer or general social interests.¹²¹ Moreover, a specially tailored injunction might simply look to strike a reasonable balance between patentee and infringer interests in

119. In the constitutional law context, the term *prophylactic rule* has inspired “a wealth of sometimes widely divergent definitions.” Mitchell N. Berman, *Constitutional Decision Rules*, 90 VA. L. REV. 1, 30 (2004). Sometimes commentators view only extraprotective rules as prophylactic rules. See, e.g., *id.* (describing a prophylactic rule as “that sort of extraconstitutional rule that overenforces what the Constitution, as judicially interpreted, would itself require”); *id.* at 40–42 (distinguishing prophylactic rules from “underenforcement rule[s]” and hybrid “overlapping rule[s]”). At least one prior commentator has characterized both subprotective and extraprotective legal rules as prophylactic, using reasoning like that presented in this Article’s text. David A. Strauss, *The Ubiquity of Prophylactic Rules*, 55 U. CHI. L. REV. 190, 204–06 (1988) (characterizing both rules of “strict scrutiny” and “rational basis review” as prophylactic rules, although the former is likely extraprotective and the latter is likely subprotective of the constitutional interests most centrally at issue).

120. The example is inspired by the fact pattern associated with *Warner-Jenkinson Co. v. Hilton Davis Chemical Co.*, 520 U.S. 17 (1997), in which the U.S. Supreme Court considered whether a purification process running at a pH of 5.0 could infringe a patent claim under the doctrine of equivalents. See *id.* at 23 (describing the case’s underlying facts).

121. A subprotective injunction from a patentee’s perspective is likely to be an extraprotective injunction from the perspective of an infringer or society. Such an injunction might, for example, provide prophylactic protection to legitimate infringer interests in pursuing a good-faith redesign without fear of being held in contempt.

a way that is likely to be more administrable and to hold greater promise for compliance than a conventional do-not-infringe order.

This last point bears emphasis. Regardless of whether a specially tailored injunction is extraprotective or subprotective of patentee interests, such an injunction can, when well crafted, substantially ease the tasks of determining an injunction's bounds and identifying violations. Because of improvements in notice and enforceability, substitution of such injunctions for conventional do-not-infringe orders can both increase compliance and reduce chilling of socially desirable redesigns. In short, even to the extent a specially tailored injunction's relative subprotection or extraprotection of underlying rights is undesirable when considered in isolation, that suboptimality might be more than "paid for" by improvements to notice, enforceability, and administrability.

2. *Legal Status of Specially Tailored Injunctions.*—Are specially tailored injunctions legitimate under U.S. patent law? Given the Federal Circuit's views on Type-2, obey-the-law injunctions and purely reparative injunctions, one might worry about the legal legitimacy of specially tailored injunctions, particularly when they are prophylactic in the sense that they likely prohibit some noninfringing activity or require other activity that is beyond what is necessary to avoid infringement. Indeed, in declaring Type-2 injunctions to be illegitimate, the Federal Circuit sometimes has used language that facially suggests that only Type-0, measure-zero or Type-1, colorable-differences injunctions are proper.¹²²

A first point in response is that prophylactic injunctions can be drawn in a way that addresses both of the Federal Circuit's grounds for rejecting Type-2, obey-the-law injunctions and purely reparative injunctions. Prophylactic injunctions can be written in a way that provides the adequate notice required by Federal Rule of Civil Procedure 65(d)¹²³ and directly operates to prevent future infringement.¹²⁴

Moreover, prophylactic injunctions—and specially tailored injunctions more generally—have a substantial foundation in the traditional equity power that § 283 affirms.¹²⁵ Courts' use of extraprotective injunctions in a variety of areas of law reflects recognition that "sometimes the chancellor can assure plaintiffs their rights only by giving them more than they are entitled to."¹²⁶

122. See *supra* text accompanying note 102.

123. See *supra* text accompanying notes 94–97.

124. See *supra* text accompanying notes 111–13.

125. See *supra* text accompanying note 112.

126. 1 DOBBS, *supra* note 37, § 2.4(7), at 121; see also *Russian Media Grp., LLC v. Cable Am., Inc.*, 598 F.3d 302, 307 (7th Cir. 2010) ("The district court may even enjoin certain otherwise lawful conduct when the defendant's conduct has demonstrated that prohibiting only unlawful conduct would not effectively protect the plaintiff's rights against future encroachment."); cf. *Taco Cabana Int'l, Inc. v. Two Pesos, Inc.*, 932 F.2d 1113, 1126 (5th Cir. 1991) ("In fashioning relief against a party who has transgressed the governing legal standards, a court of equity is free to

As Tracy Thomas has argued, a prophylactic injunction can “develop[] almost instinctively” as an alternative to remedies that amount to no more than “empty commands simply to stop [certain] behavior.”¹²⁷ High-profile support for the notion that specially tailored and at least partially prophylactic injunctions can be acceptable forms of relief has come through anti-abortion protest cases in which the U.S. Supreme Court upheld injunctions requiring that protestors keep a certain distance from clinic entrances.¹²⁸ Such injunctions illustrate the capacity of specially tailored injunctions to balance competing interests—rights to free speech, rights “to seek lawful medical or counseling services,” and concerns with public safety, order, property rights, and privacy.¹²⁹ These injunctions also illustrate the potential use of special tailoring to generate relatively clear lines that can help private parties and public officials know what constitutes compliance and when noncompliance has occurred.¹³⁰

Indeed, prophylactic injunctions might be viewed as a subset of a larger family of prophylactic measures that U.S. law frequently uses to implement and enforce legal norms. At least since 1988, various scholars have contended that “[p]rophylactic” rules are, in an important sense, the norm, not the exception.¹³¹ Whether in the context of requiring *Miranda v. Arizona*¹³²

proscribe activities that, standing alone, would have been unassailable.” (quoting *Ky. Fried Chicken Corp. v. Diversified Packaging Corp.*, 549 F.2d 368, 390 (5th Cir. 1977)); RENDLEMAN, *supra* note 79, at 427 (“Under some circumstances, a judge may grant a plaintiff an injunction that forbids defendant’s activities that are not themselves wrongs and that commands activities that are not in themselves part of the plaintiff’s substantive-law entitlement.”); 11A WRIGHT, MILLER & KANE, *supra* note 92, § 2955, at 327–28 (indicating that a “broad decree” might be justified as “the only way to prevent a statutory violation” or because “it can be drafted by the court more easily than a narrow decree”).

127. Thomas, *supra* note 109, at 104.

128. See *Schenck v. Pro-Choice Network of W. N.Y.*, 519 U.S. 357, 380 (1997) (upholding “fixed buffer zones around the doorways, driveways, and driveway entrances” of clinics); *Madsen v. Women’s Health Ctr., Inc.*, 512 U.S. 753, 776 (1994) (upholding “noise restrictions and the 36-foot buffer zone around the clinic entrances and driveway because they burden no more speech than necessary to eliminate the unlawful conduct targeted by the state court’s injunction”).

129. *Madsen*, 512 U.S. at 767–68.

130. Cf. LAYCOCK, *supra* note 78, at 290 (“Conflict and misunderstanding, testing of limits and overreaching, emotional reactions, inconsistent perceptions and accounts of what happened—all these are inevitable. The judicial need for bright-line rules can be overwhelming.”).

131. Strauss, *supra* note 119, at 195; see also Berman, *supra* note 119, at 13–14 (observing that, by the year 2000, “many constitutional theorists had become persuaded by David Strauss’s careful and powerful argument that prophylactic rules indistinguishable from *Miranda* are ubiquitous and legitimate”); cf. RICHARD H. FALLON, JR., IMPLEMENTING THE CONSTITUTION 5 (2001) (“Rather than picturing the Justices as pervasively engaged in a search for the Constitution’s one true meaning, I argue . . . that we should understand the Supreme Court’s role as a more multifaceted one of ‘implementing’ constitutional norms. . . . [L]awyers’ work involv[es] . . . the creative design of implementing strategies.”); Lawrence Gene Sager, *Fair Measure: The Legal Status of Underenforced Constitutional Norms*, 91 HARV. L. REV. 1212, 1213 (1978) (“[T]here is an important distinction between a statement which describes an ideal which is embodied in the Constitution and a statement which attempts to translate such an ideal into a workable standard for the decision of concrete issues.”).

132. 384 U.S. 436 (1966).

protection against self-incrimination,¹³³ strictly limiting content-based restrictions on speech,¹³⁴ or applying strict scrutiny to suspect forms of legal classification,¹³⁵ courts frequently enforce legal norms through prophylactic rules that, by avoiding the need for fully individualized assessment of rights' exact bounds, help limit uncertainty and improve compliance, enforceability, and administrability.¹³⁶ Consistent with the sense that prophylactic rules can render difficult legal problems more manageable, prophylactic injunctions often issue in contentious public law cases "involving schools, prisons, [or] sexual harassment."¹³⁷ In such cases, courts commonly require reporting or monitoring, new institutional policies and procedures, or personnel training to try to transform violation-fostering cultures.¹³⁸

But prophylactic injunctions also issue in cases that are more narrowly focused on commercial interests. In a leading casebook, Douglas Laycock illustrates such relief through a case in which a court protected "PepsiCo trade secrets and confidential information" by ordering a former PepsiCo employee to delay starting work for a competitor.¹³⁹ This decree thus prohibited activity beyond the disclosure or use of confidential information that the law formally forbade.

Antitrust law is an area where prophylactic injunctions are particularly well established. A structural injunction at the conclusion of an antitrust case can require the breakup of an offending firm as a means to protect against future monopolization.¹⁴⁰ Less dramatic antitrust decrees can also have

133. See, e.g., Strauss, *supra* note 119, at 190 (discussing the "'prophylactic' character [of] the *Miranda* rule").

134. See, e.g., *id.* at 198 (discussing how significant aspects of First Amendment doctrine, such as "a nearly conclusive presumption against [the] constitutionality" of most "content-based" restrictions on speech, might be viewed as prophylactic).

135. See, e.g., *id.* at 204–05 (discussing how equal protection doctrine might be viewed as embodying prophylactic rules).

136. Cf. *id.* at 200 (describing the courts' "categorical approach to content-based [speech] restrictions and the *Miranda* rules [as] relatively rigid doctrines designed to reduce the likelihood that the authorities . . . will violate the law, and designed to improve a reviewing court's chances of identifying violations where they occur").

137. See Thomas, *supra* note 109, at 99, 100 (describing prophylactic injunctions as tending to "reach[] the facilitators of harm in order to prevent continued illegality"); cf. RENDLEMAN, *supra* note 79, at 34 ("We will encounter many injunctions that forbid defendants' preparatory, ancillary, and related behavior."); Rendleman, *supra* note 5, at 89 ("Public-nuisance injunctions against street gangs often extend defendants' prohibitions beyond the criminal law . . .").

138. Thomas, *supra* note 109, at 101–02 (describing potential types of prophylactic measures).

139. LAYCOCK, *supra* note 78, at 284–87 (reproducing portions of *PepsiCo, Inc. v. Redmond*, 54 F.3d 1262 (7th Cir. 1995)).

140. HERBERT HOVENKAMP, *THE ANTITRUST ENTERPRISE: PRINCIPLE AND EXECUTION* 300 (2005) ("Early in the history of antitrust enforcement courts tended to favor 'structural' remedies in cases involving significant § 2 violations. A structural remedy . . . typically breaks the defendant firm into two or more pieces . . ."); cf. Howard A. Shelanski & J. Gregory Sidak, *Antitrust Divestiture in Network Industries*, 68 U. CHI. L. REV. 1, 15–16 (2001) (describing "structural remedies" as "redistributing competitive assets" either "by breaking the defendant company into two or more pieces" or "by requiring the defendant to sell or otherwise make available to its competitors some input, right, or facility").

prophylactic aspects. In *United States v. Microsoft Corp.*,¹⁴¹ the district court acknowledged that its consent decree included a requirement that “plainly exceed[ed] the scope of [Microsoft’s] liability”¹⁴²—namely, the requirement that “Microsoft license . . . any communications protocol installed on a Windows client which is used to interoperate or communicate with a Microsoft server operating system product without the addition of software code to the client.”¹⁴³ The court reasoned that this requirement was justified because it was “closely connected with the theory of liability . . . and further[ed] efforts to ensure that there remain[ed] no practices likely to result in monopolization in the future.”¹⁴⁴

Such examples of prophylactic injunctions from a variety of legal areas support Tracy Thomas’s notion that prophylactic injunctions are particularly likely to be warranted when two conditions apply: (1) the underlying principles of substantive law are difficult to enforce or articulate with precision,¹⁴⁵ and (2) otherwise lawful conduct appears likely to facilitate, accompany, or be difficult to distinguish from an oft-associated offense.¹⁴⁶

At least as of this writing, the Federal Circuit has not insisted on a contrary view. In *TiVo*, seven judges of the circuit’s en banc majority determined that challenges to a partial-disablement injunction as overbroad had been waived through the defendants’ failure to make them on direct appeal.¹⁴⁷ In a footnote, the majority added a statement emphasizing that such a challenge would not necessarily succeed:

We note . . . that, although we have strongly discouraged judicial restraint of noninfringing activities, we have never barred it outright and instead have repeatedly stated that district courts are in the best position to fashion an injunction tailored to prevent or remedy

141. 231 F. Supp. 2d 144 (D.D.C. 2002), *aff’d in relevant part sub nom.* Massachusetts v. Microsoft Corp., 373 F.3d 1199 (D.C. Cir. 2004).

142. *Id.* at 190.

143. *Id.* at 189.

144. *Id.* at 190 (internal quotation marks omitted). The district court’s reasoning substantially parroted the D.C. Circuit’s earlier instruction that insurance against future monopolization was one of the necessary goals of an antitrust decree. *United States v. Microsoft Corp.*, 253 F.3d 34, 103 (D.C. Cir. 2001) (“[A] remedies decree in an antitrust case must seek to . . . ensure that there remain no practices likely to result in monopolization in the future.” (citations omitted) (internal quotation marks omitted)).

145. See Tracy A. Thomas, *The Prophylactic Remedy: Normative Principles and Definitional Parameters of Broad Injunctive Relief*, 52 BUFF. L. REV. 301, 372 (2004) (“The intangible rights at issue in the prophylactic remedies cases present challenges to the court as to how to translate those rights into tangible meaning.”).

146. See *id.* at 334 (“[T]he affiliated conduct included in the prophylactic relief must demonstrate a sufficient causal nexus to the established harm.”); *id.* at 339–40 (“The causal nexus is established where the affiliated conduct bears a factual relationship to the harm and the relationship is of sufficiently close degree to justify the inclusion of the conduct in the prophylactic order.”).

147. *TiVo Inc. v. EchoStar Corp.*, 646 F.3d 869, 890 (Fed. Cir. 2011) (en banc) (“We therefore conclude that EchoStar’s arguments on overbreadth of the district court’s injunction have been waived by its failure to raise them earlier.”).

infringement. Because it is not before us in this case, we make no *en banc* holding on that issue.¹⁴⁸

What will be prophylactic injunctions' fate if the Federal Circuit squarely faces the issue? It might be a close call. The *en banc* majority's footnote suggests that the majority remained open to such relief but was at least somewhat suspicious of it. Moreover, five judges dissented from the portion of the majority opinion that contained this footnote. Their dissenting opinion used language that might be understood to indicate that the dissenters were significantly less open to the use of prophylactic injunctions:

[T]his court has recognized that an injunction is only proper to the extent it is to prevent the violation of any right secured by patent. . . . [In a prior case], we held that an injunction which precludes [the defendant] from activities that are not necessary to prevent infringement of the patented process cannot stand.¹⁴⁹

Notably, however, Federal Circuit precedent on injunctive relief is more nuanced than the dissenters' language might suggest. In the case usually cited as the source of the rule that the U.S. Patent Act does not authorize purely reparative injunctions, the circuit panel stated that an injunction to "prevent infringement of a United States patent" "can reach extraterritorial activities . . . , even if these activities do not themselves constitute infringement."¹⁵⁰ Moreover, the circuit has repeatedly upheld prophylactic injunctions in situations where an adjudged patent infringer violated a prior court order.¹⁵¹ For example, in *Additive Controls & Measurement Systems, Inc. v. Flowdata, Inc.*,¹⁵² the Federal Circuit upheld a broad injunction forbidding a contemnor "from undertaking any activities with respect to positive displacement flowmeters without first obtaining leave of court."¹⁵³

148. *Id.* at 890 n.9 (citations omitted). Does the *en banc* majority's indication that an injunction can be designed "to prevent or *remedy* infringement" mean that the judges in the majority are looking to step back from the Federal Circuit's prior rejection of purely reparative patent-infringement injunctions? Not necessarily. The majority might simply have meant to recognize that some injunctions, such as an order to destroy specified devices, can serve dual purposes of preventing further infringement and helping to correct for past infringement.

149. *Id.* at 893-94 (Dyk, J., dissenting) (fourth alteration in original) (citations omitted) (internal quotation marks omitted).

150. *Johns Hopkins Univ. v. CellPro, Inc.*, 152 F.3d 1342, 1366-67 (Fed. Cir. 1998) (emphasis omitted); *cf.* Trimble, *supra* note 96, at 367 ("U.S. courts . . . have issued orders requesting or prohibiting behavior abroad that is not infringing per se but is behavior that the courts have decided to target in order to prevent further infringements of U.S. patents.").

151. *See, e.g., Spindelfabrik Suessen-Schurr v. Schubert & Salzer Maschinenfabrik Aktiengesellschaft*, 903 F.2d 1568, 1577 (Fed. Cir. 1990) (holding that "repeated and 'flagrant' violations of the district court's earlier injunction fully justified these broad provisions" against "directly or indirectly engaging in any activity which in any way relates to the manufacture, sale, use, servicing, exhibition, demonstration, promotion or commercialization of any automated rotor spinning machines"); *see also* KIMBERLY A. MOORE ET AL., *PATENT LITIGATION AND STRATEGY* 715 (3d ed. 2008) (discussing *Spindelfabrik*).

152. 154 F.3d 1345 (Fed. Cir. 1998).

153. *Id.* at 1356.

The Federal Circuit held that this ban from an entire sphere of commercial activity reflected a “reasonabl[e] conclu[sion] that such measures were necessary . . . to compel compliance with the court’s orders.”¹⁵⁴ Although the Federal Circuit has also indicated that *broad* prophylactic orders like that in *Additive Controls* “should be used only in exceptional cases,”¹⁵⁵ this indication itself appears to leave open the possibility that comparatively *narrow* prophylactic orders can be proper in a wider range of cases. There thus appears much for litigants and judges to work out in future battles over the proper bounds of both prophylactic relief and the larger category of specially tailored relief of which prophylactic relief forms a part.

IV. District Court Practice in Issuing Patent-Infringement Injunctions

Part III described five categories of injunctions. To what extent do these injunction types appear in practice? To answer this question, I used the Lex Machina database to search for patent-infringement injunctions issued by U.S. district courts in 2010. Through systematic review of injunction-related orders, I identified 143 patent-infringement injunctions.¹⁵⁶ I obtained the text of the 143 orders via Lex Machina or PACER.¹⁵⁷

The 143 orders were coded for various characteristics, including (a) whether the order was a permanent injunction, preliminary injunction, or temporary restraining order; (b) whether, as part of a consent agreement, the parties agreed to the injunction prior to its issue—i.e., whether, in this Article’s terminology, the injunction was “consented to”; (c) whether a patent that formed part of the basis for the order was a utility patent as opposed to a design or plant patent; (d) whether patent rights at issue focused

154. *Id.*

155. *Id.*

156. Some additional search results were apparently added to the Lex Machina database after the author’s review of search results for injunctions issued in 2010 began. Thus, the results reported in this Article might not reflect all 2010 injunctions that are ultimately indicated in the Lex Machina database.

157. Four orders for injunctive relief were excluded from the ultimate dataset because their text leaves unclear whether they were motivated by concerns with patent infringement, as opposed to infringement of other rights such as trademark or trade dress. *Bon-Aire Indus., Inc. v. Mitchell Prods.*, No. 3:10-cv-01602-MLC-TJB, slip op. at 2 (D.N.J. Apr. 26, 2010) (permanently enjoining various acts involving “any hose nozzle having a trade dress that is identical to, substantially similar to, or a colorable imitation of the appearance of the ULTIMATE® hose nozzle”); *Bon-Aire Indus., Inc. v. Mitchell Prods.*, No. 3:10-cv-01602-MLC-TJB, slip op. at 2–3 (D.N.J. Apr. 13, 2010) (preliminarily enjoining, in the same case, various acts involving such a nozzle); *ICON Health & Fitness, Inc. v. Solo Sports Grp., Inc.*, No. 1:10-cv-00020-TC, slip op. at 1 (D. Utah Mar. 4, 2010) (permanently enjoining activities involving “the Elite Fitness Dual Action Upright Exercise Bike Model EB275, and any other product that incorporates the same or substantially the same features of ICON’s trade dress design contained in its Weslo Pursuit E28 bike”); *Metraflex Co. v. Flex-Hose Co.*, No. 1:10-cv-00302, slip op. at 1–2 (N.D. Ill. Feb. 16, 2010) (preliminarily enjoining “utilizing, displaying, or distributing [the] Seismic Movement Brochure,” statements about whether certain products “meet building code requirements or specifications” and reproduction of “drawings and/or photographs of U-shaped or V-shaped flexible loops that are original to and/or the copyrighted property of Metraflex”).

on a biomedical substance (BMS) or another form of subject matter (non-BMS); and (e) whether at least a portion of the order explicitly incorporated Type-0, Type-1, Type-2, purely reparative, or specially tailored language, or any combination of such language. To help ensure consistency of coding, I personally coded all results.¹⁵⁸

Arguably, there is double counting in the 143-order dataset because some of the orders were issued in the same case and even on the same day. In a single case, *Reah v. Re.source, Inc.*, a district court issued nine different permanent injunctions directed at nine different defendants in a little over two months.¹⁵⁹ For purposes of this Article, I have counted separately such same-case orders because, although many of the orders use substantially identical language, this is not true of all of them.¹⁶⁰ Observed differences suggest that parties or judges might be properly understood to have given separate consideration to each individual order's scope. In any event, same-case orders from a total of ten different cases accounted for only twenty-eight orders total. The general impressionistic significance of the results from the 143-order dataset appears unlikely to be dominated by how this counting

158. Coding appears to have generally been straightforward, but it did sometimes require review of underlying patents, briefing, or other documents. Further, characterization of injunction language as Type-1 or as a specially tailored injunction targeting correlated activity could involve some reasoning about the meaning of the injunction's text and its relation to the statutory delineation of infringing activities in 35 U.S.C. § 271. For example, if an injunction forbade "manufacturing," rather than the "making" specifically defined as infringing by § 271, I reasoned that this difference in specific language did not amount to a classification-relevant difference in scope that might result in the injunction being classified as a specially tailored injunction, rather than a do-not-infringe injunction. Likewise, I characterized various, occasional forms of injunctive language explicitly targeting "similar" or "substantially equivalent" products or processes as Type-1 even though this language did not appear precisely in more typical colorable-differences form. *Cf.*, e.g., *Extreme Tool & Eng'g, Inc. v. Bear Cub Enters., LLC*, No. 2:08-cv-257, slip op. at 2 (W.D. Mich. Mar. 29, 2010) (including "similar products that lack colorable distinctions" within the injunction's scope); *Canon Inc. v. Densigraphix Kopi Inc.*, No. 1:10-cv-00034-CMH-IDD, slip op. at 2-3 (E.D. Va. Mar. 9, 2010) (including "substantially equivalent" products within the injunction's scope); *Mannatech, Inc. v. Techmedica Health, Inc.*, No. 3:06-CV-00813-P, slip op. at 3 (N.D. Tex. Jan. 12, 2010) (including "colorable imitations" and products "substantially equivalent in composition" within the injunction's scope). These approaches to classification were consistent with my general effort to be relatively conservative in characterizing injunctive language as Type-2 or specially tailored, rather than falling within the more generally acknowledged categories of Type-0 or Type-1 do-not-infringe orders.

159. *Reah v. Re.source, Inc.*, No. 2:09-cv-00601-CW-DN (D. Utah Mar. 25, 2010) (issuing three separate consent orders directed at three different defendants); *Reah v. Re.source, Inc.*, No. 2:09-cv-00601-CW-DN, slip op. at 1-2 (D. Utah Feb. 23, 2010) (issuing a consent order directed at Datavision Computer Video, Inc.); *Reah v. Re.source, Inc.*, No. 2:09-cv-00601-CW-DN (D. Utah Jan. 20, 2010) (issuing five separate consent orders directed at five different defendants).

160. *Compare*, e.g., *Reah v. Re.source, Inc.*, No. 2:09-cv-00601-CW-DN, slip op. at 10-11 (D. Utah Mar. 25, 2010) (prohibiting LBM Corp. from "making, using, selling, offering for sale, or importing products that come within one or more claims of U.S. Patent No. 6,982,542, or otherwise infringing . . . U.S. Patent No. 6,982,542, including without limitation the Power Station and Power Station Traveller"), with *Denmel Holdings, LLC v. Re.source, Inc.*, No. 2:09-cv-00601-CW-DN (D. Utah Mar. 25, 2010) (prohibiting Electronicsshowplace.com from "making, using, offering for sale, and/or importing charging valets and/or charging stations that come within one or more claims of U.S. Patent No. 6,982,542, or otherwise infringing[ing] . . . U.S. Patent No. 6,982,542").

problem is resolved, particularly as various aspects of these same-case results seem entirely in line with those of the dataset as a whole.¹⁶¹

A. Systematic Violation of Federal Rule of Civil Procedure 65(d)

The most striking empirical result is that a substantial majority of the 143 orders appear to violate the Federal Circuit's understanding of Federal Rule of Civil Procedure 65(d).¹⁶² Eighty-two injunctions, about 57% of the total, contain Type-2 language. This figure is striking, if not shocking. It has long been suspected that despite Rule 65(d), obey-the-law injunctions are relatively common in patent law¹⁶³ and elsewhere.¹⁶⁴ But I do not know of any prior indication that courts commit apparent Rule 65(d) error in the majority of such a significant subset of cases.¹⁶⁵

1. *Consented-To, Uncontested, and Actively Opposed Orders.*—The prevalence of apparent Rule 65(d) error—hereinafter commonly described simply as “error” or “Type-2 error”—does not merely reflect a large number of consent judgments. One might posit that although courts are supposed to be attentive to the proper limits of injunctions even when issuing consent decrees,¹⁶⁶ courts might be less rigorous in enforcing Rule 65(d) when parties have agreed on the form of relief or when, for other reasons, the propriety of

161. The overall Type-2 error rate for the 28 same-case injunctions is about 54% (15 of 28). The Type-2 error rate for the 25 consented-to orders among these 28 injunctions is about 56% (14 of 25). These percentages appear substantially consistent with the approximately 58% Type-2 error rate for the other 115 non-same-case orders (67 of 115) and the approximately 57% Type-2 error rate for the 58 non-same-case consented-to orders (33 of 58). When the results for the 28 same-case injunctions and the 115 non-same-case injunctions are compared, two-sample, two-tailed *t*-tests assuming equal variances and assuming unequal variances yield, respectively, $t(141) = -0.45$, $p = 0.66$; and $t(40) = -0.44$, $p = 0.66$. An *F*-test for equality of variances did not indicate a statistically significant difference between variances ($F = 1.1$, $p = 0.41$). Likewise, when the results for the 25 same-case consented-to orders and the 58 non-same-case consented-to orders are compared, two-sample, two-tailed *t*-tests assuming equal variances and assuming unequal variances yield, respectively, $t(81) = -0.075$, $p = 0.94$; and $t(45) = -0.074$, $p = 0.94$. An *F*-test for equality of variances did not indicate a statistically significant difference between variances ($F = 1.0$, $p = 0.45$).

162. See *supra* text accompanying notes 92–108.

163. See *KSM Fastening Sys., Inc. v. H.A. Jones Co.*, 776 F.2d 1522, 1526 (Fed. Cir. 1985) (“[I]njunctions are frequently drafted or approved by the courts in general terms, broadly enjoining ‘further infringement’ of the ‘patent,’ despite the language of Rule 65(d) . . .”), *overruled on other grounds*, *TiVo Inc. v. EchoStar Corp.*, 646 F.3d 869 (Fed. Cir. 2011) (en banc).

164. See LAYCOCK, *supra* note 78, at 274 (“Defendants do not object to obey-the-law clauses as often as one would expect in light of [the case law].”).

165. Perhaps the nearest known analog is Marketa Trimble’s study of thirteen cases involving patent-infringement “[i]njunctions issued against foreign entities.” Trimble, *supra* note 96, at 339. Trimble noted in passing that at least two of the injunctions in the study contained language that, according to Federal Circuit precedent, violates Rule 65(d). *Id.* at 340.

166. See, e.g., *Harris v. City of Phila.*, 47 F.3d 1342, 1349 (3d Cir. 1995) (stating that Rule 65(d) is “also applicable to consent decrees”); *Converse Inc. v. Reebok Int’l Ltd.*, 328 F. Supp. 2d 166, 176 (D. Mass. 2004) (“[T]he Court must ensure that the consent decree conforms to the strictures of Federal Rule of Civil Procedure 65(d) . . .”). See generally LAYCOCK, *supra* note 78, at 345 (discussing the principle that parties cannot contract for an injunction otherwise beyond a court’s power to grant).

issuing an injunction is generally unopposed—for example, because of default.¹⁶⁷ At the very least, a trial judge might rightly suspect that a consented-to or default-judgment injunction is substantially less likely to be subjected to appellate scrutiny than an injunction issued over party opposition. Lack of fear of appellate reversal might result in less care to avoid legal error.¹⁶⁸

Indeed, the three categories of consented-to injunctions, otherwise-unopposed injunctions, and the rest, which I term “actively opposed injunctions,” exhibit facially distinct Type-2 error rates. In the 2010 dataset, the percentage of consented-to injunctions featuring Type-2 error is essentially the same as the overall average: about 57% (47 of 83). A higher rate of Type-2 error characterizes the dataset’s nineteen otherwise unopposed injunctions—mostly default-judgment injunctions and one temporary restraining order issued without prior notice to the other side.¹⁶⁹ These unconsented-to but unopposed injunctions have a Type-2 error rate of just over 89% (17 of 19). The Type-2 error rate for actively opposed injunctions is lower, with about 44% (18 of 41) of this last category of injunctions including Type-2 language.

Although the observed Type-2 error rates for these three subclasses of injunctions are distinct, they all seem relatively high. Moreover, these high levels do not appear to be mere flukes resulting from a limited sample size. According to a standard *t*-test, the distinction between the approximately 57% error rate observed for the eighty-three consented-to injunctions and a hypothesized 45% error rate is statistically significant at a 95% confidence

167. The question of whether an injunction should be classified as opposed or unopposed is not necessarily entirely straightforward. In this category of otherwise unopposed injunctions, I do not include injunctions whose underlying bases—e.g., the validity of the patents the injunctions sought to enforce—were contested by not presently defaulting parties that were targets of the injunction, even if, following determinations on the merits of liability, the issuance of an injunction itself does not seem to have been specifically opposed or if the merits were contested before a magistrate judge but not subsequently before the relevant district judge. *Cf.* Docket Entry No. 218, *Duramed Pharms., Inc. v. Watson Labs., Inc.*, No. 3:08-cv-00116-LRH-WGC (D. Nev. Apr. 15, 2010) (recording the filing of an “Unopposed Motion for Entry of Final Judgment”). The qualification about the parties at issue being not presently defaulting reflects contemplation of the “anomalous” procedural situation surrounding an injunction in *Ocean Innovations, Inc. v. Quarterberth, Inc.*, No. 1:03-CV-0913, slip op. at 6–7 (N.D. Ohio May 14, 2010) (memorandum opinion and order) (“[T]he procedural history and posture of this case is certainly anomalous.”), which issued after various parties had defaulted on the merits but then had appeared to contest the default. *Cf.* *Ocean Innovations, Inc. v. Quarterberth, Inc.*, No. 1:03-CV-0913 (N.D. Ohio May 14, 2010) (issuing a permanent injunction enjoining defendants from various activities).

168. *Cf.* David E. Klein & Robert J. Hume, *Fear of Reversal as an Explanation of Lower Court Compliance*, 37 *LAW & SOC’Y REV.* 579, 603 (2003) (suggesting that “trial judges in virtually every court system” might be particularly sensitive, compared to U.S. circuit court judges, to “[t]he threat of reversal”).

169. *See* Temporary Restraining Order, *Valvtechnologies, Inc. v. North*, No. 4:10-cv-03943 (S.D. Tex. Oct. 20, 2010) (issuing a temporary restraining order without notice to the defendants).

level.¹⁷⁰ The same holds true for the distinction between the approximately 44% error rate observed for the forty-one actively opposed injunctions and a hypothesized 25% error rate.¹⁷¹ Likewise, statistical analysis suggests that the over 89% error rate observed for the nineteen unconsented-to but unopposed injunctions is significantly distinct from a hypothesized error rate for such orders of 70%.¹⁷²

The lower observed error rate for actively opposed orders as opposed to consented-to orders or otherwise unopposed orders might suggest that truly adversarial proceedings in the district courts help prevent improperly drafted decrees. Litigators should probably not congratulate themselves too much, however, for this apparent gain from adversarial process. For starters, the difference between the approximately 44% error rate observed for actively opposed injunctions and the approximately 57% error rate observed for consented-to injunctions does not appear to be statistically significant even at a relatively low 85% confidence level.¹⁷³ Further, the discrepancy between the observed error rates for these categories all but disappears after the sample is truncated to eliminate patent-infringement injunctions directed exclusively to either of two idiosyncratic forms of subject matter: biomedical-substance technology and ornamental designs.¹⁷⁴ By contrast, the differences between the approximately 89% observed error rate for unconsented-to but unopposed injunctions (predominantly defaults) and either of the error rates for consented-to orders and actively opposed orders appears to be statistically significant at a 99% confidence level.¹⁷⁵ These

170. More specifically, under a one-sample, two-tailed *t*-test applied to the sample of eighty-three consented-to injunctions, a null hypothesis that the real error rate is 45% is rejected at a 95% confidence level, with $t(82) = 2.1, p = 0.04$. If a one-tailed *t*-test is used to test a null hypothesis that the real error rate is less than or equal to 45%, the null hypothesis is rejected at a 95% confidence level by a more substantial margin, $t(82) = 2.1, p = 0.02$.

171. Under a one-sample, two-tailed *t*-test applied to the sample of forty-one actively opposed injunctions, a null hypothesis that the real error rate is 25% is rejected at a 95% confidence level, with $t(40) = 2.4, p = 0.02$. If a one-tailed *t*-test is used to test a null hypothesis that the real error rate is less than or equal to 25%, the null hypothesis is rejected at a 95% confidence level by a more substantial margin, $t(40) = 2.4, p = 0.01$.

172. Under a one-sample, two-tailed *t*-test applied to the sample of nineteen unconsented-to but unopposed injunctions, a null hypothesis that the real error rate is 70% is rejected at a 95% confidence level, with $t(18) = 2.7, p = 0.015$. If a one-tailed *t*-test is used to test a null hypothesis that the real error rate is less than or equal to 70%, the null hypothesis is rejected at a 95% confidence level by a more substantial margin, $t(18) = 2.7, p = 0.0075$.

173. When the results for the eighty-three consented-to injunctions and the forty-one actively opposed injunctions are compared, two-sample, two-tailed *t*-tests assuming equal variances and assuming unequal variances yield, respectively, $t(122) = 1.3, p = 0.18$; and $t(79) = 1.3, p = 0.19$. An *F*-test for equality of variances did not indicate a statistically significant difference between variances ($F = 0.98, p = 0.47$).

174. See *infra* text accompanying notes 190–91.

175. When the results for the eighty-three consented-to injunctions and the nineteen otherwise-unopposed injunctions are compared, two-sample, two-tailed *t*-tests assuming equal variances and assuming unequal variances yield, respectively, $t(100) = -2.7, p = 0.007$; and $t(42) = -3.6, p = 0.0008$. An *F*-test for equality of variances indicated a statistically significant difference between variances at a 95% confidence level ($F = 2.15, p = 0.015$). When the results for the forty-one

differences appear to remain significant only at a lower confidence level once results are truncated to eliminate biomedical-substance and ornamental-design orders.¹⁷⁶ In any event, a conclusion that active opposition and agreement both reduce error rates relative to default but that active opposition does not reduce error rates relative to agreement would seem no great triumph for the adversarial process.

Finally, regardless of comparisons with the observed error rates for consented-to and otherwise-unopposed injunctions, the 44% observed error rate for actively opposed injunctions seems high in light of the fact that compliance with the ban on Type-2, obey-the-law injunctions does not appear to be particularly difficult. Neither Type-0, measure-zero nor Type-1, colorable-differences language is hard to draft. Yet nearly 50% of the eighteen actively opposed injunctions that feature Type-2 language (8 of 18) do not feature either Type-0 or Type-1 language. The only orders of a do-not-infringe form that these injunctions contain is of the forbidden, obey-the-law type. Even if one views it as predictable that successful plaintiffs will seek, and often obtain, facially overreaching orders for injunctive relief, one might have thought those plaintiffs would also be careful to include legally proper Type-0 or Type-1 language so that any later-discovered impropriety in injunction scope might be viewed as easily severed from an otherwise proper order. Indeed, ten of the eighteen actively opposed injunctions featuring Type-2 language do include Type-0 language, and five of these ten also include Type-1 language. Of course, in such instances, one might wonder why opposing parties and courts have together failed to “get things right” by having the offending obey-the-law language stricken while leaving the rest of the order intact.

In any event, even if legal compliance were not so easy, a statistically significant error rate of more than 25% with respect to the scope of an actively opposed remedial order might be viewed as surprisingly high. An appellate reversal rate of about 35% or so on questions of claim construction is commonly thought to signal serious problems with the way U.S. courts work.¹⁷⁷ This is true even though (a) claim construction is generally an interlocutory issue that is often difficult to resolve¹⁷⁸ and (b) claim-construction issues pursued in an appeal might be assumed, on average, to be

actively opposed injunctions and the nineteen unconsented-to but unopposed injunctions are compared, two-sample, two-tailed *t*-tests assuming equal variances and assuming unequal variances yield, respectively, $t(58) = -3.6, p = 0.0006$; and $t(53) = -4.3, p = 0.00008$. An *F*-test for equality of variances indicated a statistically significant difference between variances at a 95% confidence level ($F = 2.5, p = 0.018$).

176. See *infra* note 191 and accompanying text.

177. See, e.g., Golden, *supra* note 20, at 324 & n.15 (noting that “claim construction jurisprudence continues to bear hallmarks of unpredictability” and that “[r]eversal rates of district court claim constructions stand at roughly 34%”).

178. Cf. *id.* at 386 (“[E]ven a more coherent claim construction jurisprudence will leave room for uncertainty regarding the meanings of particular claims . . .”).

unusually tricky.¹⁷⁹ In contrast, Type-2 error can be far from even arguably subtle. In *Monsanto Co. v. Bowman*,¹⁸⁰ for example, Monsanto obtained an actively opposed permanent injunction of breathtaking breadth: rather than restricting itself to enforcement of the patent rights at issue, this order “permanently enjoined [the defendant] from making, using, selling or offering to sell any of Monsanto’s patented crop technologies.”¹⁸¹

2. *Preliminary Versus Permanent Injunctions*.—The 2010 dataset can be used to test another potential hypothesis: judges or parties should be less susceptible to Type-2 error in the preliminary-relief context. Preliminary injunctions and temporary restraining orders are rarer than permanent injunctions and, because of less lead time from the date of suit, more likely to significantly disrupt prior, reasonable expectations.¹⁸² Further, from a court’s perspective, such injunctions are naturally more suspect because they become effective before the court has made a final decision on the merits according to ordinary procedure.¹⁸³ A supposedly wronged rights holder might ultimately turn out to own no valid rights or at least no valid rights that were violated. In light of the preceding, courts and parties might be expected to be especially careful in policing the form of preliminary injunctions and temporary restraining orders.

Consistent with this expectation, the percentage of permanent injunctions featuring Type-2, obey-the-law language is greater than the percentage of preliminary injunctions or temporary restraining orders doing the same. Just over 60% of permanent injunctions (75 of 124) and about 37% of preliminary injunctions or temporary restraining orders (7 of 19) feature Type-2 language. According to a *t*-test, this difference is significant at a 90% confidence level.¹⁸⁴

179. See *id.* at 324 n.15 (“Of course, the relatively high reversal rates for patent claim construction could be explained by litigants’ greater selectivity in choosing which claim constructions to appeal, rather than any atypical failure on the part of courts.”).

180. No. 2:07-cv-283-RLY-WGH slip op. (S.D. Ind. May 12, 2010).

181. *Id.* at 1.

182. See DOUGLAS LAYCOCK, *THE DEATH OF THE IRREPARABLE INJURY RULE* 111, 116 (1991) (noting that “a preliminary order may inflict serious costs on a defendant who had little time to prepare a defense” and that “[i]t is almost universally true that courts are more willing to grant permanent injunctions than preliminary injunctions”); Golden, *supra* note 6, at 2159 n.178 (“During the several years that it can take to obtain a permanent injunction, the defendant may redesign its accused product or process multiple times, or perhaps stop manufacturing or using it simply because it has become obsolete.”).

183. See *id.* at 111 (noting that one “reason[] for being cautious with preliminary relief” is that a “court must act without a full trial, sometimes on sketchy motion papers and affidavits”).

184. When the results for the 124 permanent injunctions and the 19 preliminary injunctions (including temporary restraining orders) are compared, two-sample, two-tailed *t*-tests assuming equal variances and assuming unequal variances yield, respectively, $t(141) = 2.0$, $p = 0.053$; and $t(24) = 1.9$, $p = 0.06$. An *F*-test for equality of variances does not indicate a statistically significant difference between variances ($F = 0.98$, $p = 0.44$).

But this superficially significant difference between the observed error rates for preliminary and permanent injunctions is likely deceiving. As discussed below, once orders directed to biomedical-substance or ornamental-design patents are eliminated from consideration, the discrepancy between observed Type-2 error rates for preliminary and permanent injunctions essentially disappears.¹⁸⁵ Generally speaking, the 2010 dataset does not appear to provide strong support for a hypothesis that, under comparable circumstances, there is a significant difference in Type-2 error rates for preliminary and permanent relief.

3. *Subject-Matter Specificity of Error Rates and Earlier Error Rates Revisited.*—In addition to highlighting distinctions between preliminary and permanent relief, the 2010 dataset reinforces a common impression that the practical operation of patent law can be very technology specific.¹⁸⁶ One substantial subset of orders leaps out as one in which the observed Type-2 error rate is very low. This is the subset of orders focused on biomedical-substance technology, predominantly pharmaceuticals. Only three out of twenty-five BMS orders in the dataset use Type-2, obey-the-law language. The Type-2 error rate for BMS orders is thus only 12%. For the twenty-two BMS orders dealing with a pharmaceutical substance apparently subject to Food and Drug Administration regulation, the error rate is even lower: only one of the twenty-two orders (about 5%) contains Type-2 language.¹⁸⁷

General lack of Type-2, obey-the-law language is merely one aspect of BMS orders' idiosyncrasy. Remarkably, only two of the twenty-five BMS orders (8%) even bother to include Type-1, colorable-differences language. The overwhelming majority of BMS orders are simple Type-0, measure-zero orders lacking explicit extension even to matter "no more than colorably different" than that expressly described. There seem to be at least two related explanations for this unusual satisfaction with Type-0 orders: such orders appear to have more than measure-zero weight in this context because (a) the quantum nature of matter renders it unusually difficult if not impossible to make satisfactory "tweaks" (e.g., addition or omission of an atom here or there) to design around many BMS technologies, perhaps particularly small-molecule pharmaceuticals; and (b) as an indirect result of this fact, an injunction directed at BMS drug technology can often be directed at an Abbreviated New Drug Application, in which an applicant seeking Food

185. See *infra* notes 192–94 and accompanying text.

186. See, e.g., John M. Golden, *Innovation Dynamics, Patents, and Dynamic-Elasticity Tests for the Promotion of Progress*, 24 HARV. J.L. & TECH. 47, 105 (2010) (noting further reason "to believe that a policy mechanism like patent law will have disparate effects for different technologies and industries").

187. The majority of BMS orders (15 of 25) were actively opposed, and none involved situations of actual or effective default. Somewhat interestingly, Type-2 language appears exclusively in actively opposed BMS orders: three of the fifteen actively opposed BMS orders feature Type-2 language, whereas no consented-to BMS order includes such language.

and Drug Administration approval for drug commercialization commonly looks to establish that a proposed generic drug has the same active ingredients as, and bioequivalence to, the patentee's brand-name pharmaceutical.¹⁸⁸ These peculiarities of the BMS context appear to enable patentees to receive adequate protection without the hazy language of a Type-1, colorable-differences order. In short, whether because of relatively inherent qualities of the subject matter, heavy government regulation, or perhaps some other cause, BMS technology lends itself to a peculiarly high degree of precision in rights definition and enforcement.¹⁸⁹

Indeed, outside BMS orders, Type-2 error is rife. Over 63% of utility-patent orders coded as not involving BMS technology (67 of 105) contain Type-2, obey-the-law language. For orders relating to design patents and no other types of patents (purely design-patent orders), Type-2 language is nearly ubiquitous: twelve of the thirteen such orders in the 2010 dataset (about 92%) contain Type-2 language.

The dominance of Type-2 error for purely design-patent orders and the general lack of Type-2 error for BMS orders suggests that we should revisit the figures discussed in earlier subsections after excluding such orders. In large part because, unlike other types of patent-infringement injunctions, most BMS orders are actively opposed, the main notable effect of excluding BMS and purely design-patent orders is that the numerical discrepancy between the Type-2 error rates for consented-to and actively opposed injunctions—which was already suggested to lack statistical significance¹⁹⁰—essentially disappears. About 61% of the remaining consented-to injunctions (39 of 64) feature Type-2 error. For the remaining actively opposed injunctions, the Type-2 error rate is about 58% (15 of 26). But the Type-2 error rate for unconsented-to but unopposed injunctions (mostly from

188. See 21 U.S.C. § 355(j)(2)(A) (2006) (mandating that “an abbreviated application for a new drug shall contain . . . information to show that the active ingredients of the new drug are the same as those of the listed drug” and “information to show that the new drug is bioequivalent to the listed drug”); see also David E. Adelman & Christopher M. Holman, *Misplaced Fears in the Legislative Battle Over Affordable Biotech Drugs*, 50 IDEA 565, 580 (2010) (“In the case of traditional drugs, [follow-on drug] assessment turns on the chemical identity and purity of a generic drug (i.e., whether it is ‘bioequivalent’ and employs the ‘same’ active ingredient), both of which involve testing methods that are accurate and precise.”).

189. See JAMES BESSEN & MICHAEL J. MEURER, PATENT FAILURE: HOW JUDGES, BUREAUCRATS, AND LAWYERS PUT INNOVATORS AT RISK 152 (2008) (“The usual explanation for the superior performance of patents in [chemical and pharmaceutical] technologies is that the boundaries of chemical patents are clearer . . . —the structure of a molecule or the composition of a mixture can be defined with precision.”). One large subclass of BMS patent litigation is litigation based on § 271(e) of the U.S. Patent Act, which makes the filing of an application for Food and Drug Administration approval of a patented drug an act of infringement. 35 U.S.C. § 271(e)(2) (2006). Subsection 271(e)(4) makes distinct and exclusive provision for the remedies that are available for such an act of infringement, declaring, in effect, that in many cases the only available relief will be “injunctive relief . . . to prevent the commercial manufacture, use, offer to sell, or sale within the United States or importation into the United States of an approved drug or veterinary biological product.” *Id.* § 271(e)(4).

190. See *supra* text accompanying note 173.

defaults) remains atypically high: thirteen of the fifteen remaining injunctions in this subcategory (nearly 87%) feature Type-2 error. Unsurprisingly, the differences between the approximately 61% and 58% error rates observed for the remaining consented-to and actively opposed injunctions, respectively, do not appear to be statistically significant at any plausibly meaningful confidence level.¹⁹¹ On the other hand, the differences between the observed error rates for each of these two classes of injunctions and the observed error rate for the remaining unconsented-to but unopposed injunctions might be significant, with *t*-tests indicating such significance at a 90% or 95% confidence level.¹⁹²

With respect to preliminary versus permanent injunctions, the dataset for preliminary injunctions is so small after BMS and purely design-patent orders are excluded that little of significance can be said. For permanent injunctions outside the BMS and purely design-patent categories, the Type-2 error rates are about 65% overall (62 of 96) and about 60% for actively opposed orders (12 of 20). Meanwhile, five of the nine preliminary injunctions in the residuum (about 56%) feature Type-2 error. Three of the six of these preliminary injunctions that were actively opposed (exactly 50%) feature such error. None of the differences in comparable preliminary-versus-permanent error rates in these residual samples appear to be statistically significant.¹⁹³ Likewise, differences in the Type-2 error rates for

191. When the results for the sixty-four remaining consented-to injunctions and the twenty-six remaining actively opposed injunctions are compared, two-sample, two-tailed *t*-tests assuming equal variances and assuming unequal variances yield, respectively, $t(88) = 0.28, p = 0.78$; and $t(45) = 0.28, p = 0.78$. An *F*-test for equality of variances does not indicate a statistically significant difference between variances ($F = 0.95, p = 0.42$).

192. When the results for the sixty-four remaining consented-to injunctions and the fifteen remaining unconsented-to but unopposed injunctions are compared, two-sample, two-tailed *t*-tests assuming equal variances and assuming unequal variances yield, respectively, $t(77) = -1.9, p = 0.06$; and $t(28) = -2.3, p = 0.03$. An *F*-test for equality of variances does not indicate a statistically significant difference between variances at a 95% confidence level but does indicate such a difference at a 90% confidence level ($F = 2.0, p = 0.08$). When the results for the twenty-six remaining actively opposed injunctions and the fifteen remaining unconsented-to but unopposed injunctions are compared, two-sample, two-tailed *t*-tests assuming equal variances and assuming unequal variances yield, respectively, $t(39) = -1.9, p = 0.06$; and $t(37) = -2.2, p = 0.04$. An *F*-test for equality of variances does not indicate a statistically significant difference between variances at a 95% confidence level but does indicate such a difference at a 90% confidence level ($F = 2.1, p = 0.08$).

193. No preliminary injunctions in the residuum were consented-to. When the results for the ninety-six remaining permanent injunctions and the nine remaining preliminary injunctions (including temporary restraining orders) are compared, two-sample, two-tailed *t*-tests assuming equal variances and assuming unequal variances yield, respectively, $t(103) = 0.53, p = 0.59$; and $t(9) = 0.49, p = 0.63$. An *F*-test for equality of variances does not indicate a statistically significant difference between variances ($F = 0.83, p = 0.31$). Likewise, when the results for the twenty of these permanent injunctions that were actively opposed and the six of these preliminary injunctions that were actively opposed are compared, two-sample, two-tailed *t*-tests assuming equal variances and assuming unequal variances yield, respectively, $t(24) = 0.42, p = 0.68$; and $t(8) = 0.40, p = 0.70$. An *F*-test for equality of variances does not indicate a statistically significant difference between variances ($F = 0.84, p = 0.35$).

preliminary and permanent injunctions in the relatively small samples of BMS and purely design-patent orders do not appear statistically significant or, for that matter, very substantial even on their face.¹⁹⁴

4. *Geographic Ubiquity of Error.*—Type-2 error exhibits substantial geographic ubiquity as well as substantial technologic ubiquity. Among districts that issued at least two non-BMS injunctions in 2010, only one—the Southern District of Ohio—avoided Type-2 error.¹⁹⁵ Among districts issuing three or more non-BMS injunctions, the Southern District of California had the best batting average: its six non-BMS injunctions feature only two that contain Type-2, obey-the-law language. Among districts that issued five or more of the injunctions in the dataset, Table 1 shows that the Districts of Delaware and New Jersey were best at avoiding Type-2 error. But this might largely reflect the dominance of BMS-oriented patent cases in these districts. All five of the District of Delaware's injunctions deal with a particular form of BMS subject matter: regulated pharmaceuticals. Twelve of the District of New Jersey's injunctions can be similarly described, and the remaining three injunctions from that district are purely design-patent orders that account for all three of the district's orders using Type-2 language. In short, the relative success of the Districts of Delaware and New Jersey in complying with the Federal Circuit's interpretation of Rule 65(d) might be more attributable to technologic subject matter than to any special legal acumen on the part of these courts or the attorneys who practice before them.

194. The observed Type-2 error rates for the sixteen BMS permanent injunctions and the nine BMS preliminary injunctions (including temporary restraining orders) are approximately 13% (2 of 16) and 11% (1 of 9), respectively. The observed Type-2 error rates for the twelve purely design-patent permanent injunctions and the one purely design-patent preliminary injunction are approximately 92% and 100%, respectively. Given the closeness of the preliminary and permanent injunction error rates and the small sizes of the samples, it virtually goes without saying that the observed differences within the respective BMS and purely design-patent classes do not appear to be statistically significant.

195. See *Lexmark Int'l, Inc. v. Ink Techs. Printer Supplies, LLC*, No. 1:10-cv-00564-MRB (S.D. Ohio Dec. 21, 2010) (issuing a permanent injunction directed at Print-Rite Holdings Ltd.); *Lexmark Int'l, Inc. v. Ink Techs. Printer Supplies, LLC*, No. 1:10-cv-00564-MRB (S.D. Ohio Dec. 9, 2010) (issuing a permanent injunction directed at Virtual Imaging Products, Inc.).

Table 1. Leading Districts for Injunctive Relief in the 2010 Dataset

District	Number of Injunctions	Number of Injunctions with Type-2 Language	Number of Non-BMS and Non-Purely-Design-Patent Injunctions	Number of Non-BMS and Non-Purely-Design-Patent Injunctions with Type-2 Language
C.D. Cal.	19	11	14	7
D.N.J.	15	3	—	—
D. Utah*	10	10	9	9
E.D. Tex.	9	6	9	6
N.D. Tex.	6	5	6	5
S.D. Cal.	6	2	6	2
N.D. Ga.	5	3	5	3
S.D. Fla.	5	2	3	1
D. Del.	5	0	—	—

* Nine of the injunctions issued by the District of Utah were issued by one judge in a single case.

5. *Potential Explanations for High Rule 65(d) Error Rates.*—What explains the district courts' mass violation of the Federal Circuit's understanding of Rule 65(d)? Limitation of a violation's expected consequences provides one plausible explanation. As discussed earlier, if Type-2 error is not corrected on direct appeal, the general result is simply that Type-2, obey-the-law language will be enforced as if it were Type-1, colorable-differences language.¹⁹⁶ Consequently—and particularly as Type-1 or Type-0, measure-zero language often accompanies Type-2 language—courts and parties might commonly view Type-2 error as essentially harmless.¹⁹⁷ This might be especially true when parties have settled their differences, perhaps through a broad licensing or cross-licensing agreement, and seek an injunction essentially only to reinforce their settlement despite lack of any apparent reason for serious concern about a risk of future infringement. A perception of relative lack of harm from Type-2 error might also result in relative lack of attention by trial judges and attorneys to Federal Circuit opinions identifying and correcting such error.

196. See *supra* text accompanying notes 106–08.

197. Of course, attorneys and their clients still need to consider the risk that, whether through error or a legal course change, courts will later either enforce Type-2, obey-the-law language as written (a risk for the adjudged infringer and its attorneys) or refuse to enforce the injunction at all (a risk for the patentee and its attorneys). Given the apparent ease of avoiding such risks, one might have expected attorneys to more strictly avoid Type-2 language.

This last potential result of a perception of harmlessness dovetails with another possible explanation for the frequency of Type-2 error—namely, that the very prevalence of such error generates noncorrective inertia. The familiarity of Type-2, obey-the-law language might result in its being perceived as unsuspecting. Such a perception is likely encouraged by the fact that Type-2 orders and even Type-0 orders in combination with Type-0, measure-zero orders have pedigrees that stretch back well over a century.¹⁹⁸ The proposition that such pedigrees can contribute to relatively unthinking inertia, even when seemingly simple fixes might better serve a client's interests, appears to have some support in the observation made by contracts scholars that “standard commercial contract[s]” can “pass relatively untouched through the hands of generations of lawyers” even when, in view of later legal developments, “one might have expected the elite practicing bar to have reacted immediately and decisively” to deviate from the previously established standard form.¹⁹⁹

A further contributor to Type-2 error could be Type-2, obey-the-law orders' partial specificity. This might be important because, in applying U.S. law generally, courts commonly frown on injunctions that “do[] no more than instruct a defendant to ‘obey the law.’”²⁰⁰ But Type-2 do-not-infringe

198. Evidence of Type-2 orders appears at least as early as the mid-nineteenth century. See *Cal. Artificial Stone Paving Co. v. Molitor*, 113 U.S. 609, 613, 617–18 (1885) (refusing to determine the scope of an injunction against “making, selling, or using, or in any manner disposing of, any artificial stone-block pavements embracing the invention and improvements described in the said reissued letters patent”); *Corning v. Troy Iron & Nail Factory*, 56 U.S. (15 How.) 451, 456 (1854) (reporting the trial court's grant of an injunction against “in any manner infringing or violating any of the rights or privileges granted or secured by said patent”); see also *KSM Fastening Sys., Inc. v. H.A. Jones Co.*, 776 F.2d 1522, 1533 n.1 (Fed. Cir. 1985) (Newman, J., concurring in part) (“The majority posits the ‘unreasonableness of a decree incorporating a vague or broad prohibition against “infringement” of a “patent.”’ A century of precedent weighs to the contrary.”), *overruled on other grounds*, *TiVo Inc. v. EchoStar Corp.*, 646 F.3d 869 (Fed. Cir. 2011) (en banc). The same is true for injunctions including both Type-2 and Type-0 orders. See *Barnard v. Gibson*, 48 U.S. (7 How.) 650, 653 (1849) (reporting the trial court's grant of an injunction against “any further constructing or using in any manner . . . of the two planing-machines mentioned in said bill . . . and [against] infringing upon or violating the said patent in any way whatsoever”). The practice of limiting the effective scope of injunctions to matters judged to be infringing and only colorable variants thereof appears to have comparable lineage. See, e.g., *Crown Cork & Seal Co. of Balt. City v. Am. Cork Specialty Co.*, 211 F. 650, 653 (2d Cir. 1914) (“It has been the practice in this circuit not to deal with modifications of a machine held to be an infringement, on motions to punish for contempt, unless the change was plainly a mere colorable equivalent . . .” (citations omitted)); *Onderdonk v. Fanning*, 2 F. 568, 569 (E.D.N.Y. 1880) (concluding that a difference from a device previously adjudged to infringe “was not so plainly colorable as to entitle the plaintiff to an attachment against him for contempt”); cf. ALBERT H. WALKER, *TEXT-BOOK OF THE PATENT LAWS OF THE UNITED STATES OF AMERICA* § 708, at 555 (4th ed. 1904) (“And an attachment will not issue where the character of the defendant's doings, after the injunction, is doubtful.”).

199. G. Mitu Gulati & Robert E. Scott, *Introduction: The Three and a Half Minute Transaction: Boilerplate and the Limits of Contract Design* 5 (Columbia Law Sch., Working Paper No. 410, 2011), available at <http://ssrn.com/abstract=1945988>.

200. *Burton v. City of Belle Glade*, 178 F.3d 1175, 1201 (11th Cir. 1999) (“As [an] injunction [against racial discrimination in annexation] would do no more than instruct the City to ‘obey the law,’ we believe that it would not satisfy the specificity requirements of Rule 65(d) . . .”); see also

orders typically refrain from generally forbidding any future violations of 35 U.S.C. § 271 or the U.S. Patent Act as a whole. Instead, these orders ordinarily forbid future infringement only of specific patents or patent claims that the defendant is adjudged to have infringed already. Even among those familiar with the general rule against obey-the-law orders, this partial specificity could nurture an uninformed confidence that a patent-specific Type-2 order complies with Rule 65(d). This confidence might be particularly strong in the design-patent context because the definition of infringement for a design patent typically already has something like Type-1 scope—albeit Type-1 scope centered on a central, patent-described embodiment, rather than the embodiment previously used by an adjudged or accused infringer: U.S. design patents generally claim only one or more ornamental designs specifically illustrated by diagrams included in the patent,²⁰¹ and the test for infringement is whether an allegedly infringing design is substantially similar to one or more of those shown.²⁰² Thus, for design patents, parties and district courts might be well justified in viewing a Type-2 injunction’s literal, as well as effective, scope as fundamentally equivalent to that of a legitimate Type-1 injunction.

In any event, even with respect to what U.S. law calls utility patents, courts in other common law jurisdictions have indicated that patent-limited injunctions of Type-2 form are sufficiently detailed to provide proper notice of bases for finding contempt. Canadian courts have indicated that an order prohibiting future infringement of a particular patent or patent claim is adequately instructive.²⁰³ Likewise, courts in the United Kingdom (U.K.)

Payne v. Travenol Labs., Inc., 565 F.2d 895, 898 (5th Cir. 1978) (holding that an antidiscrimination injunction “more specific than Title VII itself only in that it does not prohibit employment discrimination based on religion and natural origin . . . cannot be sustained”); cf. LAYCOCK, *supra* note 78, at 274 (describing Rule 65(d) as “generally preclud[ing] injunctions that merely tell defendant to ‘obey the law’”).

201. See, e.g., U.S. Patent No. D618,225 S, at [57] (filed July 13, 2009) (“The ornamental design for a cellphone plug adapter, as shown and described.”); U.S. Patent No. D456,023 S, at [57] (filed July 17, 2000) (“The ornamental design for a display, as shown and described.”).

202. See *Egyptian Goddess, Inc. v. Swisa, Inc.*, 543 F.3d 665, 678 (Fed. Cir. 2008) (describing proof of infringement as requiring the patentee to “prov[e] the two designs would appear ‘substantially the same’ to the ordinary observer”).

203. *Weatherford Can. Ltd. v. Corlac Inc.*, [2010] F.C. 667 paras. 17, 20 (Can. Ont. Fed. Ct.) (observing that an injunction against “infringing [patent] claims as interpreted whether [via] the named products or not” was “consistent with other orders of this Court, as affirmed by the Court of Appeal, restraining sale and distribution of infringing products generally”); see also *Merck & Co. v. Apotex Inc.*, [1999] 293 N.R. 316 (Can. Fed. Ct. App.) (rejecting a request to narrow an injunction prohibiting infringement of a specified patent so that an adjudged infringer might participate in activities involving “newly developed compounds” not available at the time of the trial judgment); cf. ROBERT J. SHARPE, *INJUNCTIONS AND SPECIFIC PERFORMANCE* (2d ed. 1997). Justice Sharpe of the Ontario Court of Justice observes:

[I]t has often been said that for negative injunctions a general form is to be used, provided it gives sufficient guidance, and orders prohibiting the defendant from acting ‘in the manner hitherto pursued by him or in any other manner so as to cause a nuisance’ and ‘in the manner complained of . . . or otherwise so as to cause a nuisance’ have been approved by appellate courts.

have viewed “the standard form of injunction” as one that “restrain[s] the defendant from infringing the patent.”²⁰⁴ To the extent the scope of such an injunction is not entirely clear, the U.K.’s Court of Appeal has indicated that “it is the infringer who should seek guidance from the court if he wishes to sail close to the wind.”²⁰⁵ Courts in Australia have taken a similar position. The Federal Court of Australia has stated, “Particularly when the validity of the patent has been an issue, the patentee is entitled to an injunction restraining all infringement, and not just the particular form of infringement which was the subject of evidence at the trial.”²⁰⁶

Courts in these common law jurisdictions are, like U.S. courts, sensitive to the need for injunctions to have clear scope.²⁰⁷ Their acceptance of Type-2, obey-the-law injunctions thus lends support to a notion that, at least as interpreted by the Federal Circuit, Rule 65(d)’s requirements are less than intuitive.²⁰⁸

On the other hand, the positions of these other courts might be wrong or at least suboptimal. Obtaining timely and adequate prior approval as suggested by the U.K.’s Court of Appeal might not be as easy as it sounds, particularly given the potentially evolving shape of a design-around product

Id. ¶ 1.400, at 1-17 (footnotes omitted).

204. *Coflexip S.A. v. Stolt Comex Seaway MS Ltd.*, [2001] 5 R.P.C. 182 (Ct. App.) 186–87 (Eng.) (Aldous, L.J.); *see also id.* (“The draft order contained an injunction in the usual form which restrained the defendants from ‘infringing European Patent (U.K.) No. 0478742.’”); *Nutrinova Nutrition Specialties & Food Ingredients GMBH v. Scanchem UK Ltd.* (No. 2), [2000] F.S.R. 831 (Patents Ct.) 838–39 (Eng.) (noting that although “there is no rule that in any case of infringement of an intellectual property right the injunction granted will be a general injunction against infringing that right[.]” “the discretion is a wide one [and] injunctive relief . . . may be in the wide form hitherto customary”). *See generally* *Microsoft Corp. v. Plato Tech. Ltd.*, [1999] Masons C.L.R. 370 (Ct. App.) (Eng.) (Nourse, L.J.) (agreeing that cited authorities “undoubtedly support . . . as a general proposition” “that in the case of infringement of intellectual property rights . . . it is well established . . . that, once the claimant has established any infringements of his rights at all, he is entitled as of right to an injunction in the usual wide form to restrain all future infringements”).

205. *Coflexip*, [2001] 5 R.P.C. at 188.

206. *Welcome Real-Time SA v Catuity Inc.* [2001] F.C.A. 785, para. 9 (Austl.); *see also id.* para. 11 (“The invariable practice in the High Court has been to grant an injunction which simply restrained infringing the patent . . .”).

207. *See, e.g.*, SHARPE, *supra* note 203, ¶ 1.390, at 1-16 (“Quite clearly, in formulating injunction orders, the courts should avoid vague or ambiguous language which fails to give the defendant proper guidance or which in effect postpones determination of what actually constitutes a violation of the plaintiff’s rights.”).

208. Nonetheless, I can cite two instances where district courts appear to have actively corrected a party’s attempt to obtain a Type-2, obey-the-law injunction. In one, a court denied an injunction altogether. *Plastic Tubing Indus., Inc. v. Blue Diamond Indus., LLC*, No. 6:10-cv-1227-GAP-KRS (M.D. Fla. Dec. 28, 2010) (denying parties’ “Joint Motion for Entry of Consent Final Judgment and Permanent Injunction” because “[t]he proposed consent judgment [was] essentially an extremely broad ‘obey the law’ injunction”). In the other, the court blacklined Type-2, obey-the-law language in a party’s draft order. *Proveris Scientific Corp. v. InnovaSystems, Inc.*, No. 05-12424-WGY, slip op. at 2–3 (D. Mass. May 11, 2007) (striking language that would have prohibited activities involving “any other product that embodies the patented inventions recited in claims 3-10 and 13” of U.S. Patent No. 6,785,400 and striking language requiring the destruction of “any other infringing products”).

or process and the possible need to commit significant resources before the product's final shape is fully determined. Further, where the adjudged infringer and patentee are competitors, the cost to an adjudged infringer of effectively having to signal through court proceedings the nature of a future product line might be inadequately appreciated or limited by courts and the protective orders that they issue.

In any event, the fact that a legal rule is less than entirely intuitive does not necessarily mean that it should not be expected to be understood or applied. Failure to understand or apply the rule might be explained by a combination of the inertia described above and the likely harmlessness of Type-2 error in many cases. Particularly where there are a variety of other grounds for potential appeal or trial-level dispute still at hand, the information costs involved in learning or recalling a rule of potentially only marginal importance might make efforts to avoid procedural error not worth the candle.

On the other hand, at least with respect to actively opposed injunctions, one might conjecture that other forces are at work. Because the overwhelming majority of patent-infringement disputes settle before an injunction issues,²⁰⁹ one might conjecture that a seemingly surprising level of Type-2 error, even during continued adversarial litigation, reflects the idiosyncratic nature of the parties or attorneys involved in those cases that do not settle. Among other things, failure of parties to settle a case before an injunction issues could signal (1) a relatively large continuing discrepancy in their views of the case's likely disposition on and after appeal or (2) asymmetrically high stakes for the rights holder under conditions where, for strategic or other reasons, reverse payments from the rights holder to an accused infringer are not a viable settlement option.²¹⁰ Such circumstances could correlate with comparatively lower quality attorney representation or

209. See Lemley & Shapiro, *supra* note 2, at 1992 (stating that "far more patents are licensed or settled than litigated to judgment").

210. See Keith N. Hylton & Sungjoon Cho, *The Economics of Injunctive and Reverse Settlements*, 12 AM. L. & ECON. REV. 181, 185 (2010) (developing a model that corroborates standard notions that asymmetric stakes or different probability-of-outcome expectations can prevent settlement where reverse payments are not an option). Reverse payments might fail to be a viable option because they are legally prohibited by, for example, antitrust laws. See Henry N. Butler & Jeffrey Paul Jarosch, *Policy Reversal on Reverse Payments: Why Courts Should Not Follow the New DOJ Position on Reverse-Payment Settlements of Pharmaceutical Patent Litigation*, 96 IOWA L. REV. 57, 61 (2010) (noting that both the U.S. Department of Justice and the Federal Trade Commission have "pursu[ed] antitrust liability for reverse-payment settlements"). Alternatively, reverse payments might fail to be a viable option because they undermine a right holder's ability to credibly threaten others with enforcement that does not entail payment to the others to avoid or cease infringement. Cf. Kaplow & Shavell, *supra* note 10, at 722 (noting that under a liability rule, "even though I would be willing to pay Jack not to take my car if it were inadequately valued by the courts, there would be no point in paying him to desist—for Jill, or someone else, could come along and take it the next day").

lower quality use of attorney representation (from a rational profit maximizer's perspective) on one or another side of the case.²¹¹

In any event, further work is needed to determine whether cases in which Type-2 error appears are, for example, peculiarly likely to feature attorneys less familiar with the nuances of Federal Rule of Civil Procedure 65(d) and its interpretation by the Federal Circuit. For purposes of this Article, key lessons from the prevalence of Type-2 error are that current practice with respect to do-not-infringe injunctions appears to leave much room for improvement and that patentees do not seem generally content with asking for injunctions of no more than Type-1, colorable-differences scope.

B. Purely Reparative Injunctions in U.S. Practice

In stark contrast to the mass violation of Federal Circuit precedent against Type-2, obey-the-law injunctions, district courts appear generally to heed Federal Circuit precedent holding that injunctions to enforce patent rights cannot be purely reparative.²¹² Among the 143 orders in the 2010 dataset, I identified only one that contained a purely reparative injunction. Further, this purely reparative injunction had a relatively trivial form: it was a consented-to order commanding an infringer to “provide a written letter of apology . . . that recognizes [the] infringement of the patents-in-suit, and apologizes for it.”²¹³ In short, compliance with Federal Circuit precedent against purely reparative injunctions appears to be quite good.

C. Specially Tailored Injunctions in U.S. Practice

The situation with respect to specially tailored patent-infringement injunctions is more complex, in part because such orders come in many different forms. Among the orders issued in 2010, I have identified four basic subcategories of specially tailored injunctions that are discussed in more detail below: (i) correlated-activity injunctions (Type-C); (ii) destruction, disablement, or delivery injunctions (Type-D);

211. Lower quality attorney representation or lower quality use of attorney representation could explain a discrepancy in the parties' views about likely case outcomes. Less experienced attorneys or attorneys whose interests are not well aligned with those of their clients might be less likely to convey an accurate sense of the likelihood of case outcomes. See Steven Shavell, *On the Design of the Appeals Process: The Optimal Use of Discretionary Review Versus Direct Appeal*, 39 J. LEGAL STUD. 63, 81–82 (2010) (“Attorneys have a personal interest in making appeals, as that means more work for them, which suggests that to some degree attorneys will give unduly optimistic advice to clients and promote excessive appeals if direct appeal is the only avenue of appeal.”). A less experienced or less legally savvy party might be more likely to misestimate such probabilities even if properly informed. On the other hand, party inexperience or lack of legal savvy might not help explain the high frequency of a procedural error like the issuance of a Type-2, obey-the-law injunction. Unlike a decision on whether to settle, a decision on whether to make a procedural objection might be almost exclusively within the domain of the attorneys.

212. See *supra* notes 111–13 and accompanying text.

213. *Batesville Servs., Inc. v. S. Rain Casket & Funeral Supply*, No. 2:09-CV-257-PPS-APR, slip op. at 5 (N.D. Ind. July 15, 2010).

(iii) “reformulated-bounds” injunctions (Type-B); and (iv) moderated injunctions (Type-M). Overall, injunctions having one or more of these specially tailored forms appear in about one-third of the 143 orders in the dataset. By far the most common of these specially tailored forms is the correlated-activity injunction.

1. *Correlated-Activity Injunctions.*—On over thirty occasions, courts issued a Type-C, correlated-activity injunction directed at activities that overlap significantly, but not entirely, with activities that by themselves can constitute infringement. Thus, for example, one district court prohibited not only “directly or indirectly infring[ing],” but also “causing, inducing or contributing to . . . infringement . . . by others.”²¹⁴ As only specific forms of causation of others’ infringement, such as active inducement or contributory infringement, constitute infringement under the U.S. Patent Act,²¹⁵ the court’s prohibition can be viewed as at least somewhat prophylactic. The prophylactic language might have been intended to protect against situations where required elements of indirect infringement are difficult to prove but indirect infringement is justifiably suspected.

Prohibition of other types of potentially noninfringing activity might reflect bleed over from other legal regimes. The U.S. Patent Act lists five kinds of acts that form bases for the most typical claims of direct infringement: “mak[ing], us[ing], offer[ing] to sell, or sell[ing] any patented invention, within the United States or import[ing] into the United States any patented invention.”²¹⁶ District courts commonly enjoin these five kinds of activity. But in the 2010 dataset, district courts also repeatedly forbade activities such as “distributing” or “shipping” infringing items, or even merely “displaying” images of these items.²¹⁷ Such language might be a carryover from other forms of intellectual property protection like trademark or copyright. U.S. copyright law explicitly gives copyright owners an exclusive distribution right with respect to “copies or phonorecords of the copyrighted work.”²¹⁸

On the other hand, prohibition of noninfringing activities such as shipping might have independent roots in a district court’s desire to prevent

214. *Seiko Epson Corp. v. Abacus 24-7 LLC*, No. 3:09-cv-00477-BR, slip op. at 8 (D. Or. July 8, 2010) (emphasis added).

215. See 35 U.S.C. § 271 (2006) (setting out different forms of infringement).

216. *Id.* § 271(a).

217. *E.g.*, *Silverlit Toys Manufactory Ltd. v. JP Commerce, LLC*, No. 2:09-CV-08959-CAS (JCX), slip op. at 4 (C.D. Cal. Apr. 29, 2010) (enjoining, inter alia, “marketing, reproducing, distributing, receiving, forwarding, shipping, displaying (on their websites or otherwise), or in any way commercially exploiting . . . any toy helicopters that infringe one or more claims” of two specified patents); *Innovation U.S.A., Inc. v. IDO Furniture (U.S.A.) Inc.*, No. 1:09-cv-01727-JBW-RLM, slip op. at 3 (E.D.N.Y. Mar. 31, 2010) (enjoining “referencing or depicting on their website or in any future catalog, brochure, and any other form of marketing literature,” a sofa or a reclining sofa bed that infringes either of two specified design patents).

218. 17 U.S.C. § 106(3) (2006).

future infringement. Such an activity, although potentially noninfringing by itself, can be highly correlated with infringing activities such as use or sale and might in fact enable those directly infringing activities to occur. Thus, prohibition of a correlated activity like shipping might provide a patentee with more effective and easily enforceable protection without chilling too much legitimate activity by an adjudged infringer.

Type-C injunctions directed at correlated activity need not be negative injunctions like the prohibitions discussed above.²¹⁹ For example, in *O2 Micro International Ltd. v. Beyond Innovation Technology Co.*,²²⁰ the district court facilitated protection of patent rights by requiring that an adjudged infringer label specified products as “Not for Sale in, Use in, or Importation into the United States.”²²¹ Although such labeling does not necessarily prevent the indicated activities, it might very well help discourage them and thus have a negative causal correlation with their occurrence.

2. *Destruction, Disablement, or Delivery Injunctions.*—Another subtype of specially tailored injunction is the Type-D injunction requiring destruction, disablement, or delivery of specified material.²²² There are seven Type-D orders in the dataset.

As with certain Type-C orders, repeated appearance of Type-D orders might reflect the influence of legal regimes such as copyright and trademark. Unlike the U.S. Patent Act, federal copyright and trademark acts expressly provide remedies of impoundment, destruction, or other court-ordered disposition of preexisting goods.²²³

219. Cf. ANDREW BURROWS, REMEDIES FOR TORTS AND BREACH OF CONTRACT 511–12 (3d ed. 2004) (describing the difference between “prohibitory” and “mandatory” injunctions); 11A WRIGHT, MILLER & KANE, *supra* note 92, § 2942, at 57 (“[I]njunctions compelling the doing of some act, as opposed to forbidding the continuation of a course of conduct, are an ancient and familiar tool of equity courts and will be used whenever the circumstances warrant.”).

220. No. 2:04-cv-00032-CE (E.D. Tex. Sept. 27, 2010).

221. *Id.* at 2. In *Polytree (H.K.) Co. v. Forests Mfg., Ltd.*, No. 1:09-cv-03377-WSD, slip op. (N.D. Ga. Dec. 20, 2010), the court commanded the U.S. Customs and Border Patrol to prevent importation into the United States of “any Christmas tree stand imported by Defendant marked with” a number from either of two patents owned by the plaintiff. *Id.* at 39. This order could be understood to facilitate prevention of infringement by releasing the patentee and customs officers from any need to prove or confirm that tree stands so marked in fact incorporated the indicated inventions. But the order could also be understood as primarily directed toward preventing further violations of the U.S. Patent Act’s false marking statute, which, among other things, prohibits the deceptive marking of a product “without the consent of the patentee.” 35 U.S.C. § 292(a) (2006).

222. See, e.g., *Saint-Gobain Technical Fabrics Am., Inc. v. Checkmate Geosynthetics, Inc.*, No. 6:09-cv-00557-MSS-KRS, slip op. at 19 (M.D. Fla. Oct. 26, 2010) (ordering defendant to “deliver, at its expense, to Saint-Gobain’s counsel all of its infringing products and all literature, advertisements and other materials related to [the] infringing products”); *Caught Fish Enters., LLC v. Blaze Wharton Constr., Inc.*, No. 09-cv-02878-PAB-KMT (D. Colo. Feb. 24, 2010) (requiring that the adjudged infringer “ship to Caught Fish at their own cost and for destruction . . . all Accused Clamps in their possession, custody, or control”).

223. 15 U.S.C. § 1118 (2006) (empowering courts in federal trademark actions to “order that all labels, signs, prints, packages, wrappers, receptacles, and advertisements in the possession of the defendant, bearing the registered mark . . . or any reproduction, counterfeit, copy, or colorable

Again, however, a bleed-over hypothesis is not the only available explanation. Although the Federal Circuit has indicated that “ordering the repatriation and destruction of [already] *exported*” matter can be too remote from any aim of preventing infringement to be authorized by the U.S. Patent Act,²²⁴ court-ordered destruction of goods located in the United States might well be understood to be a permissible sort of specially tailored order that helps prevent further infringement.

3. *Reformulated-Bounds Injunctions*.—In another seven orders in the dataset, courts issued what I term reformulated-bounds or Type-B injunctions. Such an injunction is distinctive in that it defines the scope of technologies that it encompasses without fundamentally relying on reference to or reproduction of a preexisting formulation or instantiation—i.e., without reference to the adjudged infringing products or processes, a patent or patent claim, or a description appearing elsewhere, such as an Abbreviated New Drug Application (ANDA).²²⁵ A Type-B injunction provides its own linguistic formulation of the subject matter that it encompasses.

Some consented-to orders illustrate Type-B injunctions in short form. In *Tristar Metals, Inc. v. Edemco Dryers, Inc.*,²²⁶ the district court issued a consented-to order that tersely forbade the defendant from engaging in commercial activities related to “any pet tub having a swing ramp.”²²⁷ The injunction covered any such tub even though the relevant patent’s claims were more specific: the claims apparently reached only pet tubs having “a plurality of leg elements.”²²⁸

Likewise, in *ExitExchange Corp. v. Casale Media Inc.*,²²⁹ the district court issued a consented-to order that broadly prohibited “making, using, importing, selling, or offering to sell pop-under advertisements”²³⁰—Internet advertisements that “appear underneath the active window” on a computer

imitation thereof, and all plates, molds, matrices, and other means of making the same, shall be delivered up and destroyed”); 17 U.S.C. § 503(b) (2006) (authorizing courts in copyright-infringement cases to order “the destruction or other reasonable disposition of all copies or phonorecords found to have been made or used in violation of the copyright owner’s exclusive rights, and of all plates, molds, matrices, masters, tapes, film negatives, or other articles by means of which such copies or phonorecords may be reproduced”).

224. *Johns Hopkins Univ. v. CellPro, Inc.*, 152 F.3d 1342, 1366 (Fed. Cir. 1998) (emphasis added).

225. So-called ANDA litigation that tends to result in injunctions making such a reference is enabled by § 271(e) of the U.S. Patent Act, 35 U.S.C. § 271(e) (Supp. IV 2010), which, roughly speaking, provides patentees with “the ability to sue [generic drug manufacturers] for merely filing an application (known as an Abbreviated New Drug Application, or ANDA) with the [Food and Drug Administration].” Timothy R. Holbrook, *Possession in Patent Law*, 59 SMU L. REV. 123, 141 (2006).

226. No. 4:10-cv-00044.A, slip op. (N.D. Tex. May 20, 2010).

227. *Id.* at 2.

228. U.S. Patent No. 6,516,752 col. 4 I. 46 (filed July 2, 2001); *id.* at col. 5 I. 22.

229. No. 2:10-cv-297-TJW, slip op. (E.D. Tex. Nov. 2, 2010).

230. *Id.* at 2.

screen and thus tend “not to be seen until some or all other browser windows are closed or minimized.”²³¹ The issued injunction omits specific limitations appearing in the corresponding patent’s claims, such as a limitation involving “a time interval beginning incrementally before said advertisement has completely finished loading.”²³²

Consequently, in both *Tristar* and *ExitExchange*, the issued injunction’s scope appears to depart from that of the more detailed patent claims. At least facially, the claims appear directed to more particular forms of pop-ups or pop-under advertisements than the corresponding injunction forbids. Apparently, therefore, each of these injunctions prohibits a spectrum of noninfringing activity.

The case of *Vertical Doors, Inc. v. Howitt*²³³ offers an example of a somewhat longer-form Type-B injunction.²³⁴ But of greatest immediate interest is a linguistically much more dramatic example of a Type-B injunction that appeared in *Stone Strong, LLC v. Del Zotto Products of Florida, Inc.*²³⁵ In this case, the court issued an unconsented-to order that defined the scope of its prohibition through extensive description of the physical characteristics of forbidden concrete blocks.²³⁶ Specifically, the court enjoined the defendant from the following:

[M]aking, using, offering to sell, selling within the United States or importing into the United States a precast concrete block with a front surface, first and second side surfaces, a top surface, a bottom surface and a back surface that contains a lifting device protruding from the top surface and a recess or notch in the bottom surface:

- i. where at least one recess or notch can be positioned to receive within it at least one lifting device from another block; and
- ii. where the width (“w”) of the recess or notch . . . is less than or equal to either:

231. Jennifer Yannone, Comment, *The Future of Unauthorized Pop-Up Advertisements Remains Uncertain as Courts Reach Conflicting Outcomes*, 7 TUL. J. TECH. & INTELL. PROP. 281, 282 (2005).

232. U.S. Patent No. 7,386,555 col. 15 II. 28–30 (filed Feb. 23, 2004); see also *id.* at col. 14 I. 64 to col. 15 I. 16 (claiming a “system for Internet advertising” comprising “a media that interacts with a display device to display to a user at least one browser,” “a script handler that invokes a post-session procedure” that “open[s] a second browser in a . . . background window,” and “an event handler that . . . loads [an] advertisement into said second browser”).

233. No. 2:09-cv-04685-JVS-AN, slip op. (C.D. Cal. Jan. 5, 2010).

234. Compare *id.* at 2 (forbidding activities involving “any vertical door conversion kits or vertical door hinges intended to allow opening of a vehicle door outward (i.e., in a horizontal motion like typical car doors), and then upward (i.e., in a vertical motion), and that are designed to be bolted on to the vehicle frame and door, as opposed to welded”), with U.S. Patent No. 6,845,547 col. 12 II. 46–64 (filed Nov. 26, 2002) (claiming a “vehicle door hinge for a vehicle door and frame, the hinge comprising: a chassis mounting plate securely fastened to such vehicle frame; a swingarm securely fastened to such vehicle door[,]” and other elements).

235. No. 5:08-cv-00503-WTH-DAB, slip op. (M.D. Fla. Nov. 19, 2010).

236. *Id.* at 1–2 (forbidding various activities involving “a precast concrete block” having various specified features).

1. the distance from the front of the block to the back end of the lifting device (“a”) . . . ; or
2. the distance from the back of the block to the front end of the lifting device (“b”)²³⁷

This injunction’s descriptive language does not track precisely the language of any relevant patent claim.²³⁸ For example, claim 1 of plaintiff Stone Strong’s U.S. Patent No. 6,796,098 covers:

1. A block comprising:

a front surface;

first and second side surfaces coupled to the front surface;

a top surface coupled to the front surface and to the first and second side surfaces, wherein the top surface includes at least one alignment device, each alignment device comprising a device for lifting the block when the block is being placed;

a bottom surface coupled to the front surface and to the first and second side surfaces, the bottom surface including at least one recess positioned to receive at least one alignment device of a previously-placed block to align the block with respect to the previously-placed block; and

a back surface coupled to the first and second side surfaces, to the top surface, and to the bottom surface.²³⁹

Comparison of the *Stone Strong* injunction with claim 1 reveals a number of differences. For example, claim 1 includes limiting language, such as the requirement of an “alignment device,” that the injunction facially lacks. On the other hand, the injunction’s requirement of a specific relationship between the width of a block’s “recess or notch” and other dimensions is missing from claim 1’s explicit language.

Such differences might reflect a deliberate trade-off. The dimensional constraints of the injunction might be viewed as a way of capturing, in comparatively unambiguous terms, at least a subset of situations in which the “lifting device” in combination with the “recess or notch” will tend to operate as an “alignment device.”

Stone Strong’s dimensional language thus demonstrates how a Type-B, reformulated-bounds injunction can effectively replace claim language with substitute language that at least arguably increases the clarity of an injunction’s scope. Such increased clarity might be advantageous for both parties as well as the courts. Gains from increased clarity can compensate a party for the broader or narrower scope of the injunction relative to that of a

237. *Id.* (emphasis omitted).

238. Compare *id.*, with U.S. Patent No. 7,073,304 col. 17 I. 35 to col. 20 I. 65 (filed Sept. 23, 2004) (listing claims); U.S. Patent No. 6,796,098 col. 13 I. 41 to col. 16 I. 43 (filed Feb. 27, 2003) (same).

239. ’098 Patent, col. 13 II. 41–57.

conventional Type-1, colorable-differences order. Indeed, the fact that *ExitExchange* and *Vertical Doors, Inc. v. Howitt*, in contrast with *Stone Strong*, both involve consent decrees appears to confirm that parties on opposing sides of an injunction can agree to prefer the apparently clearer boundaries of a Type-B injunction to a Type-1 injunction's "no more than colorable differences" haziness.²⁴⁰

4. *Moderated Injunctions*.—A fourth type of specially tailored injunction is what this Article terms a "moderated injunction"—an injunction that includes an explicit carve out for infringing (or likely infringing) behavior. At least five orders in the 2010 dataset exhibit such moderated terms. An order restricting the use of tarpless fumigation includes a carve out for "anyone . . . walk[ing] by a tarpless fumigation wearing a detection device solely for the purpose of personal safety."²⁴¹ Another order directed to the recording of medical data specifically limits the order's prohibition to use "in any clinical applications in the United States," thereby providing a carve out for nonclinical uses.²⁴² A temporary restraining order forbidding infringing activities such as use, offers to sell, and importation specifically provides that "this restraint does not apply to manufacture."²⁴³ Somewhat similarly, two injunctions involving pharmaceuticals specifically limit their prohibitions of making the patented invention to prohibitions of "*commercial* manufacture."²⁴⁴ With respect to these last two injunctions, Congress itself might be credited as the cause of moderation: the courts issued these injunctions in accordance with special statutory language that provides a limited authorization for injunctions to "prevent the *commercial* manufacture, use,

240. Parties routinely agree on contract terms that set bounds for authorized or unauthorized behavior that do not precisely align with intellectual property rights. See Jeanne C. Fromer, *Claiming Intellectual Property*, 76 U. CHI. L. REV. 719, 749–50 (2009) (describing such practice by copyright owners); cf. Mark R. Patterson, *Contractual Expansion of the Scope of Patent Infringement Through Field-of-Use Licensing*, 49 WM. & MARY L. REV. 157, 160 (2007) ("[P]atentees have used field-of-use licensing to prevent purchasers of patented products from repairing those products, and to enforce distribution restraints."). Consequently, reflection of such behavior in consent decrees that result from agreements between parties should probably not be surprising.

241. *Structural Tenting Corp. v. Termite Doctor, LLC*, No. 1:09-cv-21285-MGC, slip op. at 2 (S.D. Fla. Aug. 17, 2010).

242. *DatCard Sys., Inc. v. Codonics, Inc.*, No. 8:08-cv-00063-AHS-RNB, slip op. at 1 (C.D. Cal. Jan. 21, 2010).

243. *King Pharms., Inc. v. Sandoz, Inc.*, No. 3:08-cv-05974-GEB-DEA, slip op. at 2 (D.N.J. Apr. 6, 2010).

244. *Hoffman-La Roche Inc. v. Cobalt Pharms. Inc.*, No. 2:07-cv-04539-SRC-MAS, slip op. at 2 (D.N.J. Nov. 10, 2010) (emphasis added) ("preliminarily enjoin[ing] . . . engag[ement] in the commercial manufacture, use, offer for sale, or sale within the United States of any products that are the subject of Cobalt's Abbreviated New Drug Applications Nos. 79-002 and 79-003"); *Sandoz Inc. v. Boehringer Ingelheim Int'l GMBH*, No. 3:10-cv-00437-TJC-JBT, slip op. at 1–2 (M.D. Fla. July 29, 2010) (emphasis added) (prohibiting "engag[ement] in the commercial manufacture, use, offer to sell, or sale within the United States, or importation into the United States, of any drug product containing pramipexole or a pharmaceutically acceptable salt thereof").

offer to sell, or sale within the United States or importation into the United States of an approved drug or veterinary biological product."²⁴⁵

V. Crafting the Optimal Injunction

We now come to a basic normative question: How, from a policy standpoint, does one determine the optimal scope and form of a patent-infringement injunction? Response to this question is complicated by the fact that a diligent policy maker must balance a host of concerns, including concerns relating to likely effects of injunctions on adverse parties' behaviors or to courts' limitations in issuing and enforcing injunctive relief. Most fundamentally, questions of patent-infringement injunction scope raise concerns of overdeterrence and underdeterrence that are pervasive in law. Part II has indicated how a narrow injunction can leave a rational infringer with substantial reason to pursue a course of action that, though unlikely to constitute contempt, will likely be determined to constitute infringement. On the other hand, a broad injunction can combine with the potential severity of contempt sanctions to deter future activity that is unlikely to result in a finding of either infringement or contempt. The latter possibility can be of special concern if, as assumed for purposes of simplicity here,²⁴⁶ society rationally wishes to allow or encourage design-arounds even where a design-around only barely avoids infringement.²⁴⁷

One could argue, however, that with respect to an adjudged infringer, concerns with overdeterrence are generally ill founded. A judgment of infringement means that a patent claim has been held not invalid and not unenforceable despite whatever challenges an adjudged infringer chose to make. The claim's scope has been clarified at least to the extent necessary to support a holding of infringement. Moreover, the adjudged infringer cannot justifiably claim that it remains unaware of the patent or its potential relevance to the infringer's activities.²⁴⁸ Thus, various concerns of notice, clarity, and uncertainty²⁴⁹ have been addressed. Especially if an infringer's

245. 35 U.S.C. § 271(e)(4)(B) (2006) (emphasis added); *see also supra* note 189.

246. *See supra* text accompanying notes 30–36.

247. Likewise, for purposes of simplicity, this Article assumes, at least as a general matter, that society rationally wishes to discourage actually infringing activity. Of course, if one believes that patent rights are systematically overbroad, one might question this assumption and seek narrower or otherwise weaker injunctions without vexing about any felt need to balance lost protection for patentees against the gains from reduced chilling and "taxation" of infringing or, at least, potentially infringing activity.

248. *Cf.* Omri Ben-Shahar, *Damages for Unlicensed Use* 23 (Univ. of Chi. John M. Olin Law & Econ., Working Paper No. 534, 2010), available at <http://ssrn.com/abstract=1677667> (indicating an argument that "already-licensed parties" should "face harsher remedies for unlicensed use than the ones strangers face" because already-licensed parties can be expected, generally speaking, to have lower transaction costs for additional licensing).

249. Herbert Hovenkamp, Response, *Notice and Patent Remedies*, 88 TEXAS L. REV. SEE ALSO 221, 224 (2011) (emphasizing the importance of notice concerns in relation to remedies for patent

past conduct suggests a substantial probability of future violations, why not follow foreign common law jurisdictions and favor issuance of a patent-limited Type-2, obey-the-law order—i.e., an injunction against further infringement of the specific patents or patent claims already adjudged to have been infringed?

One reason why not is that under current law and practice, notice of the existence and potential relevance of a specific patent or patent claim does not equate to clear notice of the scope of associated patent rights.²⁵⁰ Consequently, a patent-limited Type-2, obey-the-law order tends to be less instructive than an order not to trespass further on Blackacre. An individual patent claim can encompass a great variety of dissimilar-looking embodiments of the claimed invention. Moreover, claim limitations are frequently less than crystal clear. Even seemingly simple questions of patent scope can be surprisingly difficult to answer. In an actual suit alleging infringement of a Gillette patent by the four-blade Schick Quattro, much initial argument centered on an apparently simple question—whether Gillette's patent could cover a razor having more than three blades as well as a razor having three and only three blades.²⁵¹ The distribution of judicial “votes” on the issue indicates that the question was far from easy, at least at the preliminary-injunction stage: the district court judge and one Federal Circuit judge believed that Gillette's patent could not cover a four-blade razor, but two Federal Circuit judges thought otherwise.²⁵²

Even if prior litigation has clarified—and perhaps even nailed down—the scope of patent claims along a number of dimensions, not all dimensions of a claim will necessarily have been addressed. The selection of claim terms subjected to judicial interpretation will reflect particular characteristics of accused matter or the specific nature of challenges to validity or enforceability. Unaddressed questions of claim validity and scope can generate uncertainty comparable to that which existed before prior litigation. Hence, for example, if Energizer Holdings, the defendant in the Gillette suit, had previously produced a razor with exactly three blades, prior infringement litigation relating to that three-blade razor might have done nothing to clarify whether the Gillette patent could cover a razor with more than three blades as

infringement); see also BESSEN & MEURER, *supra* note 189, at 46 (“A successful property system establishes clear, easily determined rights.”).

250. See Dan L. Burk & Mark A. Lemley, *Fence Posts or Sign Posts? Rethinking Patent Claim Construction*, 157 U. PA. L. REV. 1743, 1748 (2009) (“Those who are intimate with the patent system have long understood that it is simply impossible to define boundaries of invention with the physical or descriptive precision of defining the boundaries of real property.”).

251. *Gillette Co. v. Energizer Holdings, Inc.*, 405 F.3d 1367, 1368 (Fed. Cir. 2005) (“The district court denied Gillette's motion for a preliminary injunction because it found that the claims . . . covered only a three-bladed razor . . .”).

252. See *id.* (holding that “the district court erred in construing the claims”); *id.* at 1382 (Archer, J., dissenting) (“[T]he specification makes abundantly clear that the invention . . . was a razor having three blades, no more . . .”).

such a question would most likely have been irrelevant to whether the razor at hand infringed.

Thus, to the extent patent law policy makers wish to ensure that market actors, including adjudged infringers, feel free to develop and disseminate innovations whose infringement or noninfringement of another's patent rights is uncertain, those policy makers should worry about the possibility of overdeterrent injunctions. The potential severity of contempt sanctions and uncertainty about the scope of an injunction that, in effect as well as by its terms, forbids future infringement of a patent or patent claim could deter good-faith, socially productive activity in which an infringer would have otherwise engaged.²⁵³ To limit the likelihood or severity of overdeterrence, courts might wisely seek to ensure that, generally speaking, patent-infringement injunctions do not stray too far from the immediate environs of matter already adjudged to infringe—matter for which relevant questions of patent-claim scope have been addressed. In most cases, the haziness of a resulting Type-1 order might not be much of a concern because the onward march of technology or market interest, agreement between the parties, or other circumstances make activity that both tests the boundaries of an injunction and substantially upsets the patentee relatively unlikely. Thus, Type-1, colorable-differences injunctions arguably represent an appropriate default, one that protects core social interests while saving on the costs of providing an exact, *ex ante* specification of what society means to protect.²⁵⁴

An advocate for Type-2, obey-the-law injunctions might counter that an adjudged infringer is not without recourse to limit uncertainty. If the infringer is in doubt about whether later contemplated activity is at risk, the infringer can petition for clarification or modification of the original order.²⁵⁵ Alternatively, the infringer can contract with the patentee for a blanket license for activities that might otherwise violate the patentee's previously infringed patents.

But such solutions are neither always feasible nor always socially desirable. Whether because of information costs, developed distrust between parties, strategic behavior, or conditions that provide a basis for "holdup" or

253. *Contra* Earl C. Dudley, Jr., *Getting Beyond the Civil/Criminal Distinction: A New Approach to the Regulation of Indirect Contempts*, 79 VA. L. REV. 1025, 1025–27 (1993) (describing exercises of contempt power as lacking meaningful constraint).

254. *Cf.* Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557, 569 (1992) (observing that where "there is a given cost of determining the appropriate content of the law *ex ante*," "rules are more expensive to promulgate than standards").

255. *See* FED. R. CIV. P. 60(b)(5) ("On motion and just terms, the court may relieve a party or its legal representative from a final judgment, order, or proceeding for the . . . [reason that] . . . applying it prospectively is no longer equitable . . ."); *see also* *Smith Corona Corp. v. Pelikan, Inc.*, 784 F. Supp. 452, 486 (M.D. Tenn. 1992) ("Pelikan's motion to clarify injunction is granted . . ."); *RENDELMAN, supra* note 79, at 491 ("The Supreme Court's decision that a Chancery court has inherent power to modify or dissolve an injunction came in 1932 Rule 60(b)(5) became effective . . . six years later.").

“holdout,”²⁵⁶ reasonable licenses do not always occur. Uncertainty about an injunction’s scope can exacerbate difficulties in coming to a satisfactory agreement, and the cost of clarifying that scope through new judicial process might be unjustifiably great for resource-strapped courts and private parties alike. In short, possibilities for private contracting or judicial clarification dilute but do not eliminate bases for believing that Type-1, colorable-differences injunctions provide a sensible default.

Moreover, there is good reason for demanding that the Type-1 *effective scope* of a Type-2, obey-the-law injunction be indicated on the face of the order itself, rather than be imposed through later, narrowing interpretation or application. The general language of prohibition characteristic of Type-2 orders might encourage unsophisticated parties to believe that the effective scope of an injunction is broader than it is. Even if a relatively sophisticated party were informed that it is a “legal slam dunk” that obey-the-law language’s effective scope will be narrower than its plain meaning suggests, that party might justifiably hesitate to rely on such information. Slam dunks are sometimes missed, and private parties might be used to having to discount the certainty of legal representations.

Further, broad obey-the-law language might invite error by a court itself. Such language might encourage a district court to believe that valid grounds for finding contempt extend substantially beyond activities explicitly prohibited by a Type-1, colorable-differences order. Even assuming that a district court later recognizes that there are Type-1 limitations on grounds for finding contempt, Type-2, obey-the-law language might subtly distort how a district court understands those limitations. In the shadow of Type-2 language, limitation of an order’s effective scope might seem more a response to concerns with limiting false positives—improper holdings of contempt—than a response to Rule 65(d) concerns with notice. A court that has persuaded itself that an enjoined party’s new course of action is infringing might be quicker to discount worries about false positives than concerns about adequate notice. The latter concerns more firmly require a court to consider how things looked to the infringer when the injunction issued, not merely how things look to the court later.

So the case for a default rule in favor of Type-1, colorable-differences injunctions seems fairly strong. Nonetheless, Type-1 injunctions fall far short of providing a complete answer to problems of injunction scope. Most obviously, they fall short because their “no more than merely colorable differences” language explicitly incorporates a continuing reason for uncertainty: the question of whether a difference is merely colorable. Even if

256. Cf. Lemley & Shapiro, *supra* note 2, at 1993 (“Injunction threats often involve a strong element of *holdup* in the common circumstance in which the defendant has already invested heavily to design, manufacture, market, and sell the product with the allegedly infringing feature.”). See generally Calabresi & Melamed, *supra* note 4, at 1107 (describing “moving from a property rule to a liability rule” as a potential solution to a “holdout problem”).

this phrase could be precisely defined in the abstract, its meaning in actual practice would often be unclear. This follows from the fact that determination of what is merely colorable commonly requires reference to associated patent claims. But as discussed above, patent claims themselves often contain latent ambiguity.²⁵⁷ Would addition of a fourth blade to a previously infringing three-blade razor constitute a colorable change? The answer presumably depends on whether one believes that the fourth blade is somehow relevant to the patent claims at issue. Resulting uncertainty about what constitutes a “merely colorable difference” can leave an uncomfortable degree of uncertainty about a Type-1 injunction’s scope.

What alternatives might enable a court to avoid such uncertainty? A court might decline to issue any injunction at all. Although this might seem so flawed a response that it is scarcely worth mentioning, the district court in *MercExchange, L.L.C. v. eBay, Inc.*²⁵⁸ seems originally to have followed this course.²⁵⁹ A substantial cause for the district court’s initial denial of an injunction was its fear that such an order would simply be no more than a prelude to “contempt hearing after contempt hearing.”²⁶⁰

A problem with the no-injunction solution to fears of later contempt proceedings is that this solution seems disproportionately likely to lead to denials of injunctions in precisely those situations where an injunction is most needed: situations where an adjudged infringer will foreseeably engage in activity likely to raise related concerns of infringement. Even if circumstances exist where an injunction will lead to more, rather than less, socially wasteful litigation, there is an externality concern with respect to correct judicial identification of those circumstances. Denial of an injunction will likely shift the burden of further foreseeable litigation to another judge, whereas a grant of an injunction will likely require the granting judge to shoulder at least a portion of that burden in a later contempt proceeding. Thus, to the extent a judge wishes to limit his or her involvement in vexatious proceedings, a judge trying to decide whether foreseeable future litigation justifies denying rather than granting an injunction might have self-interested reasons to favor denial over a grant.

257. See *supra* text accompanying notes 250–52.

258. 275 F. Supp. 2d 695 (E.D. Va. 2003), *aff’d in part, rev’d in part, vacated in part*, 401 F.3d 1323 (Fed. Cir. 2005), *vacated*, 547 U.S. 388 (2006).

259. *Id.* at 714.

260. *Id.* More fully, the district judge stated,

If the court did enjoin the defendants here, the court would essentially be opening a Pandora’s box of new problems. This case has been one of the more, if not the most, contentious cases that this court has ever presided over. . . . The court predicts that if it granted the plaintiff’s request for a permanent injunction, the battle would continue to be as contentious as ever. . . . The court envisions contempt hearing after contempt hearing. . . . This will result in extraordinary costs to the parties, as well as considerable judicial resources.

Id.

Compared to denying an injunction altogether, a less extreme alternative is to consider how an order might be narrowed or otherwise crafted to avoid the problems that injunctive relief might present. Courts sometimes will delay the full effectiveness of injunctions to avoid some of the special disruption or other hardship that an immediately effective order might cause.²⁶¹ But in deciding whether to issue an injunction and when to make it effective, courts should perhaps more often consider how the specific scope of an injunction can affect the balance of concerns that favor or disfavor injunctive relief. Taking scope into consideration in deciding on grant or denial might require district courts to alter existing practices—or at least to more consistently follow best practices. District courts have perhaps too often decided whether to issue an injunction without substantially considering the details of the potential injunction's scope, indeed even combining the decision to issue an injunction with a call for a subsequent proposal for the injunction's form.²⁶² Once one recognizes that patent-infringement injunctions can come in different shapes and sizes, it seems logical to quite uniformly demand the provision of a draft order in advance of deciding whether an injunction should issue and, in making the latter

261. See *Broadcom Corp. v. Qualcomm Inc.*, 543 F.3d 683, 704 (Fed. Cir. 2008) (upholding a district court's order granting an injunction in part because a "sunset provision" permitting infringement to continue for twenty months after the jury verdict alleviated hardship to the adjudged infringer); *Metso Minerals, Inc. v. Powerscreen Int'l Distribution Ltd.*, 788 F. Supp. 2d 71, 77 (E.D.N.Y. 2011) (holding that, in light of the defendants' representation "that they will have completed the redesign of their screener line by the end of June 2011," the balance of hardships and considerations of public interest made "appropriate . . . delay[ing] the implementation of injunctive relief until July 11, 2011"); *B. Braun Melsungen AG v. Terumo Med. Corp.*, 778 F. Supp. 2d 506, 524 (D. Del. 2011) (ruling in favor of the adjudged infringer's proposal of "a 'sunset' period of fifteen months, allowing [the infringer] to continue to sell the [infringing device] in [a] market segment where it is presently sold").

262. See, e.g., *QBAS Co. v. C Walters Intercoastal Corp.*, No. 8:10-cv-00406-AG-MLG, slip op. at 23 (C.D. Cal. Dec. 16, 2010) (deciding to grant a preliminary injunction but deciding to "determine an appropriate bond amount [only] after Plaintiffs file a proposed preliminary injunction order"); *Stone Strong, LLC v. Del Zotto Prods. of Fla., Inc.*, No. 5:08-cv-503-oc-10DAB, slip op. at 8 (M.D. Fla. Oct. 25, 2010) (holding that "the Plaintiff is entitled to injunctive relief" and giving the plaintiff "[ten] days within which to file and submit a proposed, final injunctive decree"); *Mytee Prods., Inc. v. Harris Research, Inc.*, No. 3:06-cv-01854-CAB, slip op. at 7 (S.D. Cal. Jan. 20, 2010) (granting a motion for permanent injunction and calling for subsequent submission of "[a] proposed permanent injunction"); *Flexiteek Ams., Inc. v. Plasteak, Inc.*, No. 0:08-cv-60996-JIC, slip op. at 18–19 (S.D. Fla. Sept. 15, 2009) (concluding that an injunction should issue and ordering plaintiffs to file a proposed injunction). *But see Metso Minerals*, 788 F. Supp. 2d at 77 ("In addition to their arguments opposing injunctive relief in whole, the defendants also oppose a number of specific aspects of the plaintiff's proposed injunction."); *B. Braun Melsungen*, 778 F. Supp. 2d at 525–26 (reviewing the parties' proposals for the scope of injunctive relief and granting a permanent injunction "only to the extent of entering an injunction on the terms proposed by [the defendant]"). After deciding to grant an injunction, a district court might call for a specific proposal for the form of the on an injunction even though a proposed order has previously been submitted. Compare *Mytee Prods.*, slip op. at 7 (deciding to issue an injunction and calling for subsequent submission of "[a] proposed permanent injunction"), with *Mytee Prods., Inc. v. Harris Research, Inc.*, No. 3:06-cv-01854-CAB (S.D. Cal. Nov. 20, 2009) (proposed permanent injunction) (presenting a draft order prohibiting infringement of various patent claims).

decision, to consider more carefully the proposed scope of a specific draft order as well as, possibly, different specific scopes that this draft order might be amended to have.²⁶³ Attempting to evaluate the balance of hardships and public interest concerns associated with an injunction without substantial reference to a specific proposal for the form of an injunction might be ill considered in cases where tailoring an injunction might make a difference.

How might an injunction be tailored? An obvious first possibility would be to limit its scope to that of a “true” Type-0, measure-zero order. Such an order would specify that it is not to be enforced against anything but the exact products or processes already held to infringe.

As already discussed, however, the true Type-0 alternative is often likely to be not much of an alternative at all.²⁶⁴ Outside relatively idiosyncratic fields such as pharmaceuticals, there frequently is a virtually limitless pool of minor variations that can distinguish new products or processes without significant change in functionality. Under such circumstances, a true Type-0 order is likely to be essentially worthless. Permitting the possibility of contempt to be averted by, for example, offering three-blade razors that are gray rather than black could make the whole process of awarding injunctive relief a fundamentally empty gesture.

Specially tailored injunctions provide courts with a potential way to escape the Type-0 versus Type-1 dilemma. By crafting injunction-specific language, a court can avoid the haziness of a Type-1 injunction’s “no more than colorable differences” language while also providing relief that forbids more than a measure-zero range of conduct. As seen with the dimensional limitations in the *Stone Strong* injunction,²⁶⁵ a successfully drafted specially tailored injunction can possess comparatively clear limits that provide safe havens for a broad range of potential future activities. These havens can reduce the possibility that a combination of uncertainty and infringer risk aversion will cause an injunction’s deterrent effect to overshoot its mark.

For those worried that specially tailored orders will tend to favor patentee interests at excessive expense to society, it bears emphasizing that a specially tailored injunction can be broader, narrower, or simultaneously broader and narrower than alternatives such as a Type-1, colorable-differences injunction. An order to destroy certain existing articles or to label certain future articles as “Not to Be Sold in the United States” can both require more and provide less than patent rights abstractly demand. By generating a special set of metes and bounds for purposes of injunctive relief, a Type-B, reformulated-bounds injunction can provide an independent

263. *Cf.* ION, Inc. v. Sercel, Inc., No. 5:06-CV-236-DF, slip op. at 14 (E.D. Tex. Sept. 16, 2010) (agreeing with objections to the scope of “ION’s proposed injunction” and observing that “exempting customers that have already purchased infringing sensors from the scope of the injunction mitigates the adverse practical and economic effects”).

264. *See supra* text accompanying notes 85–86.

265. *See supra* text accompanying notes 235–40.

description of forbidden products or processes that is simultaneously extraproductive and subprotective relative to a Type-1 injunction.²⁶⁶ If these gains or losses in breadth are appropriately supplemented by increased clarity, the resulting order can be easier for a court to enforce and perhaps even preferable for all parties. One way in which a court in the position of the *eBay* district court can try to forestall contentious contempt proceedings is to devise a specially tailored injunction that protects vital patentee interests while also placing relatively clear limits on what the adjudged infringer must or must not do.

Indeed, there are at least two types of circumstances in which, despite generally higher drafting costs, specially tailored injunctions seem particularly likely to be preferable to Type-1 do-not-infringe orders. Roughly speaking, these circumstances involve situations in which the uncertain bounds of Type-1 orders can yield substantial overdeterrence or underdeterrence.²⁶⁷

The first type of situation, like that in *eBay*, is one in which an adjudged infringer wishes to pursue one or more design-arounds that are apparently both (a) questionable enough legally and (b) significant enough for the patentee that their pursuit will likely trigger contempt proceedings. To the extent society wishes to protect against relitigation of an almost unaltered course of infringement but also to provide relative safe harbors for design-around activity that do not require a petition for prior approval, a specially tailored injunction might be desirable.

More particularly, if we assume that compared to a Type-1, colorable-differences order, a specially tailored order will provide clearer instruction that helps reduce later attorney advice or litigation costs and that also helps generate increased compliance with issued injunctions, we might model the costs and benefits of a specially tailored order as follows. Let D represent the extra drafting cost for a specially tailored order relative to a Type-1 order; S represent the savings in later attorney advice or litigation costs that the specially tailored order is expected to generate; G represent the expected value, outside of any litigation or attorney advice savings, of any increase in legal compliance—whether by the adjudged infringer or by other parties²⁶⁸—

266. See *supra* text accompanying notes 225–40.

267. Cf. Kaplow, *supra* note 254, at 576 n.42 (“If, for example, there is uncertainty concerning what an adjudicator would deem to be due care, there may be a tendency . . . to take care that is excessive relative to the expected due care requirement, although it is also possible that individuals would take less care than the expected due care requirement.”).

268. A specially tailored order might help promote increased legal compliance by third parties if, for example, violation or arguable violation of Type-1 orders encourages disrespect of patent rights or court orders more generally. Cf. Edward K. Cheng, *Structural Laws and the Puzzle of Regulating Behavior*, 100 NW. U. L. REV. 655, 656–57 (2006) (suggesting that a symptom of “excessive use of criminalization” is “casual lawbreaking by ordinary citizens [that] promotes an unhealthy disrespect for the law”); Joseph P. Liu, *Owning Digital Copies: Copyright Law and the Incidents of Copy Ownership*, 42 WM. & MARY L. REV. 1245, 1307 (2001) (“[A] significant divergence between norms of usage and legal requirements may lead to a disrespect for the law.”).

that is associated with the specially tailored order; and A represent the expected value (positive or negative) of the decrease or increase in design-around activity predicted to result from replacement of a Type-1 injunction with the specially tailored order. Then, assuming these terms together reflect all relevant social effects of replacement of a Type-1 order with a specially tailored order, we find that the difference in social value between the specially tailored and Type-1 orders, Δ_{ST-T1} , has the following formula:

$$\Delta_{ST-T1} = (S - D) + G + A.$$

In accordance with this formula, the *eBay*-type situation might be understood as one where the magnitude of the legal savings, S , is likely to be substantial if the specially tailored order will head off any need for contempt proceedings and even more substantial if the specially tailored order, perhaps because its relevant scope is clear, limits even the need for attorney advice in anticipation of contempt proceedings. Meanwhile, if the patent system is assumed to be far from totally dysfunctional and if, as implicitly posited, the value of reducing the chilling of likely design-around activity is positive, a specially tailored order that increases the probability of compliance and reduces chilling might reasonably be expected to correspond to a non-negative value for the sum $G + A$. Under such circumstances, nontrivial drafting costs could be incurred without sacrificing a necessary condition for the social desirability of special tailoring, $\Delta_{ST-T1} > 0$.

A second type of situation where a specially tailored injunction might be thought desirable is one where there is real concern that a Type-1, colorable-differences order will be inadequate to protect patentee rights. In other words, the second type of situation is one where the principal concern is not the potential for undue chilling of socially desirable design-around activity but instead the potential for socially undesirable erosion of the value of patent rights.

Situations in which a previously enjoined party has been found in contempt might be viewed as falling in this class both because violation of a prior Type-1 order has proven that order to have insufficient deterrent effect and because the district court has had a chance to assess whether circumstances, exceptional or otherwise, justify or excuse the enjoined party's continuing unlawful activity. Moreover, where violation of patent rights and a prior injunction appears recalcitrant, the expected gain from increased compliance might in part be a systemic gain from a more general signal to third parties of courts' willingness and capacity to ensure that patent

The value of increased legal compliance might generally be thought likely to be positive although, under a dysfunctional or even merely imperfect legal regime, this is not necessarily always so. Cf. RONALD DWORKIN, A MATTER OF PRINCIPLE 106 (1985) ("Almost everyone will agree that if a particular decision is very wicked, people should disobey it." (emphasis omitted)); Hovenkamp, *supra* note 249, at 223 ("As a result of overly aggressive substantive rules, many antitrust violations are also economically efficient."); Eric A. Posner & Alan O. Sykes, *Efficient Breach of International Law: Optimal Remedies, "Legalized Noncompliance," and Related Issues*, 110 MICH. L. REV. 243, 287 (2011) (positing "examples of 'efficient breach' in international law").

rights are ultimately respected. In other words, the value of G in such situations might be unusually large—larger than the stakes of individual parties directly affected might alone suggest. Thus, even a sizable negative value for A might not prevent a broad, specially tailored injunction from being desirable. Further, willingness to sweep broadly might well keep drafting costs, D , relatively low even while keeping the clarity of the specially tailored order and thus attorney advice or litigation savings, S , relatively high. The result might be a well-justified expectation that Δ_{ST-T1} for a broad, prophylactic order will be positive. Such a conclusion accords with the Federal Circuit's recognition of the validity of broad, reformulated-bounds injunctions issued in response to violation of a prior injunctive order.²⁶⁹

To the extent a court worries that a specially tailored injunction will nonetheless prove overreaching, the court can take corrective steps. First, the court can include a sunset provision—a specific time limit on the injunction's effectiveness absent further court action.²⁷⁰ The court can also signal a special willingness to reconsider the injunction's scope in light of changed circumstances or new information.²⁷¹ Finally, to the extent injunction overreach is a concern, a district court can opt for a narrower, rather than broader, specially tailored injunction, perhaps in combination with a Type-1, colorable-differences injunction that the court plans to construe narrowly. Such a hybrid approach might accord well with equity's traditional use of anti-opportunism safety valves—here in the shape of a somewhat fuzzy Type-1 order—to complement at least aspirationally clearer, *ex ante* rules that might, by themselves, be too easy for an ill-intentioned actor to circumvent.²⁷²

Appellate judges can also take steps to ensure that district courts use specially tailored orders properly. If there is legitimate cause for concern that district courts might issue such orders too readily or rashly, appellate courts might seek to limit the practice by demanding that both the use of a specially tailored injunction and its scope be provided with special degrees of

269. See *supra* text accompanying notes 151–54.

270. See LAYCOCK, *supra* note 78, at 287 (discussing a six-month limit on an injunction against a former employee joining a competitor).

271. Cf. FED. R. CIV. P. 60(b)(5) (“On motion and just terms, the court may relieve a party or its legal representative from a final judgment, order, or proceeding for the . . . [reason that] . . . applying it prospectively is no longer equitable . . .”); SHARPE, *supra* note 203, ¶ 1.450, at 1-18 to 1-19 (“If necessary, the court can make an order by which it implicitly undertakes to review the circumstances and the obligation imposed as matters proceed.”). See generally DAVID I. LEVINE ET AL., REMEDIES: PUBLIC AND PRIVATE 244 (5th ed. 2009) (discussing “[t]he uncertainty of when to apply [a] rigid standard . . . to modifications of injunctions and consent decrees and when a more flexible standard is appropriate”).

272. See Henry E. Smith, An Economic Analysis of Law Versus Equity 39 (Mar. 10, 2010) (unpublished manuscript), available at <http://extranet.isnie.org/uploads/isnie2010/smith.pdf> (“Equity applies in a smaller domain [than law] with an eye to deterring opportunism, but where it applies it is vague and *ex post*.”); cf. Kaplow, *supra* note 254, at 618 (observing that fraud “may be easier to commit if there are known rigid rules that a fraudulent actor can carefully circumvent”).

explicit justification.²⁷³ In accordance with such a requirement, district courts issuing specially tailored relief and parties seeking such relief would need to expend an unusual degree of effort. The need for this unusual level of effort should itself act as a brake on excessive use of specially tailored injunctions, which, compared to a Type-1 default injunction, are already independently likely to entail higher “promulgation costs” for parties and district courts.²⁷⁴ Moreover, the resulting generation of a more complete trial record should better enable the Federal Circuit to use abuse-of-discretion review to keep the practice within reasonable bounds.²⁷⁵

Indeed, judicious combination of specially tailored injunctions with Type-1, colorable-differences injunctions might have the added benefit of fostering a more principled and predictable jurisprudence on the scope of the latter, while also protecting against a subsequently discovered, opportunistic means of avoiding a specially tailored injunction’s force.²⁷⁶ Inclusion of a specially tailored order such as a destruction order could provide greater assurance that the patentee has gotten *some* prospective advantage from prior successful litigation, even if the court later denies a motion for contempt and thus requires the patentee to launch a wholly new suit against a previously adjudged infringer. Judges might thus be better insulated from fears that without a relatively broad understanding of “no more than colorable differences,” their orders are too easily circumvented.

Still, can we really expect courts to craft specially tailored injunctions that are “no more burdensome . . . than necessary to provide complete relief”?²⁷⁷ As with any practically applicable rule or standard, there is undoubtedly risk that a specially tailored injunction will either overreach or underreach relative to what is necessary to achieve optimal enforcement of a patentee’s rights.²⁷⁸ But as illustrated by the specially tailored patent-

273. See *infra* text accompanying notes 291–94.

274. Cf. Kaplow, *supra* note 254, at 572 (“The difference in [legal] promulgation costs favors standards, whereas that in enforcement costs favors rules.”).

275. Cf. Steven Shavell, *Optimal Discretion in the Application of Rules*, 9 AM. L. & ECON. REV. 175, 178 (2007) (describing how an appeals process can “induce[] decisions to conform to the socially desirable, at least within the range governed by the cost of an appeal”).

276. Cf. Mark Gergen, John M. Golden & Henry E. Smith, *The Supreme Court’s Accidental Revolution? The Test for Permanent Injunctions*, 112 COLUM. L. REV. 203, 237 (2012) (“A major theme in equity has been the need to correct for party opportunism, and injunctions partake of this overarching purpose.” (footnote omitted)).

277. *Califano v. Yamasaki*, 442 U.S. 682, 702 (1979) (rejecting an argument that “nationwide class relief [was] inconsistent with the rule that injunctive relief should be no more burdensome to the defendant than necessary to provide complete relief to the plaintiffs”); see also *Madsen v. Women’s Health Ctr., Inc.*, 512 U.S. 753, 765 (1994) (describing *Califano* as stating a “general rule”); cf. *Dayton Bd. of Educ. v. Brinkman*, 433 U.S. 406, 418 (1977) (“In effect, the Court of Appeals imposed a remedy which we think is entirely out of proportion to the constitutional violations found by the District Court . . .”).

278. See Kaplow, *supra* note 254, at 594 (indicating the likelihood of imperfect tailoring, as a practical matter, under any form of legal doctrine by characterizing “imagin[ing] countless factors (arguments) that a decisionmaker might take into account” under a sufficiently inclusive standard as “a romantic perspective, hardly a valid depiction of actual decisionmaking”).

infringement injunctions issued by district courts in 2010,²⁷⁹ the forms of specially tailored injunctions generally contemplated here seem likely to remain reasonably tightly tied to both the actual offense and the underlying rights at issue: these are not the sort of broadly ambitious institutional-reform or structural injunctions whose issuance or nonissuance has a greater tendency to spark heated contentions of judicial usurpation or abdication. Indeed, the risks of overprotection or underprotection through the forms of specially tailored injunctions contemplated here seem to be relatively comparable to those present with respect to conventional Type-1 injunctions. The latter's hazy "no more than colorable differences" language leads not only to immediate uncertainty but also to the possibility of an undesirably broad or narrow reading in later contempt proceedings.

Problems in choosing injunction form thus bear substantial analogy to problems in deciding the extent to which patent scope should be determined through central claims, peripheral claims, or some combination of the two. Like Type-1, colorable-differences injunctions, central claims describe embodiments to which infringing matter needs to be substantially related.²⁸⁰ Like some specially tailored injunctions, peripheral claims seek to use language to mark the perimeter of matter that legal entitlements encompass.²⁸¹

The analogy is imperfect, however, because the circumstances in which the relevant delimiting language is developed and applied differ substantially. The case for specially tailored injunctions might be thought generally stronger than that for peripheral claiming because of an arguably greater need for precisely delimited scope to avoid overdeterrence from a threat of contempt and also because judicial gatekeepers for specially tailored injunctions have substantial advantages over patent examiners responsible for allowing peripheral claims. Specially tailored injunctions are directed at particular parties who have already engaged in specific forms of behavior that the court, after claim construction and resolution of any challenges to validity and enforceability, has adjudged to be infringing.²⁸² Except in default situations, both the patentee and the adjudged infringer are before the court and able to provide information relating to their proximate future interests. Thus, the court, in contrast to a patent examiner, might relatively cheaply be able to determine the nature of potentially infringing conduct that is likely to be of future concern. In contrast, initial examination of patent

279. See *supra* notes 213–44 and accompanying text.

280. See Golden, *supra* note 20, at 348–49 (“In central claiming, claims describe or point to representative embodiments of the inventive idea.”).

281. See *id.* at 349 (“In peripheral claiming, claims indicate the literal boundaries of patent rights.”).

282. Cf. RENDLEMAN, *supra* note 79, at 424 (observing that relative to “a statute, a rule, a regulation, a will, and a contract . . . , an injunction can be more specific because it is drafted to control the behavior of a known defendant engaged in an actual lawsuit”).

rights generally occurs without involvement of any potential infringer.²⁸³ Indeed, such an infringer's relevant products or processes might not yet even be contemplated. Relative to an examiner, a court is thus likely to have a significant cost advantage in seeking to provide a relatively clear, *ex ante* description of prohibited behavior.²⁸⁴

Further, with the threat of infringement by a particular adjudged infringer already an established fact, the benefits of such *ex ante* specification might be expected to be greater and to have less need for time discounting.²⁸⁵ Given a commonly observed oligopolistic structure for established technology markets²⁸⁶ and a potential patentee proclivity for pursuing industry players having significant market share,²⁸⁷ the adjudged infringer might well represent a significant share of a valuable commercial market, with the result that an injunction against the infringer might have significant implications for the operation of the market as a whole. In sum, the cost-benefit analysis for *ex ante* specification of prohibited conduct might be relatively favorable in the injunction context, even though the injunction, unlike the underlying patent, targets only the infringing party.

Moreover, a judicial gatekeeper for injunctions has a further comparative advantage. The judicial gatekeeper, unlike the examiner, will likely oversee enforcement or modification of the legal claims that an injunction creates and thus can act in the future to ensure that specially tailored language achieves its intended purpose.²⁸⁸

283. See John M. Golden, *Patentable Subject Matter and Institutional Choice*, 89 TEXAS L. REV. 1041, 1098 (2011) (“The *ex parte* nature of the dominant forms of USPTO proceedings—examination proceedings and *ex parte* reexamination proceedings—means that USPTO examiners primarily interact with parties seeking to obtain patent rights.”); Hovenkamp, *supra* note 249, at 223 (“[A]pplications are evaluated in largely *ex parte* proceedings by overworked government officials . . .”).

284. See Kaplow, *supra* note 254, at 585 (“Whether a law should be given content *ex ante* or *ex post* involves determining whether information should be gathered and processed before or after individuals act.”).

285. Kaplow has noted that time discounting “tends to favor standards, which have a cost advantage at the first stage and disadvantages at [later] stages.” *Id.* at 572 n.33.

286. See ALFRED D. CHANDLER, JR., *INVENTING THE ELECTRONIC CENTURY: THE EPIC STORY OF THE CONSUMER ELECTRONICS AND COMPUTER INDUSTRIES* 4–5 (2005) (“Once . . . core companies establish a viable national industry, entrepreneurial start-ups are rarely able to enter,” with a result “that only a small number of enterprises defines the evolving paths of learning in which the products of new technical knowledge are commercialized.”).

287. See Michael Bednarek, *Responding to Recent Trends in the IP Realm: New Considerations for IP Lawyers and Clients*, in *INTELLECTUAL PROPERTY LAW 2011: TOP LAWYERS ON TRENDS AND KEY STRATEGIES FOR THE UPCOMING YEAR* 41, 51 (2011) (stating that patentees who do not practice their inventions “typical[ly] target high-profile companies with deep pockets”).

288. See Rendleman, *supra* note 5, at 74 (noting that judges have “considerable discretion in administering an injunction; a judge enforcing an injunction against a recusant defendant may choose between modifying the injunction, granting a second injunction, holding the defendant in contempt, imposing a civil contempt or a criminal contempt sanction, and deciding whether to displace the defendant with a receiver”).

These advantages might not be decisive. The common roots and consequences of problems in defining claim and injunction scope run deep. They appear to be grounded in familiar difficulties associated with choosing between relatively rigid rules and comparatively flexible standards.²⁸⁹ The fundamental nature of these difficulties suggests that, with respect to efforts to optimize claim and injunction scope, common problems might dominate over differential advantages. Hence, we might justifiably suspect that decades-long debates over the merits of central claiming and peripheral claiming²⁹⁰—as well as the merits of Type-1-like doctrines of equivalents²⁹¹—suggest that concerns of injunction scope will lack easy universal answers.

We can console ourselves, however, by noting that questions of what type of injunction a court should issue are really just a subset of questions that courts encounter all the time in trying to provide well-tailored yet effective remedies in individual cases. As compared with a Type-1, colorable-differences injunction, a specially tailored injunction, or at least a specially tailored injunction of the reformulated-bounds stripe, tends to front-load questions about an injunction's effective scope. Thus, particularly when specially tailored language is not a product of consent, a court should probably take special care to ensure that the injunction is properly tailored. Model language for certain types of specially tailored injunctions, such as those prohibiting activities frequently correlated with infringement, might help both parties and courts achieve results whose proportions are reasonably tailored, well understood, and relatively cheaply attained.²⁹² More generally, when a court adopts an innovative or sui generis form of relief, such as that

289. See, e.g., Daniel A. Crane, *Rules Versus Standards in Antitrust Adjudication*, 64 WASH. & LEE L. REV. 49, 51 (2007) (suggesting that an existing tendency to shift toward standards-based adjudication in antitrust law will likely be followed by a “swing back toward rules”); Kaplow, *supra* note 254, at 560 (assuming, for purposes of analysis, that “the only distinction between rules and standards is the extent to which efforts to give content to the law are undertaken before or after individuals act” (emphasis omitted)).

290. See, e.g., Burk & Lemley, *supra* note 250, at 1746 (suggesting that U.S. patent law might improve through a return to central claiming, under which “the patentee discloses the central features of the invention . . . and the courts determine how much protection the patent is entitled to”); Fromer, *supra* note 240, at 772 (arguing that patent law should “incorporate further aspects of central claiming”). See generally Golden, *supra* note 20, at 349 (describing a protracted historical shift from central to peripheral claiming).

291. See, e.g., Michael J. Meurer & Craig Allen Nard, *Invention, Refinement and Patent Claim Scope: A New Perspective on the Doctrine of Equivalents*, 93 GEO. L.J. 1947, 2010 (2005) (calling for “balanc[ing] refinement cost savings and innovative incentives created by the [doctrine of equivalents] against the harm to competition and rent-seeking costs created by the doctrine”); Doug Lichtman, *Substitutes for the Doctrine of Equivalents: A Response to Meurer and Nard*, 93 GEO. L.J. 2030, 2030 (2005) (“[T]he doctrine of equivalents should play a significant role in patent interpretation.”).

292. See Kaplow, *supra* note 254, at 580 (observing that spending “relatively little” on the design of a rule “unlikely to apply to many or any acts” “softens the disadvantage of rules . . . by reducing the promulgation cost differential” but might also “reduc[e] the benefit of rules with regard to inducing individuals to behave in a socially optimal manner”).

almost necessarily embodied by a reformulated-bounds injunction, the court should be required to articulate a justification for the injunction's scope that enables the court's reasoning to be scrutinized on appeal.²⁹³ The court's justification might be expected to explain why a specially tailored injunction's specific language is likely to advance interests such as notice, compliance, enforceability, and administrability, and why these advantages outweigh risks of overreaching or underreaching relative to a conventional Type-1 order.²⁹⁴ In addition to facilitating appellate review, articulation of such reasoning should remind the trial court of the need for circumspection in the crafting of injunctive relief, thereby helping prevent abuse of specially tailored injunctions' multifariously malleable form.

In any event, the key point is that the remedial quiver of district courts appears to contain an additional arrow that the Federal Circuit and commentators have often neglected to discuss—the capacity to craft a specially tailored injunction. An individual court in an individual case will have to do the best it can to determine whether a Type-1, colorable-differences or specially tailored injunction will better place the parties and society in an appropriate “rightful position.”²⁹⁵ Concerns of overdeterrence and underdeterrence can justify viewing a Type-1 injunction as the default. But particular circumstances, perhaps including a party's willingness to draft a sensible specially tailored order, can mean that a specially tailored injunction will better balance concerns of rights protection, rights limitation, notice, enforceability, and administrability.

VI. Conclusion

The scope of patent-infringement injunctions is a crucial aspect of the United States' system of patent remedies. Concerns of notice, effective rights enforcement, efficient legal administration, and avoidance of patent

293. See Rendleman, *supra* note 5, at 94 (“In selecting and measuring a remedy, the judge should articulate reasons and apply standards.”). *But cf.* R. Shep Melnick, *Taking Remedies Seriously: Can Courts Control Public Schools?*, in FROM SCHOOLHOUSE TO COURTHOUSE: THE JUDICIARY'S ROLE IN AMERICAN EDUCATION 17, 26 (Joshua M. Dunn & Martin R. West eds., 2009) (“Given the potential sweep and coerciveness of structural injunctions, [U.S.] Supreme Court guidance on the use of injunctions has remained remarkably vague.”).

294. See Thomas, *supra* note 145, at 365–69 (discussing how, in deciding on an appropriate remedy, a “court uses traditional decisionmaking tools such as cost-benefit analysis or balancing of the equities . . . to select from among the possible alternatives”); *cf. id.* at 332 (arguing that prophylactic remedies should be “narrowly targeted at redressing the proven harm” and have “a sufficient causal nexus to the established harm”).

295. David S. Schoenbrod, *The Measure of an Injunction: A Principle to Replace Balancing the Equities and Tailoring the Remedy*, 72 MINN. L. REV. 627, 678 (1988) (“The injunction's aim must be the plaintiff's rightful position, but to achieve that aim, its terms may impose conditions . . . that require actions going beyond the plaintiff's rightful position.”); see also Tracy A. Thomas, *Understanding Prophylactic Remedies Through the Looking Glass of Bush v. Gore*, 11 WM. & MARY BILL RTS. J. 343, 389 (2002) (“The right level of protection commonly accepted for injunctive remedies is the return of the plaintiff to her rightful position . . .”).

overreach combine to generate difficult legal and policy questions about proper injunction scope.

Nonetheless, questions of patent-infringement injunctions' scope have not previously attracted significant attention. Too often in legal contexts not restricted to patent law, injunctions are characterized as simple off switches that generally foreclose the likelihood of any further rights violation and that do essentially nothing more than prevent such a future rights violation. As a result, commentators often invoke a substantially misleading dichotomy between injunctions and monetary relief, a dichotomy that tends to ignore the reality that an injunction can be expected to be only as powerful as the sanctions, often monetary, that back the injunctions up. In many situations, an injunction operates essentially as a mere gateway to higher-than-normal monetary sanctions delivered with higher-than-normal speed. Failure to recognize this fact can contribute to failure to recognize that injunctions can leave their targets with multiple plausible options that might involve a risk of further infringement.

In the patent-infringement context, where the precise scopes of both underlying legal rights and of the injunction itself are commonly uncertain, violation of an injunction typically only risks sanction for civil contempt. Criminal contempt can fail to loom as a realistic possibility. Where any holding of civil contempt is likely to be enforced through monetary sanctions that an enjoined party expects to find bearable, facile distinctions between property rules enforced by injunctions and liability rules enforced by monetary relief can melt away. As the model in subpart II(B) illustrates, an enjoined party's post-injunction decision making can substantially parallel that of a party subject only to the law's underlying legal commands. Consequently, many questions about proper injunction scope can parallel legal and policy questions about proper patent scope. Similar conclusions might follow in other legal regimes where the scope of underlying rights or their proper application is likely to remain unclear or legitimately contested even after an initial judgment of rights violation.

In the meantime, Part III's description of different injunction types and Part IV's study of recent district court practice suggest that there is much room for improvement in how patent-infringement injunctions are crafted. With high frequency, U.S. district courts have issued obey-the-law injunctions that appear to defy Federal Circuit precedent. In a dataset including 143 patent-infringement injunctions issued by district courts in 2010, over 57% feature such apparent error. Once one excludes injunctions focused on either biomedical-substance technology (for which obey-the-law injunctions are rare) or ornamental designs only (for which the rate of apparent error is near 100%), the rate of apparent error rises above 60%. On the other hand, differences between the apparent error rates for preliminary and permanent injunctions or between injunctions that are actively opposed and those that are embedded in consent decrees are not statistically significant after such exclusions.

The high rate at which courts issue obey-the-law injunctions seems surprising because courts already have a decent and readily deployed alternative. The Federal Circuit has approved general use of colorable-differences injunctions that forbid infringement in ways already adjudged to infringe or “no more than colorably different” from them.

Of course, colorable-differences injunctions are rarely ideal. Indeed, specially tailored injunctions that deviate from this form can allow courts to better balance interests in notice, administrability, rights enforcement, and rights limitation.²⁹⁶ Specially tailored injunctions can replace or supplement colorable-differences orders by (1) making affirmative or negative commands regarding activities correlated with rights violations; (2) requiring destruction, disablement, or delivery of specified materials; (3) providing an injunction-specific linguistic formulation of injunction scope; or (4) providing explicit carve outs to protect legitimate infringer or societal interests. Courts should recognize that, under appropriate circumstances, such specially tailored injunctions are both legally permissible and potentially desirable. The result can be a relative win for all concerned—one that helps square some of the circles that patent law creates.

296. See Thomas, *supra* note 145, at 372 (arguing that compliance, notice, and enforcement advantages “make [prophylactic remedies] particularly effective . . . to enforce intangible rights”).

Book Reviews

Still Convicting the Innocent

CONVICTING THE INNOCENT: WHERE CRIMINAL PROSECUTIONS GO WRONG.
By Brandon L. Garrett. Cambridge, Massachusetts: Harvard University Press, 2011. 367 pages. \$39.95.

Reviewed by Jennifer E. Laurin*

"Our society has a high degree of confidence in its criminal trials, in no small part because the Constitution offers unparalleled protections against convicting the innocent."¹

"The availability of technologies not available at trial cannot mean that every criminal conviction, or even every criminal conviction involving biological evidence, is suddenly in doubt. The dilemma is how to harness DNA's power to prove innocence without unnecessarily overthrowing the established system of criminal justice."²

Introduction

It is somewhat remarkable to contemplate that only two decades ago, the fundamental reliability of the American criminal process and its exemplary protections for innocent defendants could still be blithely expressed as a widely held article of faith. In that short space of time, the figure of the innocent wrongly convicted has moved from peripheral invisibility to looming centrality in understanding and discourse about the criminal justice system. It is a collective consciousness shift that is unimaginable—certainly in so short a timeframe—without the advent of DNA technology and its deployment to very publicly establish the innocence of more than 270 convicted individuals over the past two decades.³ None of the traditional error-detection mechanisms that our criminal justice system features—jury acquittals, appellate reversals, even executive pardons—compare to the DNA exoneration in terms of the scientific certainty it projects or the public spectacle it generates. In the pre-DNA dark ages of the American criminal justice

* Assistant Professor, The University of Texas School of Law. For helpful feedback on earlier drafts, thanks to Willy Forbath, Colin Starger, Jordan Steiker, and Nicholas Stepp. Thanks as well to Ryan Goldstein for his typical (and soon to be dearly missed) ace research assistance.

1. *Herrera v. Collins*, 506 U.S. 390, 420 (1993) (O'Connor, J., concurring).

2. *Dist. Attorney's Office for the Third Judicial Dist. v. Osborne*, 129 S. Ct. 2308, 2316 (2009).

3. See *Innocence Project Case Profiles*, INNOCENCE PROJECT, <http://www.innocenceproject.org/know> (providing statistics of DNA exonerations).

system, the innocent prisoner wrongly convicted was, in the words of Judge Learned Hand, a “ghost”⁴—a specter revealed through largely unseen mechanisms, its existence contested by nonbelievers.

It was against the backdrop of that far more skeptical attitude toward the phenomenon of wrongful convictions that Yale Law School professor Edwin M. Borchard published his 1932 study, *Convicting the Innocent*. Presenting sixty-two instances of American wrongful convictions culled from media reports or other chance encounters with an account of erroneous conviction, the book boldly aimed to prove the existence of Judge Hand’s ghost in the face of widespread “supposition” that “[i]nnocent men are never convicted.”⁵ Borchard brought to light the stories of innocents ensnared, of investigations and prosecutions gone bad, and of the near indifference of society to the issue of recompense for the wrongly imprisoned. Moreover, he identified patterns of contributory factors identifiable across the dataset—eyewitness misidentification, unreliable physical and testimonial evidence, false confessions, investigative and prosecutorial overreaching, and poor defense lawyering—and offered recommendations for criminal justice reform.⁶ As the first systematic attempt to document and explain the capacity for breakdown in the criminal justice system, the book is a classic in its field, a touchstone for future examination of the criminal justice system’s capacity for error.⁷

Our criminal justice system today is, by important measures, better calibrated for reliable and fair outcomes than the one reviewed by Borchard: criminal defendants enjoy dramatically expanded procedural protections, mechanisms of social stratification relevant to criminal enforcement and punishment—including de jure racial discrimination—have at least formally receded, and scientific advances permit more accurate assessments of guilt and innocence. But as the continuing stream of DNA exonerations reflects, these advances have not rid the criminal justice system of error. And so, three-quarters of a century after Borchard first aimed to illuminate and explain the most dire of criminal justice system failures, University of

4. *United States v. Garsson*, 291 F. 646, 649 (S.D.N.Y. 1923) (“Our procedure has been always haunted by the ghost of the innocent man convicted. It is an unreal dream.”).

5. EDWIN M. BORCHARD, *CONVICTING THE INNOCENT*, at v (1932) (quoting the reported words of a Massachusetts prosecutor).

6. *Id.* at 367–77.

7. See Joseph D. Grano, Kirby, Biggers, and Ash: *Do Any Constitutional Safeguards Remain Against the Danger of Convicting the Innocent?*, 72 MICH. L. REV. 717, 723 (1974) (asserting four decades after the publication of *Convicting the Innocent* that the book “still constitutes the best graphic study of the problem” of wrongful conviction); Richard A. Leo & Jon B. Gould, *Studying Wrongful Convictions: Learning from Social Science*, 7 OHIO ST. J. CRIM. L. 7, 11 (2009) (“Borchard, in effect, created the template that would be used to study wrongful convictions for many years to come: identify wrongful conviction cases, describe their legal causes, and propose reforms to prevent future miscarriages.”); see also Francis A. Allen, Book Review, 24 U. CHI. L. REV. 779, 779 (1957) (reviewing JEROME FRANK & BARBARA FRANK, *NOT GUILTY* (1957)) (asserting that the later work on innocence by Judge Jerome Frank and his daughter “necessarily invite[d] comparison with Borchard’s classic work”).

Virginia School of Law professor Brandon Garrett offers his own volume of the same title⁸ with similar goals—but in the context of the new world of DNA technology.

Garrett examines the first 250 cases in which convictions were overturned based on postconviction DNA testing that excluded the prisoner as the source of biological evidence attributable to the perpetrator of the crime.⁹ The sheer existence of these cases, and the very public way in which each came to light, has undeniably put to rest any controversy surrounding Borchard's threshold question of *whether* wrongful convictions occur. Today, Garrett contends, the critical inquiries concern *why* conceded miscarriages of justice take place and what, if anything, can be done by way of prevention.¹⁰ More particularly, a crucial, gnawing question is posed by the wave of DNA exonerations of the last two decades: whether “the first 250 DNA exonerations result[ed] from unfortunate but nevertheless unusual circumstances” or rather were “the result of entrenched practices that criminal courts rely upon every day.”¹¹ Garrett convincingly argues that the answer is unquestionably the latter. The characteristics of these failed prosecutions that caused errors to be committed *and* to go undetected—until the fortuitous event of DNA testing—are representative rather than idiosyncratic, systemic rather than episodic.

In substantiation of this contention, *Convicting the Innocent* reports exhaustive research into the background of each exoneration—including review of some 207 complete trial transcripts¹²—in the most detailed portrait to date of the individual and shared characteristics of wrongful convictions. It is a portrait that reveals pervasive patterns of upstream missteps and misconduct that systematically compromised the reliability of guilt assessments as defendants moved through the processes of investigation, charging and arrest, and trial. It is a portrait also of the structures and dynamics that, again systematically, serve to shield those upstream errors from downstream scrutiny by prosecutors, defense attorneys, or courts. And it is a portrait, according to Garrett, that should give us every reason to believe that similar errors are infecting other criminal cases in which DNA evidence will never surface as an arbiter of truth and that counsels a bold path forward for reliability-enhancing reforms to the criminal justice system—in particular, to mechanisms by which law enforcement and prosecutors investigate and shape cases.

8. BRANDON L. GARRETT, *CONVICTING THE INNOCENT: WHERE CRIMINAL PROSECUTIONS GO WRONG* (2011).

9. *See id.* at 285 (defining the parameters of “exoneration”).

10. *Id.* at 6–7; *see also id.* at 270 (describing the problems within the criminal justice system that lead to wrongful convictions).

11. *Id.* at 6.

12. *Id.* at 286.

Given the stature of Borchard's work, Garrett's titular appropriation is bold. But it is deserved. The era of DNA has moved conversations about error in the criminal justice system from "whether" to "why." In the context of that contemporary inquiry, Garrett's *Convicting the Innocent* deserves to occupy—as it arguably already does¹³—the place of prominence that Borchard's did in the debates of old.

On the other hand, *Convicting the Innocent* enters the fray at a time of significant public cognizance of DNA and the phenomenon of wrongful conviction. Even the Supreme Court, with its often-glacial tendencies toward currents of change, has begun to confront the reliability challenges raised by DNA. Yet to date, as the epigraph from *District Attorney's Office for the Third Judicial District v. Osborne*¹⁴ exemplifies,¹⁵ the problem has been treated as raising more of a management problem than a fundamental challenge to legal doctrine. And while in most policy quarters there is far greater awareness of the risk factors associated with wrongful conviction, little action has been taken in response. There is a sense in which generating momentum for reform requires not so much—or at least not only—more information about *how* error is created but novel and specific ideas about how to generate change. And whatever advantages DNA might hold as a conversation starter in this regard, it offers little assistance in generating and bringing to fruition a workable agenda for criminal justice reform. Or so this Review will suggest.

Part I of this Review sketches a brief overview of *Convicting the Innocent* and identifies and contextualizes the book's major contributions to the extant literature on wrongful convictions. In sum, the book offers the most empirically rich and conceptually nuanced descriptive account to date of the machinery of wrongful conviction. And to ongoing debates over the direction of criminal justice reform, it offers a persuasive and sustained

13. Garrett's larger project—this book and its precursor law review articles—has already attained a level of influence, at least as measured by Supreme Court citations, that rivals his progenitor. See *Perry v. New Hampshire*, 132 S. Ct. 716, 738–39 & n.6 (2012) (Sotomayor, J., dissenting) (citing findings of GARRETT, *supra* note 8, with respect to the role of eyewitness-identification evidence in wrongful convictions); *Melendez-Diaz v. Massachusetts*, 129 S. Ct. 2527, 2537 (2009) (citing Garrett's work on forensic science as supporting the importance of the right of confrontation with respect to potentially flawed scientific evidence); *Dist. Attorney's Office for the Third Judicial Dist. v. Osborne*, 129 S. Ct. 2308, 2316 (2009) (citing Garrett for the proposition that state-level legislative efforts provide an adequate opportunity for postconviction DNA testing in the absence of a constitutional right to such testing); *id.* at 2337 n.9 (Stevens, J., dissenting) (citing, in support of the proposition that DNA exonerations counsel that due process embrace finality over accuracy, Garrett's demonstration that "in 50% of cases in which DNA evidence exonerated a convicted person, reviewing courts had commented on the exoneree's likely guilt and in 10% of the cases had described the evidence supporting conviction as 'overwhelming'"); *Baze v. Rees*, 553 U.S. 35, 86 (2008) (Stevens, J., concurring in judgment) (citing, in connection with Justice Stevens's own reconsideration of the constitutionality of the death penalty, Garrett as demonstrating "the exoneration of an unacceptable number of defendants found guilty of capital offenses").

14. 129 S. Ct. 2308 (2009).

15. See *supra* text accompanying note 2.

critique of the position, embodied by the epigraphs above, that the “unparalleled” procedural protections of the American jury system are effective checks on substantive accuracy as well. And yet, *Convicting the Innocent* does not offer quite as powerful an explanatory lens as Garrett sometimes claims, and it does not advance the ball of criminal justice policy reform as far as it might. Part II of this Review suggests that the project is hampered in fulfillment of its descriptive and prescriptive agendas by constraints intrinsic to the data at Garrett’s disposal and by limitations that Garrett’s own framing and methodology impose. Part II further offers that modest but important qualifiers and additions to Garrett’s agenda could enhance the prospect that his worthy contribution to criminal-justice-reform conversations will translate into positive and much-needed change.

I. *Convicting the Innocent*: Summary and Context

Convicting the Innocent presents the fruits of an exhaustive examination of the process by which 250 individuals were investigated, prosecuted, convicted, and ultimately exonerated in criminal cases. The sources assembled by Garrett for the study are as impressive as the rigor he brings to assessment of them: complete trial transcripts in almost 90% of the 234 cases that went to trial; confession statements, interrogation transcripts, laboratory reports, and other investigative documents; and additionally—in the cases of sixteen of the exonerees who pleaded guilty—pretrial hearings, testimony in co-defendant trials, and other evidence that likely would have been presented at these defendants’ trials.¹⁶ Garrett also analyzed all written judicial decisions from the appeals and postconviction proceedings in these cases.¹⁷ In addition to marshalling primary materials, Garrett filled in missing dates, demographic details, and other information with news reports and interviews with attorneys.¹⁸ Garrett and a veritable army of research assistants¹⁹ extracted and coded details concerning defendant and victim demographics, the character of the evidence amassed, legal arguments made, and a litany of other variables, which then were analyzed and aggregated.²⁰ The book is organized around what that work revealed to be the most prevalent error-generating factors in these cases, with separate chapters discussing postconviction proceedings in the exonerees’ cases and surveying the prospects for reform aimed at preventing future wrongful convictions.

Convicting the Innocent stands against the backdrop of a substantial body of historical and contemporary efforts to document and explain wrongful convictions generally and the phenomenon of DNA exonerations in

16. GARRETT, *supra* note 8, at 286–87.

17. *Id.* at 287.

18. *Id.* at 286–87.

19. *See id.* at 352–53 (acknowledging more than two dozen former research assistants).

20. *Id.* at 287.

particular.²¹ A critical question, therefore, is whether it tells us anything that we did not already know about wrongful convictions or, inferentially, the criminal justice system that produces them.

On this score, it must be said that the results of Garrett's study, at least in broad outline, essentially confirm what the reader familiar with *any* prior analysis of wrongful convictions already knew. The key evidentiary pathologies that emerge from the dataset and that frame the first five substantive chapters of *Convicting the Innocent*—eyewitness misidentification, flawed scientific evidence, informant testimony, false confessions, and weak defenses²²—are consistent with the “canonical” list of factors that featured in Borchard's work and that of every subsequent scholar of wrongful convictions.²³ Equally resonant with prior wrongful conviction scholarship is Garrett's critique of the post-trial adjudicatory mechanisms that failed to detect and correct the miscarriages of justice that occurred in these cases, discussed in Chapters Seven and Eight.²⁴ And, at the prescriptive level, the array of proposed fixes that Garrett entertains in Chapter Nine—full recording of interrogations, implementation of eyewitness-identification procedures that comport with the findings of contemporary scientific research on minimizing suggestiveness, overhaul of a neglected forensic science infrastructure, and the development of institutional capacity to investigate the causes of system breakdown when wrongful convictions occur²⁵—tracks a template of proposals in circulation at least since DNA

21. See *supra* note 7 and accompanying text; *infra* note 23 and accompanying text.

22. GARRETT, *supra* note 8, at 8–10, 279–83.

23. The “canonical” characterization is Samuel Gross's. Samuel R. Gross, *Convicting the Innocent*, 4 ANN. REV. L. & SOC. SCI. 173, 186 (2008). For evidence of the consensus, see, for example, AM. BAR ASS'N CRIMINAL JUSTICE SECTION, *ACHIEVING JUSTICE: FREEING THE INNOCENT, CONVICTING THE GUILTY*, at xv–xxviii (2006) [hereinafter ABA REPORT]; BORCHARD, *supra* note 5, at 367–78; EDWARD CONNORS ET AL., *CONVICTED BY JURIES, EXONERATED BY SCIENCE: CASE STUDIES IN THE USE OF DNA EVIDENCE TO ESTABLISH INNOCENCE AFTER TRIAL* 15–19 (1996); JIM DWYER ET AL., *ACTUAL INNOCENCE: FIVE DAYS TO EXECUTION AND OTHER DISPATCHES FROM THE WRONGLY CONVICTED* 246–50 (2000); FRANK & FRANK, *supra* note 7, at 31; C. RONALD HUFF ET AL., *CONVICTED BUT INNOCENT: WRONGFUL CONVICTION AND PUBLIC POLICY* 64 tbl.3.3 (1996); and Samuel R. Gross et al., *Exonerations in the United States: 1989 Through 2003*, 95 J. CRIM. L. & CRIMINOLOGY 523, 527 (2005) [hereinafter Gross et al., *Exonerations*]. In fact, the consensus is so secure that a recently released casebook on wrongful convictions—the first of its kind—is organized around this diagnosis. JAMES R. ACKER & ALLISON D. REDLICH, *WRONGFUL CONVICTION: LAW, SCIENCE, AND POLICY*, at viii–xi (2011) (describing Part II of the book, titled “Leading Correlates and Causes of Wrongful Convictions,” as being divided into chapters focused on eyewitness identification, false confessions, police and prosecutorial conduct, defense inadequacy, forensic evidence, and informants).

24. GARRETT, *supra* note 8, at 178–240; see also BORCHARD, *supra* note 5, at 375–78 (advocating for review by appellate courts on issues of fact as well as law, arguing that “[a]ppeals for errors of law only often defeat the interests of justice”); DWYER ET AL., *supra* note 23, at 218–20 (describing the problem of “courts at every level . . . being pressured to shut their doors to death row appeals”).

25. GARRETT, *supra* note 8, at 241–74.

technology and the “innocence movement” seized center stage in criminal justice reform efforts.²⁶

But while the broad outlines of *Convicting the Innocent*’s substantive contributions are familiar, its deeper aims are neither duplicative nor derivative. The book seeks not simply to describe wrongful convictions but also to explain their genesis—a task that prior work has either not attempted or has executed only at the level of anecdote.²⁷ Garrett leverages the rich descriptive data available to him to generate statistics about not only the presence of certain problematic forms of evidence across cases but also the operation of dynamics underlying the generation and use of that evidence. To these ends, *Convicting the Innocent* makes two particularly distinctive contributions to the wrongful-conviction literature. First, the volume and transparency of the empirical data that it excavates and analyzes, and the role that this data plays in framing the book’s argument, distinguish the book from nearly all of its predecessors in the field and greatly enhance its current and enduring value as a scholarly resource. Second, the portrait that Garrett adeptly paints of the systemic dynamics that produce and then occlude accuracy breakdowns in the criminal justice system adds specificity, nuance, and a powerful theoretical framework to our understanding of the causes of wrongful convictions. These contributions will be explored in turn.

A. *Advancing Our Empirical Understanding*

Garrett’s work with these 250 cases represents the most comprehensive published empirical analysis that has been done on DNA exonerations. Other scholars have long bemoaned the general information deficit in this arena and the significant difficulties entailed by any systematic effort to analyze and quantify error in the criminal justice system.²⁸ These challenges

26. See ABA REPORT, *supra* note 23, at xvii–xxix (detailing issues with wrongful convictions and suggested reforms); DWYER ET AL., *supra* note 23, at 255–60 (listing proposed reforms to protect the innocent similar to those proposed by Garrett); *Priority Issues*, INNOCENCE PROJECT, <http://www.innocenceproject.org/fix/Priority-Issues.php> (identifying seven areas of reform to help prevent future wrongful convictions).

27. See *infra* Part II.

28. See, e.g., Hugo Adam Bedau & Michael L. Radelet, *Miscarriages of Justice in Potentially Capital Cases*, 40 STAN. L. REV. 21, 29 (1987) (describing underappreciated practical hurdles in identifying and analyzing wrongful convictions); Paul G. Cassell, *The Guilty and the “Innocent”*: *An Examination of Alleged Cases of Wrongful Conviction from False Confessions*, 22 HARV. J.L. & PUB. POL’Y 523, 587 n.392 (1999) (describing prohibitive difficulty in obtaining primary source materials for a group of twenty-nine cases); Saul M. Kassin et al., *Police-Induced Confessions: Risk Factors and Recommendations*, 34 LAW & HUM. BEHAV. 3, 5 (2010) (“[E]ven if one could identify a nonrandom set of hotly contested and possibly false confessions, it is often difficult if not impossible as a practical matter to obtain the primary case materials (e.g., police reports; pretrial and trial transcripts; and electronic recordings of the interrogations) . . .”); Richard A. Leo, *Rethinking the Study of Miscarriages of Justice: Developing a Criminology of Wrongful Conviction*, 21 J. CONTEMP. CRIM. JUST. 201, 216–17 (2005) (discussing barriers to the empirical study of wrongful convictions including difficulty in locating cases, proving innocence, and obtaining “primary case materials such as police reports, pretrial and trial transcripts, medical records, and other forms of physical evidence”). I personally encountered these difficulties in individual cases

include not only the obvious difficulty of identifying an appropriate metric for establishing innocence (the nasty problem of proving the negative) but also the equally daunting tasks of locating and assembling the primary source materials that would be necessary for such an analysis. Trial transcripts are not routinely produced, and they are stored such that accessing them often requires access to resourceful court clerks, trial or appellate attorneys, or even individual court reporters. Likewise, records like police reports, physical evidence, laboratory reports, and witness or defendant statements are not uniformly maintained in any one location, particularly once cases have made the rounds through state and federal court in appellate and postconviction litigation. Add to this the complication that each of the thousands of federal, state, and local criminal justice actors and institutions follow independent practices governing retention, tracking, and access to criminal case records.

Garrett possesses a number of advantages in staring down these obstacles. DNA is, at least in the cases that have resulted in convictions being vacated, far less contestable than any other available metric for assessing innocence. DNA exonerations are unusual, discrete, and well-documented events in the life cycle of a criminal case. And a substantial institutional infrastructure has developed for collecting the information necessary for fielding DNA-based claims of innocence—namely, the sixty-odd innocence projects that now exist around the country.²⁹ As a result of leveraging these advantages, the breadth and nuance of the data collection and analysis on display in *Convicting the Innocent* is unprecedented and a stand-alone contribution to the field. Garrett has made both his raw data and the full results of his analysis publicly available through online archives containing scanned transcripts and investigative documents as well as aggregate and case-by-case analyses of the variables examined in his study.³⁰ The archives present opportunity not only for independent review of Garrett's analysis and conclusions but also for future efforts to assess and learn from wrongful convictions.

As the outgrowth of a quantitative empirical study, *Convicting the Innocent* stands apart from the predominant approach of prior examinations

when, as a practicing attorney, I represented wrongly convicted individuals—some of whom are included in Garrett's dataset—in civil rights actions. In the interest of full disclosure, Garrett practiced at the same firm, though our tenures did not overlap. Others have noted that the significant effort required to obtain the type of data that Garrett has managed to assemble has to date posed a barrier to the systematic analysis of wrongful convictions.

29. See Steven A. Krieger, *Why Our Justice System Convicts Innocent People, and the Challenges Faced by Innocence Projects Trying to Exonerate Them*, 14 NEW CRIM. L. REV. 333, 364, 367–70 (2011) (identifying a number of innocence projects and describing the typical process of case evaluation, including assessment of transcripts, police reports, appellate opinions, and other primary sources).

30. See "*Convicting the Innocent*": *Data and Materials*, U. VA. SCH. LAW, http://www.law.virginia.edu/html/librarysite/garrett_innocent.htm (providing links to detailed chapter-by-chapter data, research appendices, and resources).

of wrongful convictions that are fundamentally narrative driven.³¹ Garrett argues his case first and foremost from the strength of his data and, in so doing, avoids some of the downsides that accompany more anecdotal and dramatic accounts of miscarriages of justice—in particular, the lack of a mechanism for contextualizing idiosyncrasy in the cases, and an underanalyzed and underparticularized account of the complex causal links among various evidentiary, procedural, and structural features of cases and their outcomes.

Nevertheless, Garrett is not immune to the power of a good story, which is skillfully deployed in a supporting role. Extended case descriptions introduce each chapter, and shorter narrative snippets pepper the development of his argument to exemplify trends that the data reveals. Garrett's discussion of exoneree Kennedy Brewer's case is illustrative. The case is first noted briefly in Garrett's chapter on forensic evidence as one of seven convictions resting on bite-mark comparison, or forensic odontology—a forensic methodology that has long been alleged to lack scientific validation or standards for practice but that remains in use in criminal investigations.³² Brewer's story is revisited in fuller form to introduce Chapter Seven's discussion of how exonerees' cases fared in appellate and postconviction proceedings, exemplifying the many instances in which courts rejected claims of trial error and actual innocence, and in which prosecutors resisted DNA testing of evidence that could establish innocence.³³ Significantly, the bite-mark evidence makes a troubling reappearance in this postconviction context. Despite its importance in Brewer's trial, it occupied only a minor place in the litany of claims advanced in appellate and habeas proceedings, exemplifying the trend identified by Garrett of substantive evidentiary and innocence-based claims taking a backseat to procedural grounds for error.³⁴ More disturbing, it was that very evidence—scientifically dubious even by contemporary standards—that courts and prosecutors consistently pointed to as providing overwhelming evidence of guilt to countermand Brewer's assertions of innocence.³⁵ Garrett thus humanizes and makes three-dimensional the argument substantiated in drier form by the aggregate descriptive statistics he has assembled: that “once central evidence is contaminated at the earliest stages of a case, the damage cannot be easily discovered or reversed.”³⁶

Not insignificantly, the interweaving of data- and narrative-driven argument in *Convicting the Innocent* also positions the book to reach a

31. See Leo & Gould, *supra* note 7, at 14–17 (noting and critiquing the primarily narrative focus of wrongful-conviction scholarship).

32. GARRETT, *supra* note 8, at 102–05; see also NAT'L RESEARCH COUNCIL OF THE NAT'L ACADS., STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 173–76 (2009) (describing the methodology of and the lack of scientific validation for forensic odontology).

33. GARRETT, *supra* note 8, at 178–80.

34. *Id.* at 182–94.

35. *Id.* at 178–80.

36. *Id.* at 272.

broader audience than might consume more traditional legal-academic work. To be sure, the book is a work of serious legal scholarship, and certainly it lacks the dramatic narrative character of nonfiction wrongful-conviction titles that have gained currency among the popular readership—think Jim Dwyer, Peter Neufeld, and Barry Scheck’s *Actual Innocence*, or more recently John Grisham’s *The Innocent Man*. But Garrett’s light footnoting, his helpful explanations of legal technicalities and investigative techniques, and his accessible explanations of relevant criminal procedure doctrine are just some of the features of *Convicting the Innocent* that will make the volume accessible to nonlegal academics, policy makers, and students.³⁷

It is well to note on this score that *Convicting the Innocent* is not Garrett’s first published study of this dataset. Several prior law review articles have presented subsets of Garrett’s analysis, treating (albeit with a smaller number of available exonerations) the issues of false confessions, forensic science, and appellate review that *Convicting the Innocent* takes up in Chapters Two, Four, and Seven.³⁸ In addition to updating those prior studies to include intervening exonerations, Garrett’s book-length analysis gives sustained attention to important factors that were not the subject of prior articles—in particular, eyewitness identification, informant testimony, and the trial-level decisions of defense counsel.³⁹ And there are analytical advantages to presenting Garrett’s full analysis in a comprehensive volume that cuts across the subcategories of cases and evidence contained within the DNA-exonerate group. The collection of chapters in this unified treatment of Garrett’s study highlights the extent to which the causes of any one wrongful conviction are multivariate, mutually reinforcing, and structural.

B. Contamination and Contagion

In claiming that the dataset of DNA exonerations opens a “unique window on the underside of our criminal justice system,”⁴⁰ Garrett directly challenges a diverse array of commentators who have expressed strong skepticism that the cases where we have demonstrably “gotten it wrong” offer any lessons for run-of-the-mill American criminal justice. Staunch defenders of the adequacy of status quo safeguards are fairly represented by Justice

37. See, e.g., *id.* at 194–96 (explaining the stages of review of criminal convictions).

38. See Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 VA. L. REV. 1, 11–12 (2009) (summarizing the role that flawed forensic science played in exonerees’ convictions); Brandon L. Garrett, *Judging Innocence*, 108 COLUM. L. REV. 55, 60–61 (2008) [hereinafter Garrett, *Judging Innocence*] (describing and assessing overwhelmingly unsuccessful legal challenges to convictions advanced by exonerees in appellate and postconviction proceedings); Brandon L. Garrett, *The Substance of False Confessions*, 62 STAN. L. REV. 1051, 1053 (2010) (examining the role that false confessions played in exonerees’ convictions).

39. See GARRETT, *supra* note 8, at 280–83 (presenting charts related to factors including types of flawed evidence, reliability of identifications, and types of eyewitness misidentifications); *id.* at 351 (listing the subjects of prior articles).

40. *Id.* at 13.

Scalia's expressed confidence that wrongful-conviction studies demonstrate only the happy fact that errors in the criminal justice system are statistical outliers—dividing putative exonerations by total convictions in the relevant time period yields “a success rate of 99.973 percent”⁴¹—and that the system “works” to identify them before they are irrevocable.⁴² But even those who concede the prevalence and problem of wrongful convictions, who suspect that error is pervasive rather than episodic, and who apply themselves to the study of the characteristics of such cases, have grown introspective about the limited ability of that dataset to yield reliable accounts of causation or other information that could ever reliably be generalized to criminal convictions generally. Representative is the “gloomy” message of Sam Gross and Barbara O'Brien: because “exonerations are highly unrepresentative of wrongful convictions in general” and because no reliable data exists to permit comparison between wrongful and rightful convictions, “[w]e do not know much about false convictions, and it will be difficult to learn more.”⁴³ A separate chorus of voices, equally cognizant of the need for criminal justice reform, has nevertheless expressed significant skepticism that reasoning from the lessons of exonerations is a helpful or strategically sound path, cautioning that elevating accuracy as the preeminent value in criminal justice has for a variety of reasons not redounded to the benefit of most criminal defendants—innocent or guilty.⁴⁴

The totality of these critiques generates the concern that the utility of Garrett's project might be quite limited. Indeed, Garrett's data reveals even more vividly than prior studies that, far from a “random audit,”⁴⁵ DNA exonerations are highly unrepresentative of the broader universe of criminal convictions. All but four of the 250 exonerees were men, 70% of which

41. *Kansas v. Marsh*, 548 U.S. 163, 198 (2006) (Scalia, J., concurring) (quoting Joshua Marquis, *The Innocent and the Shammed*, N.Y. TIMES, Jan. 26, 2006, at A23).

42. *See id.* at 199 (“[W]ith regard to the punishment of death in the current American system, [the possibility of wrongful execution] has been reduced to an insignificant minimum.”).

43. Samuel R. Gross & Barbara O'Brien, *Frequency and Predictors of False Conviction: Why We Know So Little, and New Data on Capital Cases*, 5 J. EMPIRICAL LEGAL STUD. 927, 937–40, 958 (2008); *see also* Simon A. Cole, *Cultural Consequences of Miscarriages of Justice*, 27 BEHAV. SCI. & L. 431, 445 (2009) (“[T]he crucial issue that faces serious social scientific scholarship on miscarriages of justice is the problem of generalizability . . .”).

44. *See, e.g.*, Margaret Raymond, *The Problem with Innocence*, 49 CLEV. ST. L. REV. 449, 455 (2001) (“[T]he prevalent display of [postconviction DNA testing] has the potential to send an enduring and unrealistic message: that criminal defendants can and, perhaps, should offer substantial, convincing, and irrefutable proof of their own innocence, ideally, evidence that is as substantial, convincing, and irrefutable as DNA evidence.”); Carol S. Steiker & Jordan M. Steiker, *The Seduction of Innocence: The Attraction and Limitations of the Focus on Innocence in Capital Punishment Law and Advocacy*, 95 J. CRIM. L. & CRIMINOLOGY 587, 609–18, 621–23 (2005) (cautioning that innocence focus leads to harmless error and other reliability-based bars to appellate and postconviction litigation, and to erosion of political support for reforms enjoyed by guilty defendants such as access to counsel).

45. *See* Richard A. Rosen, *Innocence and Death*, 82 N.C. L. REV. 61, 69–70 (2003) (arguing that “DNA testing has provided what can best be described as a random audit of convictions” that “had previously been obtained by legally sufficient evidence”).

were men of color, convicted in only thirty-three states.⁴⁶ Ninety-eight percent of defendants in Garrett's dataset were convicted of either rape, murder, or both crimes;⁴⁷ these offenses account for less than 2% and less than 1% of state convictions, respectively.⁴⁸ Nearly all exonerees were convicted following trials;⁴⁹ felony plea rates generally are at 94%, and even among rapes and murders, the percentage of negotiated resolutions is 84% and 61%, respectively.⁵⁰ Seventeen exonerees (7%) were sentenced to death and eighty (32%) were sentenced to life in prison⁵¹—again, significantly higher proportions than capital and life sentences among all convicted murderers.⁵² And of course, these cases are outliers along their most significant unifying dimension: a DNA exoneration requires physical evidence to test, evidence that is collected in a small minority of criminal cases; in an even smaller minority is such evidence retained over the years and decades; and in an even smaller minority does such evidence have the factual capacity to illuminate with any precision the identity of the crime's perpetrator.⁵³

Nevertheless, Garrett wants to challenge the views of “hardened souls” who view wrongful convictions as either uninformative or untroubling.⁵⁴ To the most ardent skeptics, he argues—quite rightly—that the calculations that permit Justice Scalia a restful night's sleep use far too large a denominator: precisely because cases concluding in DNA exoneration are not representative, the relevant comparison group should be, at its largest, rape and murder cases—a small subset of total prosecutions in the United States

46. GARRETT, *supra* note 8, at 6. *But see* SEAN ROSENMERKEL ET AL., U.S. DEP'T OF JUSTICE, FELONY SENTENCES IN STATE COURTS, 2006—STATISTICAL TABLES 17 tbl.3.2 (2009) (describing the gender and race of persons convicted of felonies in state courts in 2006 and reporting that 83% of all offenses were committed by males and 60% of all offenses were committed by whites).

47. GARRETT, *supra* note 8, at 278 fig.A.2.

48. *See* ROSENMERKEL ET AL., *supra* note 46, at 3 tbl.1.1 (reporting the estimated number of felony convictions in state courts in 2006 and reporting that rape accounted for 1.3% and murder accounted for 0.6% of all of these convictions).

49. *See* GARRETT, *supra* note 8, at 286 (indicating that 234 of the 250 exonerees had a criminal trial).

50. ROSENMERKEL ET AL., *supra* note 46, at 25 tbl.4.1.

51. GARRETT, *supra* note 8, at 5.

52. ROSENMERKEL ET AL., *supra* note 46, at 28 tbl.4.4 (reporting that 2% of felons convicted of murder or nonnegligent manslaughter were sentenced to death and 23% were sentenced to life in prison).

53. *See* JOSEPH PETERSON ET AL., NAT'L INST. OF JUSTICE, THE ROLE AND IMPACT OF FORENSIC EVIDENCE IN THE CRIMINAL JUSTICE PROCESS 3–7 (2010), available at <https://www.ncjrs.gov/pdffiles1/nij/grants/231977.pdf> (finding that there was physical evidence collected in 30% of aggravated assaults, 20% of burglaries, 25% of robberies, 97% of homicides, and 64% of rapes, with lower percentages in each category representing DNA or other biological evidence); David A. Schroeder & Michael D. White, *Exploring the Use of DNA Evidence in Homicide Investigations: Implications for Detective Work and Case Clearance*, 12 POLICE Q. 319, 327 & tbl.1 (2009) (reporting that physical evidence was collected in only between half and two-thirds of Manhattan homicide investigations surveyed).

54. GARRETT, *supra* note 8, at 262.

that also disproportionately relies on the problematic categories of evidence that Garrett assesses.⁵⁵ Furthermore, viewing exonerations as evidence of a system that “works” blinks reality. In the overwhelming majority of DNA exonerations, ordinary appellate and postconviction processes failed to “detect” innocence; the right outcome followed only from the fluke of testable biological evidence being available and the good luck of clearing the gauntlet involved in obtaining DNA testing and release based on exculpatory results—stringent statutory requirements and prosecutorial and judicial resistance among them.⁵⁶ There is nothing about such idiosyncratic dynamics that should reassure us that the system is working to reliably identify and remediate error.⁵⁷

But the major thrust of Garrett’s argument, and the chief contribution of *Convicting the Innocent*, is directed at those who doubt that yet more detail about the nature of wrongful convictions can advance general understanding of how well our criminal justice system works. Though readily conceding the limits of statistical generalizability from his study,⁵⁸ Garrett rests on good old analytic skills as well as the quantity and qualitative nuance of his data to significantly enhance our understanding of *how* factors long known to be prevalent in wrongful convictions lead to erroneous results. Garrett makes the case that the dynamics driving that process are fundamentally “systemic” in nature—generated not by individual decision making or idiosyncrasies of particular cases but rather by processes endemic to how criminal cases are investigated, prosecuted, and adjudicated.⁵⁹ The appropriate metaphor for errors that come to infect wrongful convictions is not the “bad apple” but rather the “Whack-a-Mole”: removing a particular case or actor from the equation will not prevent the error from popping up farther up or down the road.⁶⁰

Central to Garrett’s case in support of this thesis, and an independent conceptual contribution of the book, is the dynamic of “contamination” that he identifies and traces.⁶¹ The idea underlying the term as he uses it is that the probative value we assign to evidence in a criminal case rests on a set of assumptions about the integrity of its production, which can themselves be undermined by certain influences. In scientific testing, of course, the results of a test involving introduction of a reagent to a substance—say, to urine being evaluated for the presence of illegal drugs—are informative and reliable only if the substance tested is unadulterated by foreign material that could

55. *Id.* at 264.

56. *See id.* at 225–31 (describing statutory and procedural barriers to postconviction DNA testing).

57. *Id.* at 263.

58. *Id.* at 288.

59. *Id.* at 265–68.

60. *Id.* at 265–66.

61. *Id.* at 21.

trigger a positive result. Analogously, police are trained to evaluate the relevance of reliability of suspect confessions largely on the basis of their substantive content and the likelihood that someone who had not committed a crime would know the details provided. If a suspect learned details through media accounts, street gossip, or (most troublingly) from investigators, this introduction of a foreign substance—a source other than that which would indicate the suspect's independent knowledge of the crime—"contaminates" the confession or statement and undermines its probative value. Similarly, eyewitness-identification procedures are designed to test a witness's memory, to evaluate the strength of the identification based upon an implicit estimate of the odds that the witness was drawing on something *other than* a memory generated when viewing criminal conduct. That evaluation is contaminated by express or implicit encouragement to select a particular suspect on the basis of a poorly constructed procedure or direct suggestion from police.

The cases in Garrett's dataset were rife with contamination of this sort. For example, in all but two false confessions and all but two instances of inculpatory testimony from jailhouse informants, exonerees were alleged to have revealed significant details concerning the crime⁶²—details that, DNA now shows, could not have been independently known by them. And in 78% of examined cases involving eyewitness evidence, there was evidence that police administering the identification procedures had contaminated the results with conduct that scientific research and legal doctrine alike recognize as suggestive, such as making remarks indicating who should be selected from a lineup or composing a lineup in a way that made the defendant stand out from other individuals.⁶³

Even more troubling is the occluded and contagious nature of contamination that Garrett's analysis reveals. At the investigative stage, contamination in one respect often spreads to falsely bolster other evidence in the case. This occurred in the Central Park Jogger case—notorious first for the brutal crime that gave rise to the case and later for the circumstances underlying the wrongful conviction of young men for the crime.⁶⁴ The defendants were told in their interrogations about weak fingerprint evidence found on a victim's "satin" jogging shorts and were thereby not only convinced to confess but also provided with a nonpublic detail that came to appear in, and falsely strengthen the credibility of, their statements.⁶⁵ Moreover, in this case and others, contamination repeatedly evaded detection

62. *Id.* at 20, 130–34.

63. *Id.* at 49; *see also* United States v. Wade, 388 U.S. 218, 232–35 (1967) (describing suggestive practices).

64. *McCray v. City of New York*, Nos. 03 Civ. 9685(DAB), 03 Civ. 9974(DAB), 03 Civ. 10080(DAB), 2007 WL 4352748 (S.D.N.Y. 2007); *see* Sharon L. Davies, *The Reality of False Confessions—Lessons of the Central Park Jogger Case*, 30 N.Y.U. REV. L. & SOC. CHANGE 209, 213–16 (2006) (describing the attack and summarizing the subsequent investigation).

65. GARRETT, *supra* note 8, at 22, 153.

because no contemporaneous documentation was available to permit scrutiny of the processes by which the problematic evidence was generated.⁶⁶ Thus, even assuming competent and adequately resourced counsel—a generous assumption at best⁶⁷—defendants’ ability to challenge seemingly damning evidence at trial (or at the charging and plea bargaining stages⁶⁸) was highly compromised.⁶⁹

Following conviction, a perverse synergy between contamination and doctrines of criminal procedure and appellate review further shielded error from detection. Garrett’s data reveals that few exonerees even challenged the most troubling evidentiary features of their convictions on appeal.⁷⁰ Whether or not they did, contamination often operated to block judicial scrutiny of the troublesome evidence. Thus, for example, appellate and postconviction challenges to eyewitness-identification evidence were made in approximately half of the cases where such evidence was presented.⁷¹ Those claims that were brought almost uniformly foundered on application of the Supreme Court’s *Manson v. Brathwaite*⁷² test for applying the Due Process Clause to identification testimony, whereby even the use of suggestive identification procedures does not preclude admissibility of the eyewitness evidence so long as other factors—“indicia of reliability”⁷³—corroborate the identification.⁷⁴ But Garrett demonstrates that, far from indicating reliability, the corroborative factors considered by courts are in fact *themselves* likely to have been affected by contamination.⁷⁵ So, too, did contamination thwart challenges to confessions, as courts repeatedly pointed to a defendant’s apparent recounting of nonpublic facts as evidence that the

66. See, e.g., *id.* at 43, 68, 142–43 (providing examples such as the failure to record interrogations from start to finish, the failure to record eyewitnesses’ initial description of the culprit, and the failure to require prosecutors to disclose information regarding informants).

67. See, e.g., Eve Brensike Primus, *Structural Reform in Criminal Defense: Reallocating Ineffective Assistance of Counsel Claims*, 92 CORNELL L. REV. 679, 686–87 (2006) (discussing prevalence and causes, individual and structural, of deficiencies in defense-counsel representation).

68. See GARRETT, *supra* note 8, at 150 (reporting that nineteen exonerees pleaded guilty). Garrett does not recount the number of exonerees who discussed pleas at any point in their prosecutions—information that would be nearly impossible to reliably assemble. But given that rapes and murders have far lower plea rates than most crimes, see *supra* note 50 and accompanying text, it is plausible to suspect that many exonerees were not offered the opportunity to consider this disposition of their cases.

69. *Id.* at 272–73.

70. *Id.* at 184.

71. *Id.*

72. 432 U.S. 98 (1977).

73. See *Perry v. New Hampshire*, 132 S. Ct. 716, 720 (2012) (“But if the indicia of reliability are strong enough to outweigh the corrupting effect of the police-arranged suggestive circumstances, the identification evidence ordinarily will be admitted, and the jury will ultimately determine its worth.”).

74. See *Manson*, 432 U.S. at 110–14 (rejecting a “per se” bar to identification evidence procured with suggestive procedures).

75. GARRETT, *supra* note 8, at 62–63, 188.

statements were, in their totality, uncoerced and reliable.⁷⁶ More broadly, the courts' application of harmless-error doctrines to an array of asserted trial errors—a feature of some 50% of cases⁷⁷—led them to place emphasis on the apparent strength of contaminated evidence to affirm convictions in spite of procedural error.⁷⁸

Garrett's analysis thus moves beyond the "tautological" account of causation that is a feature of much of the wrongful-conviction literature.⁷⁹ While Garrett's data is significant, an equally great contribution is the explanatory account he weaves to reveal how the introduction of error in a criminal case cannot readily be undone and why traditional reliance on court-centered error-correction devices—either at trial or in multiple iterations of appellate and postconviction review—is therefore misplaced. And while the 250 DNA exonerations are in some respects extraordinary, Garrett makes the case that nothing about their uniqueness was causally relevant to erroneous outcomes. The cogent and disturbing inference is that "[t]he problems that occurred in these cases," as Garrett contends, "are just as likely to infect" other cases "where DNA will never be available."⁸⁰

II. The Limitations of *Convicting the Innocent*

Notwithstanding the significant contributions of *Convicting the Innocent*, there are limitations to Garrett's study and the extent to which it in fact advances our "understand[ing] [of] why criminal prosecutions can go wrong—and how we can avoid convicting the innocent."⁸¹ To a large extent, these limitations reflect not failures of execution but rather constraints that are intrinsic to the project. That is to say, if *Convicting the Innocent* goes as far as one can in using wrongful convictions as a lens into the criminal justice system, it may unintentionally make the case that this distance is ultimately modest and the tools it deploys of limited assistance in diagnosing accuracy-based criminal justice failures. But at the same time, there is a sense of disappointment to be felt by those sympathetic to the project's normative agenda, stemming from the book's failure to chart a more novel and ambitious course for reform. This part takes each category of criticism in turn.

A. Limited Diagnostics

Convicting the Innocent enhances our understanding of how wrongful convictions are produced in our criminal justice system. But to what extent does it truly reveal causes of error, in the sense of being able to predict that

76. *Id.* at 39–40.

77. *Id.* at 185, 201–02.

78. *Id.* at 202, 211.

79. Gross & O'Brien, *supra* note 43, at 932.

80. GARRETT, *supra* note 8, at 265.

81. *Id.* at 13.

reforms aimed at altering or eradicating certain factors will enhance accuracy? Garrett advances fairly strong claims in this regard.⁸² But there are at least three reasons to doubt that *Convicting the Innocent* is really positioned to deliver on that promise.

First, however persuasive Garrett's causal account may be, it is ultimately unproven—and likely unprovable—from the data available to him. As others have catalogued, there are numerous barriers to rigorous application of social science methodologies in this arena, most notably the lack of comparators: we have no idea how frequently the variables isolated by Garrett (and others) are present in accurate acquittals of the innocent—or for that matter, accurate convictions of the guilty.⁸³ If suggestive identification procedures were prevalent in cases in which we had as much confidence in the accuracy of their outcomes as we have in the inaccuracy of the DNA exonerees' trials, one would be hard-pressed to characterize this factor as a likely cause of errors in the latter group.⁸⁴ To be sure, there is good reason that Garrett does not pursue this line of inquiry: detailed information concerning the evidence that features in most criminal cases, whether ending in acquittal or conviction, is practically unobtainable, and the outcomes in such cases are not ordinarily susceptible to confirmation via mechanisms like DNA testing. To his credit, Garrett attempted to make use of a fascinating dataset: the approximately 50% of individuals for whom the Innocence Project has obtained DNA testing whose DNA tests *confirm guilt*.⁸⁵ This small dataset, however, ultimately proved inadequate to construct a comparison to the exoneree group.⁸⁶ Perhaps over time this unusual sample of identifiable guilt “confirmations” will grow to the point that it can be exploited for more probative causal analyses. But for now, in the absence of

82. See, e.g., *id.* (positioning the book to explain “why criminal prosecutions can go wrong—and how we can avoid” error); *id.* at 201 (suggesting that the fact of DNA exoneration reveals incorrect determinations of “harmless error” and shows that the errors in trials in fact played a role in convictions); *id.* at 274 (“The errors in these exonerees’ cases were . . . caused by systemic failures.”).

83. See Leo & Gould, *supra* note 7, at 18 (making this point with regard to wrongful-conviction literature generally).

84. See *id.* at 20–24 (discussing the limited causal conclusion that can be drawn in the absence of comparisons of variable prevalence among varied case outcomes).

85. See GARRETT, *supra* note 8, at 233–34.

86. Garrett, *Judging Innocence*, *supra* note 38, at 141 & n.293. Garrett does exploit what points of statistical comparison are fairly available to him. In previously published work that is described in *Convicting the Innocent*, Garrett compared the success rates of DNA exonerees in appellate and postconviction litigation with those of randomly selected litigants convicted of rape and murder in cases with no DNA testing; this “matched-comparison” technique established that innocent defendants raised comparable claims and fared no better in litigation as compared to their presumptively non-innocent counterparts. See GARRETT, *supra* note 8, at 198 (reporting results of the study); Garrett, *Judging Innocence*, *supra* note 38, at 69–116 (describing the study and reporting results); Leo & Gould, *supra* note 7, at 22–23 (calling Garrett’s study the “most comprehensive” available matched-comparison analysis of wrongful convictions).

comparators, we cannot meaningfully test the proposition that the factors Garrett isolates are generating error.

Even assuming that the dynamics described in *Convicting the Innocent* are, as a group, predictive of erroneous outcomes, Garrett's data reveals the near impossibility of isolating and assessing the significance of any *single* factor in a given case. Illustrative on this score is the Jeffrey Deskovic case. As Garrett describes in Chapter Two, the sixteen-year-old Deskovic gave false inculpatory statements that were introduced against him in his trial for the rape and murder of his high school classmate.⁸⁷ Garrett describes how the trial transcripts revealed the "central[ity] [of] Deskovic's alleged admissions . . . to the State's case," noting that it was "the only evidence connecting Deskovic to the crime."⁸⁸ But there is even more to this story than what Garrett tells—as revealed by, among other sources, a postexoneration case review conducted at the behest of the district attorney's office that convicted Deskovic.⁸⁹ Delving deeper problematizes the premise that Deskovic's confession drove the tragic outcome in his case.

At the time of the investigation, DNA testing was performed on semen recovered from the victim. Remarkably, the results excluded Deskovic as a potential source, and the jury in his case heard this evidence.⁹⁰ (Microscopic hair examination also revealed, and the jury learned, that hairs found on the victim's body could not have been shed by Deskovic.)⁹¹ But the state also introduced "questionable" forensic evidence designed to establish circumstantially that the fifteen-year-old victim might have had consensual sex prior to her death and that this partner was the source of the semen.⁹² In a sense, this strand of the narrative reveals the hydraulic force that Deskovic's confession had in the case: once obtained, it set the prosecution on a course from which even DNA evidence did not prompt reexamination—except to pursue investigative avenues to reconcile the science with the admission.⁹³ But on the other hand, it is clear that to understand what went wrong in Deskovic's case, one must examine not only the pathologies of false confessions but also issues concerning faulty forensic science and investigative tunnel vision—among other factors.

But there also are more idiosyncratic features of the case. Garrett does not discuss that Deskovic's criminal trial was marred by the loss of evidence in the state's custody: the victim's bra, a description of which had been an

87. GARRETT, *supra* note 8, at 14–18.

88. *Id.* at 16.

89. To view this review, please see LESLIE CROCKER SNYDER ET AL., REPORT ON THE CONVICTION OF JEFFREY DESKOVIC (2007), available at <http://truthinjustice.org/Jeffrey-Deskovic-Comm-Rpt.pdf>. From 2006 to 2009, I was one of several lawyers representing Deskovic in ongoing civil litigation stemming from his conviction.

90. *Id.* at 32.

91. *Id.* at 33.

92. *Id.* at 21–24.

93. GARRETT, *supra* note 8, at 16–17.

important, allegedly corroborative detail in Deskovic's confession.⁹⁴ Had that evidence been available to the defense at trial, Deskovic's lawyer might have been able to argue to the jury that the actual bra made Deskovic's account of the crime impossible. And then there was the suspect "profile" developed by police early on in the investigation—a profile that matched Deskovic in important respects and that ultimately, once the true perpetrator was identified, proved to be grossly inaccurate.⁹⁵ Had the profile not been developed, Deskovic might never have become a suspect in the first instance.

The Deskovic case thus highlights the extent to which errors in criminal adjudications might well be fueled by factors other than those that Garrett highlights on the basis of their patterned recurrence. What of the erroneous suspect profile that may have strengthened investigators' commitment to focus on Deskovic in the first instance, or the lost evidence that limited the defense's trial strategy? To what extent were these aspects of the investigation and prosecution independently significant forces that might have generated error even in the absence of a false confession? The fact that some features of the 250 exonerations are amenable to categorization across the dataset does not in and of itself make those features more causally significant in any given case; and conversely, the fact that other dynamics in any given case are idiosyncratic does not render them less consequential.

As Garrett undoubtedly appreciates, Deskovic's trial was surely far from unique in featuring multilayered and convergent decisions and errors by stakeholders as well as mundane and uncategorizable but potentially consequential events.⁹⁶ Indeed, as the previous part argued, one of *Convicting the Innocent's* most notable contributions is its effort to document and explain some of the forces driving this overdetermination. But at the same time, Garrett repeatedly points to discrete categories of evidence as material to the erroneous outcomes in these cases: confessions, eyewitness identifications, forensic analysis, and informant testimony all independently have their turn as the "central" evidence in cases where they appear.⁹⁷ And in further service of his causal account, Garrett emphasizes that which can be categorically grouped and deemphasizes that which is idiosyncratic.

There is thus a tension between Garrett's effort to generate a nuanced, descriptive portrait and his interest in asserting broader claims about the independent significance of the factors that he highlights. At the theoretical level, it necessarily undermines the extent to which *Convicting the Innocent* can explain causal relationships even within its dataset, much less within the

94. SNYDER ET AL., *supra* note 89, at 29–30.

95. *Id.* at 10–11.

96. See Brandon L. Garrett, *Innocence, Harmless Error, and Federal Wrongful Conviction Law*, 2005 WIS. L. REV. 35, 66–67 (discussing the "high burden" faced by defendants who have to argue against many pieces of significant evidence).

97. See GARRETT, *supra* note 8, at 16, 79, 91, 139 (describing cases in which the respective factors were essential to acquiring false convictions).

larger universe of criminal prosecutions. In a practical sense, and as the next subpart will explore more fully, the tension begs a remedial question: if we want to ensure that there are no more Jeffrey Deskovics, do we accomplish this goal through reform (either at the individual-department level or more broadly) of interrogation practices? Or will a more comprehensive reexamination of investigative and prosecutorial tools and tactics be required? At times, Garrett's isolation of purported "causes" appears to suggest confidence in the former proposition, but more contextual reflection on cases like Deskovic's suggests the latter is the necessary course.

Finally, just as Garrett gives short shrift to causes that, although idiosyncratic, might well be highly explanatory, he pays little attention to upstream forces that drive the causal events he does identify. *Why* do police engage in suggestive identification practices or feed facts to suspects and witnesses? *Why* do forensic scientists—even those practicing in sound and validated disciplines—overstate the probative value of their conclusions? We lack an understanding of what might be termed "root cause" in Garrett's dataset or more generally. With the exception of the final chapter's brief assertion that psychological research into cognitive biases held by police and prosecutors may have explanatory value,⁹⁸ *Convicting the Innocent* does not develop its causal analysis in this respect.

In spite of these holes in its causal account, it must be said that *Convicting the Innocent* does no worse than prior studies of wrongful convictions⁹⁹—and in many respects, as argued above, does much better. But the limitations that the book displays in this regard are nevertheless important in that, for at least two reasons, they impede fulfillment of Garrett's announced agenda of "better understand[ing] why criminal prosecutions can go wrong—and how we can avoid convicting the innocent."¹⁰⁰

First and most generally, to unqualifiedly designate the factors analyzed in *Convicting the Innocent* as "causes" falsely suggests that we possess a deeper understanding than we do of the nature of wrongful convictions. Of course, that Garrett does not fully answer all questions on the table does not itself undermine the ambitious and valuable work he does undertake. But the impressive fact that his project is already something of a touchstone for conversations in this field¹⁰¹ means that misapprehension of the work it leaves undone might impede sustained examination of the criminal justice system beyond the parameters that Garrett specifies. This is particularly a concern given the susceptibility among all stakeholders (law enforcement, lawyers, judges and juries, policy makers, the media, the general public, and even scholars) to the allure of DNA, to statistical overclaiming, and to addressing

98. *Id.* at 266–67.

99. See Leo & Gould, *supra* note 7, at 19–21 (criticizing the entire field of wrongful-conviction literature for the thinness of its causal analysis).

100. GARRETT, *supra* note 8, at 13.

101. See *supra* note 13 and accompanying text.

the low-hanging fruit of reform—i.e., deficiencies in criminal investigations and adjudications that we *can* see, name, categorize, and explain. Consider on this score the frequency with which Garrett’s work (among others’) is cited in support of the proposition that “eyewitness misidentification is ‘the single greatest cause of wrongful convictions in this country.’”¹⁰² Increasing attention to identifying and addressing risk factors in wrongful convictions is, undoubtedly, to be cheered. But such (certainly inadvertent) overclaiming about the extent of our causal understanding could well divert attention from less obvious and even more structural factors than what wrongful-conviction scholarship has commonly highlighted—in the former category, a variety of “facially unobtrusive” procedural rules,¹⁰³ and in the latter, dynamics such as cognitive limitations of jurors, or more fundamentally forces of racial and class inequality¹⁰⁴—factors that might operate independent of, or even drive, the variables Garrett identifies.

Legal scholars and social scientists alike are attempting the challenging task of assessing the causes of criminal adjudicative error, including through application of social science methodologies to data other than that offered by

102. Perry v. New Hampshire, 132 S. Ct. 716, 738–39 & n.6 (2012) (Sotomayor, J., dissenting) (quoting State v. Henderson, 27 A.3d 872, 885 (2011), and citing additional state court decisions as well as GARRETT, *supra* note 8); see also Gross et al., *Exonerations*, *supra* note 23, at 542 (“The most common cause of wrongful convictions is eyewitness misidentification. This is not news. It was first shown in 1932 by Professor Edwin Borchard in his classic book *Convicting the Innocent*, and it is apparent again in our data: In 64% of these exonerations (219/340), at least one eyewitness misidentified the defendant. The pattern, however, is heavily lopsided. Almost 90% of the rape cases (107/121), but only half of the homicides (102/205), included at least one eyewitness misidentification.” (footnote omitted)); Cynthia E. Jones, *The Right Remedy for the Wrongly Convicted: Judicial Sanctions for Destruction of DNA Evidence*, 77 *FORDHAM L. REV.* 2893, 2928 & n.199 (2009) (describing eyewitness identification, non-DNA forensic evidence, informant testimony, and confessions as “the leading causes of wrongful convictions” and citing Garrett’s prior work in support).

103. See, e.g., Andrew M. Siegel, *Moving Down the Wedge of Injustice: A Proposal for a Third Generation of Wrongful Convictions Scholarship and Advocacy*, 42 *AM. CRIM. L. REV.* 1219, 1226 (2005) (“[F]acially unobtrusive procedural guidelines and structuring provisions operate to distort incentives, obscure relevant information, and bias results.”); see also Andrew D. Leipold, *How the Pretrial Process Contributes to Wrongful Convictions*, 42 *AM. CRIM. L. REV.* 1123, 1124 (2005) (examining the impact of pretrial procedural rules on the rate of wrongful convictions); Michael D. Pepson & John N. Sharifi, *Lego v. Twomey: The Improbable Relationship Between an Obscure Supreme Court Decision and Wrongful Convictions*, 47 *AM. CRIM. L. REV.* 1185, 1187 (2010) (arguing that raising the burden of proof in pretrial evidentiary hearings “to beyond a reasonable doubt . . . [is] perhaps the most effective means of minimizing wrongful convictions”).

104. See generally DAVID COLE, *NO EQUAL JUSTICE: RACE AND CLASS IN THE AMERICAN CRIMINAL JUSTICE SYSTEM* (1999) (arguing that structural racism and inequality undermine administration of criminal justice); WILLIAM J. STUNTZ, *THE COLLAPSE OF AMERICAN CRIMINAL JUSTICE* (2011) (pointing to structural racism and inequality in administration of criminal law but asserting ultimately that evisceration of local democratic control over the criminal justice system is the root cause of these failures); Dan Simon, *The Limited Diagnosticity of Criminal Trials*, 64 *VAND. L. REV.* 143, 146 (2011) (assessing a variety of structural limitations that cause the criminal trial to “fall[] short of delivering the level of diagnosticity that befits its epistemic demands and the certitude that it proclaims”).

known exonerations.¹⁰⁵ And there is every indication in *Convicting the Innocent* that Garrett himself is hopeful that further research along these lines will continue and that our understanding of wrongful convictions will grow in the aftermath of his study.¹⁰⁶ For that to occur, it is critical that his audience also know that Garrett's work has not occupied the field.

A second worry about the limited explanatory reach of Garrett's analysis is potentially more concerning. Garrett is clear-eyed about the unrepresentative nature of the prosecutions that he studies.¹⁰⁷ He nevertheless contends, with some persuasive force, that the lessons to be learned from these 250 cases are broadly applicable—at least to other rape and murder prosecutions, and perhaps as well to other crimes such as robberies that rarely feature dispositive biological evidence (and so are poor candidates for DNA exoneration) but that frequently feature reliance on identifications, confessions, and non-DNA forensic science.¹⁰⁸ But even this broader universe is still a tiny fraction of all criminal prosecutions, the overwhelming majority of which are (1) unlikely to proceed past the earliest stages of investigation (and thus do not entail the gathering of forensic evidence, taking of suspect and witness statements, conducting of identification procedures, and so forth) and (2) likely to resolve in plea bargains.¹⁰⁹ Assuming that wrongful convictions are also to be found within this broader set of cases—and ample circumstantial evidence supports that premise¹¹⁰—the “causes” that Garrett isolates and aims to remediate bear little relevance.

105. See generally, e.g., Karl Ask & Pär Anders Granhag, *Motivational Sources of Confirmation Bias in Criminal Investigations: The Need for Cognitive Closure*, 2 J. INVESTIGATIVE PSYCHOL. & OFFENDER PROFILING 43 (2005) (examining the source of investigative tunnel vision); Alafair S. Burke, *Improving Prosecutorial Decision Making: Some Lessons of Cognitive Science*, 47 WM. & MARY L. REV. 1587 (2006) (describing cognitive bias in the work of prosecutors); Itiel E. Dror et al., *Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications*, 156 FORENSIC SCI. INT'L 74 (2006) (analyzing cognitive bias with regard to forensic scientists); Keith A. Findley & Michael S. Scott, *The Multiple Dimensions of Tunnel Vision in Criminal Cases*, 2006 WIS. L. REV. 291 (discussing cognitive biases affecting actors at all stages of criminal investigation and adjudication). For an explanation of how social scientists use aggregated case studies, matched-comparison samples, and path analysis to understand causation in wrongful convictions, see Leo & Gould, *supra* note 7, at 19–25.

106. GARRETT, *supra* note 8, at 289.

107. *Id.* at 288–89.

108. *Id.* at 262–65.

109. See *supra* notes 48, 50, 53 and accompanying text.

110. Garrett points to a Department of Justice study finding that when DNA testing was conducted in federal criminal investigations 25% of primary suspects were eliminated prior to trial. GARRETT, *supra* note 8, at 12 (citing CONNORS, *supra* note 23, at 20). Other scholars have examined various manifestations of accuracy concerns in guilty pleas, with different critical frameworks. See, e.g., Albert W. Alschuler, *The Prosecutor's Role in Plea Bargaining*, 36 U. CHI. L. REV. 50, 60–66 (1968) (examining systemic pressure on innocent defendants to plead guilty); Josh Bowers, *Punishing the Innocent*, 156 U. PA. L. REV. 1117, 1134 (2008) (acknowledging the prevalence of wrongful convictions pursuant to guilty pleas in low-level offenses and questioning the premise that such events must be eliminated); Kevin C. McMunigal, *Disclosure and Accuracy in the Guilty Plea Process*, 40 HASTINGS L.J. 957, 965–67 (1989) (exploring and seeking to ameliorate accuracy concerns in plea bargaining); Robert E. Scott & William J. Stuntz, *Plea Bargaining as*

Conversely, these 250 cases almost certainly feature their own distinctive etiology of error. Thus, for example, while informant testimony featured in only 21% of the trials in Garrett's study, dwarfed by the percentages of trials that included forensic evidence and identification testimony,¹¹¹ this category of evidence is almost certainly significantly more prevalent in drug cases—cases that are absent from Garrett's set, that are virtually unamenable to illumination through DNA testing, and that comprise 33% of felony sentences in state courts.¹¹² Also largely uncaptured by Garrett's study are convictions procured through plea, which scholars have long asserted may exhibit a particular set of risks for innocent defendants.¹¹³

That the nature of the dataset that Garrett works with places these far more typical criminal cases beyond the scope of his study is not in and of itself a deficiency of the work. But it is all too easy to let run-of-the-mill injustice fall off the radar screen of reform. For many of the same reasons that they do not feature in Garrett's analysis, *most* criminal investigations lack distinguishing features or adequate stakes to attract attention. The plight of defendants who plead guilty to crimes is unlikely to garner the personal or political sympathy necessary to rally policy makers or, for that matter, advocates around their cause.¹¹⁴ Given limited resources—fiscal, political, and otherwise—reform priorities are likely to be zero-sum, and efforts that redound to the benefit of cases typified by Garrett's study may decrease the likelihood that more common and less attractive issues will be addressed. And yet, from the standpoint of sheer numerosity, the extent of the “innocence problem” among these convicts may well dwarf not only Garrett's dataset but also any analogous convictions not yet identified as erroneous. Again, given the current and prospective profile of *Convicting the Innocent*, it is a shame that the volume does not say more to shine a light on the more workaday ills that will remain unaddressed even in a universe where the lessons of Garrett's study are fully internalized.

Contract, 101 YALE L.J. 1909, 1911 (1992) (developing the argument that structural features of plea bargaining lead innocent defendants to be offered and accept the same deals as the guilty).

111. GARRETT, *supra* note 8, at 279 fig.A.5.

112. See ROSENMERKEL ET AL., *supra* note 46, at 3 tbl.1.1. To the author's knowledge, no empirical data exists quantifying the prevalence of informant testimony in drug cases. The proposition that it is prevalent, however, would seem uncontroversial. See ALEXANDRA NATAPOFF, SNITCHING 25–26 (2009) (asserting in comprehensive work on informants that drug enforcement is a primary arena in which informant evidence is used and that it is uniquely occluded and unregulated in this field).

113. See, e.g., Bowers, *supra* note 110, at 1119–21 (discussing typical characteristics of innocent defendants who plead guilty); Gross & O'Brien, *supra* note 43, at 931 (“[I]t may well be that a major cause of these comparatively low-level miscarriages of justice is the prospect of prolonged pretrial detention by innocent defendants who are unable to post bail.”); Scott & Stuntz, *supra* note 110, at 1911 (assessing “strategic impediments to efficient bargains [that] lead to a pooling of guilty and innocent defendants” and to members of both categories “being offered (and taking) the same deals”).

114. Cf. Emily Hughes, *Innocence Unmodified*, 89 N.C. L. REV. 1083, 1089–92 (2011) (arguing that the innocence movement has neglected important concerns surrounding guilty pleas).

B. *Limited Prognostics*

The prescriptive vision of *Convicting the Innocent* is corollary to its diagnoses and in particular its focus on the introduction and obfuscation of investigative error in criminal cases. As Part I of this Review suggested, Garrett's proposals—from the recording of interrogations, to the adoption of scientifically grounded eyewitness-identification protocols, to an overhaul of forensic science—are as familiar as they are extensive, tracking what is essentially the standard reform package advanced in the DNA era.¹¹⁵ Importantly, Garrett shares with that innocence-driven reform paradigm the view that priority must be placed on “reforming criminal investigations to prevent wrongful convictions in the first instance,”¹¹⁶ and a tendency to deemphasize courts and judicial doctrine as important staging grounds for reform. That *Convicting the Innocent*'s prescriptions are not innovative is, again, not in itself a deficiency. Garrett aims in part to demonstrate that the smattering of states and localities that have reexamined their criminal justice practices in the wake of DNA exonerations have largely adopted precisely the proposals he advances, such that while reform may effect a “sea change,”¹¹⁷ it is neither unprecedented nor infeasible. What emerges as a briefly sketched portrait of a decade and a half of sporadic but nevertheless substantively radical criminal justice reform is both illuminating and refreshingly optimistic, particularly in showcasing the extent to which political actors, and not just courts, have sidestepped ordinary barriers to taking on police and prosecutorial practices.¹¹⁸

But, given that what little remedial action has been spurred by the DNA-driven revelation of wrongful convictions is largely consistent with what Garrett would hope to see, the question is how *Convicting the Innocent* advances the reform agenda beyond simply exhorting less ambitious jurisdictions to step up to the plate. With the issues that Garrett tackles already on the radar screen of the major players in a conceptual sense, the current juncture calls for drilling down on the details—both substantive, such as the specifics of best practices that should be adopted, and strategic, such as whether reformers should be concentrating their efforts on legislatures, courts, the voluntary goodwill of law enforcement, or elsewhere. Furthermore, any change will occur in a context of limited political and financial capital, and in a climate where most defense-friendly reforms are the product of hard-fought political compromise with powerful opposing interests. Therefore, the path forward calls for principled and information-driven decisions about the inevitable prioritizations and trade-offs that will

115. See *supra* notes 22–26 and accompanying text.

116. GARRETT, *supra* note 8, at 211.

117. *Id.* at 252.

118. *Id.* at 241–62. See generally William J. Stuntz, *The Uneasy Relationship Between Criminal Procedure and Criminal Justice*, 107 YALE L.J. 1 (1997) (describing legislative incentives toward easing rather than heightening the state's burden in prosecution).

be necessitated, and for consideration of a range of second-best alternatives to the full menu of regulation and reform that Garrett would ideally envision. Unfortunately, in three critical respects, *Convicting the Innocent* sidesteps clearly presented opportunities to wade more deeply into those most relevant currents.

First, Garrett's account of reforms already undertaken eschews the pointillist technique that makes his descriptive account of the problem so rich and instead paints with a decidedly broad brush. He aims, of course, simply to give an aggregate sense of the terrain. But a consequence is that he glosses over variations that are highly salient within the terms of his own framework for analysis and critique. So, for example, Garrett points to the fact that 18 states and the District of Columbia, along with some 500 police departments, require or encourage full or partial recording of interrogations.¹¹⁹ But an important lesson of Garrett's study is that *partial* recording of interrogations is likely not only to be insufficient but also counterproductive in ensuring that only voluntary and reliable statements are taken: permitting police to interview or question a suspect "off tape" and then record subsequent statements by the suspect—a situation that occurred in the Deskovic case discussed above¹²⁰—risks generating a record that strengthens the credibility of the final product via "contamination" and insulates contamination from scrutiny.

Similarly, in the context of eyewitness identification, Garrett identifies several states that have enacted reform either by judicial action, statute, agency decision, or some combination thereof. But the book does not discuss important differences among these jurisdictions. Thus, New Jersey's path toward "landmark reform" is described in detail, from the state attorney general's promulgation of guidelines requiring police departments to adopt identification procedures tracking the current social science research on eyewitness fallibility, to the New Jersey Supreme Court's subsequent moves toward requiring electronic recording of identification procedures and cautionary jury instructions, and then implementing wholesale revision of judicial eyewitness-identification doctrine.¹²¹ We then learn that six additional states have passed statutes in response to misidentification. But Garrett's shout-out to these jurisdictions does not disclose that among them only North Carolina approaches New Jersey's level of comprehensiveness, while West Virginia and Illinois declined to address the critical issue of how witnesses view and select suspects during a lineup—a core feature of Garrett's recommendations.¹²² Had Garrett's book gone to press just months

119. See GARRETT, *supra* note 8, at 248.

120. See *supra* notes 87–95 and accompanying text.

121. See GARRETT, *supra* note 8, at 250–51.

122. *Id.* at 248–52; see 725 ILL. COMP. STAT. ANN. 5/107A-5 (West 2006) (detailing a lineup and photo-spread procedure that omits instructions on how witnesses view and select suspects); N.C. GEN. STAT. § 15A-284.52 (2009) (specifying extensive lineup procedures, including how witnesses are to view and select suspects); W. VA. CODE ANN. § 62-1E-2 (LexisNexis 2010)

later, it could also have noted the example of Texas's legislatively enacted reforms. Consistent with Garrett's recommendations, Texas now requires police departments to develop and adopt "a detailed written policy regarding the administration of photograph and live lineup identification procedures" and further requires that those policies be "based on . . . credible field, academic, or laboratory research on eyewitness memory" and "best practices designed to reduce erroneous eyewitness identifications."¹²³ Unlike New Jersey, however, the details are left ultimately to the discretion of individual departments, and admissibility of eyewitness testimony is expressly *not* contingent on compliance with a department's adopted policies.¹²⁴

It seems clear that Garrett does not view all of these approaches as equally exemplary. However, he does not say so, and in presenting them without elaboration or critique he misleadingly advances incomplete reform efforts as models. On the other hand, it may be that Garrett is willing to concede that incremental measures short of the ideal may be tolerable—as a matter of the social science research, as a matter of political feasibility, or for some combination of these or other reasons. In that case, the relevant trade-offs are well worth discussing. In either event, there is a missed opportunity to provide more nuanced guidance for a path forward.

A closely related criticism stems from Garrett's inattention to the puzzle of what conditions are necessary to spur and sustain reform of the sort that he seeks. Garrett asserts that "exonerations"—and DNA exonerations in particular—"are reshaping criminal procedure."¹²⁵ One might think the path forward is as "simple" as generating dramatic revelations of error and then standing and watching the reform percolate (in which case the most direct path to reform might be greater access to postconviction innocence review, a notion that Garrett does entertain¹²⁶). But of course this is not the case.

Thus, we learn that a spate of seven DNA exonerations pushed North Carolina down a radical path, leading first to the creation of a permanent body to investigate wrongful convictions and propose responsive systemic

(itemizing the state's eyewitness-identification procedures—a set of procedures that does not include a standardized process through which witnesses view and select suspects during a lineup). Both Illinois and West Virginia authorized further study on these issues. See INNOCENCE PROJECT, REEVALUATING LINEUPS: WHY WITNESSES MAKE MISTAKES AND HOW TO REDUCE THE CHANCE OF A MISIDENTIFICATION 23 (2009), available at http://www.innocenceproject.org/docs/Eyewitness_ID_Report.pdf (finding that West Virginia created a task force in 2007 to "study and identify additional best practices for eyewitness identification"); SHERI H. MECKLENBURG, ILL. STATE POLICE, REPORT TO THE LEGISLATURE OF THE STATE OF ILLINOIS: THE ILLINOIS PILOT PROGRAM ON SEQUENTIAL DOUBLE-BLIND IDENTIFICATION PROCEDURES 8-9 (2006), available at <http://www.chicagopolice.org/IL%20Pilot%20on%20Eyewitness%20ID.pdf> (stating that the Illinois Legislature mandated a "Pilot Study on 'the effectiveness of the sequential method for photograph and live lineup procedures'"). To my knowledge, neither jurisdiction has acted on that research.

123. TEX. CRIM. PROC. CODE ANN. art. 38.20, § 3(a), (c) (West Supp. 2011).

124. *Id.* § 5(b).

125. GARRETT, *supra* note 8, at 244.

126. See *id.* at 239, 241-44 (advocating for expansion of postconviction DNA testing and endorsing innocence commissions as a path to greater exoneration opportunities).

reforms, and subsequently to the formation of the Innocence Inquiry Commission to review prisoners' claims of innocence and make recommendations for exoneration outside the procedural limitations imposed by ordinary judicial review.¹²⁷ In Texas, by contrast, the response to forty-four DNA exonerations has been far more halting¹²⁸: a statutorily created body to review the causes of wrongful conviction has generated eleven recommendations largely tracking Garrett's own;¹²⁹ only one, eyewitness-identification reform, has been acted upon by the legislature.¹³⁰ Are there intractable differences that account for this disparity in the track record of reform? Are there any lessons to be gleaned from the North Carolina example and translated to the distinctive political and institutional contexts presented by other jurisdictions?

There are occasions, too, when Garrett does not simply sidestep but in fact glosses over fascinating and important strategic dynamics. For example, in the context of discussing eyewitness-identification reform, he asserts that following a spate of exonerations, "New Jersey began a project of revamping its criminal procedure rules."¹³¹ The statement suggests a kind of deliberate and consensus-driven effort when in fact the substance, course, and pace of reform in New Jersey were quite contested—particularly between prosecutors and law enforcement on the one hand and the trailblazing judiciary on the other.¹³² Though full exploration of the dynamics and trajectory of these conflicts is certainly beyond Garrett's inquiry, recognition of them, at least, is called for in the context of his otherwise-nuanced account.

These are difficult questions, but they are as susceptible to interrogation as the difficult systemic dynamics that much of *Convicting the Innocent* is devoted to untangling. The generality of Garrett's narrative on this front must not cloud the critical presence of these issues on the radar screens of scholars and reformers. Those who seek the overhaul of criminal justice practices in response to revealed systemic deficiencies in reliability are not

127. *Id.* at 241–44.

128. *Exonerations by State*, INNOCENCE PROJECT, <http://www.innocenceproject.org/news/StateView.php>.

129. TIMOTHY COLE ADVISORY PANEL ON WRONGFUL CONVICTIONS, REPORT TO THE TEXAS TASK FORCE ON INDIGENT DEFENSE, at ii (Aug. 2010), *available at* <http://www.courts.state.tx.us/tfdc/pdf/FINALTCAPreport.pdf>. More radical reform took place voluntarily in Dallas County, the locality that has produced the overwhelming bulk of Texas's exonerations, and which has now institutionalized open-file discovery, cold-case review, and a variety of investigative reforms as a matter of district-attorney policy. GARRETT, *supra* note 8, at 259.

130. Tex. H.B. 215, 82d Leg., R.S. (2011); *see supra* notes 123–24 and accompanying text.

131. GARRETT, *supra* note 8, at 250.

132. *See, e.g.*, State v. Henderson, 27 A.3d 872, 884–85, 912–15 (2011) (recounting that notwithstanding the Attorney General's adoption of eyewitness-identification guidelines, the state "argue[d] vigorously" against judicial imposition of a "presumption of impermissible suggestiveness" for breach of guidelines and likewise opposed judicial revision of eyewitness-admissibility factors and maintained opposition through several rounds of litigation).

short on ideas but rather on tactics, and application of Garrett's analytical rigor to that pressing challenge would have been welcome.

But the most significant hole in *Convicting the Innocent's* prescription for reform may lie in its too-offhanded treatment of courts and legal doctrine. To be sure, this is not an inadvertent oversight. One of the lessons of Garrett's diagnostic account is that adjudication can serve only as a "backstop" accompanying direct reform of the primarily investigative practices that generate error, and so it is to that latter task that Garrett directs his primary prescriptive attention.¹³³ Moreover, it is understandable that Garrett would aim to push back on the tendency of legal scholarship to be overly attuned to the work of courts and insufficiently attuned to the work of other institutions that critically shape legal outcomes. In this sense, *Convicting the Innocent* is of a piece with, and indeed advances, a new generation of criminal law scholarship that has questioned the descriptive accuracy and normative desirability of court-mediated, procedurally oriented notions of criminal "justice" inherited from the Warren Court's constitutional-criminal-procedure legacy.¹³⁴ Garrett's analysis of the tendency of error to evade judicial detection provides further reason to reject the centrality of adjudicatory mechanisms in understanding the most salient dynamics of criminal justice.

But even as peripheral players, it is important that the contribution of courts be synergistic with—or at a minimum not undermining of—upstream reliability-enhancing reforms. It is this sentiment that leads Garrett to urge that courts refocus the lens of criminal procedure doctrine to screen critical evidence—confessions, identifications, informant testimony, and expert testimony—for substantive reliability.¹³⁵ In this respect, again, Garrett's proposal is resonant with the remedial agenda around which much of the innocence-focused reform movement has already coalesced¹³⁶—though for the reasons set forth in Part I of this Review, *Convicting the Innocent* provides an important and persuasive grounding of that call in the empirical realities of the criminal process. But if the final chapter of *Convicting the Innocent* makes a plausible case for the feasibility of even seemingly radical

133. GARRETT, *supra* note 8, at 211.

134. See, e.g., Daniel Richman, *Prosecutors and Their Agents, Agents and Their Prosecutors*, 103 COLUM. L. REV. 749, 750–51 (2003) (describing the "growing recognition that the road to criminal justice reform lies not through the battleground of defendant rights . . . but through attention to . . . 'administrative-inquisitorial structures that in fact process most American criminal cases'" (quoting Gerard E. Lynch, *Our Administrative System of Criminal Justice*, 66 FORDHAM L. REV. 2117, 2151 (1998))); William J. Stuntz, *The Political Constitution of Criminal Justice*, 119 HARV. L. REV. 780, 818–19 (2006) (arguing that Warren Court criminal procedure doctrine had perverse effects on substantive accuracy).

135. GARRETT, *supra* note 8, at 248, 252, 255–56.

136. See, e.g., Keith A. Findley, *Toward a New Paradigm of Criminal Justice: How the Innocence Movement Merges Crime Control and Due Process*, 41 TEX. TECH L. REV. 133, 134, 147–72 (2008) (reviewing scholarship and reform proposals that reflect how "the Innocence Movement alters our understanding of the criminal justice system by giving us a new paradigm—a Reliability Model based on best practices").

investigative reforms, the same is not so in regard to this proposed judicial agenda. Garrett himself concedes that two decades of DNA exonerations have left us still with a Supreme Court that seems “complacent” in the face of evidence that should alarm.¹³⁷ A number of data points from recent Supreme Court history suggest that this characterization is too generous.

In fact, the Court has all but affirmatively rejected the perspective that Garrett urges and has done so in the face of express invitations (including through amicus briefs from the Innocence Project and related organizations¹³⁸) to consider whether the phenomenon of DNA exoneration challenges the normal presumption that provision of fair procedures rather than substantive evidentiary scrutiny is adequate to ensure constitutionally fair criminal adjudication. Two terms ago, in *District Attorney’s Office for the Third Judicial District v. Osborne*, the Court rejected the due process claim of a convicted prisoner challenging Alaska’s refusal to permit him access to available biological evidence in order to substantiate his claimed innocence with DNA testing.¹³⁹ The Court squarely resisted the premise that DNA evidence and its capacity to conclusively resolve questions of factual truth should upend the long-standing value of legal finality or the long-standing aversion of federal courts to second-guessing the adequacy of state criminal procedures.¹⁴⁰ More recently, and more relevantly, this term in *Perry v. New Hampshire*,¹⁴¹ the Court resoundingly, by an eight-to-one margin, ruled against a defendant’s claim that the Constitution demands inquiry into the reliability of eyewitness-identification evidence obtained under unreliable conditions that were not orchestrated by the police.¹⁴² Justice Ginsburg’s opinion for the majority squarely rejected the notion of a constitutional principle of reliability in a context—that of eyewitness identification—where the innocence movement has arguably made the most headway in generating data to support the dangers of judicial permissiveness toward admissibility.

In light of these fairly clear signals that constitutional-criminal-procedure doctrine will meet the challenge of DNA and innocence as, in the spirit of the epigraphical quote from *Osborne*,¹⁴³ a management problem rather than a revolution, there is an emptiness to Garrett’s call. To urge a wholesale judicial rethinking of doctrinal premises in a climate with such dim prospects for success risks leaving reformers empty-handed with respect

137. GARRETT, *supra* note 8, at 262, 269.

138. *E.g.*, Brief of Amicus Curiae The Innocence Network in Support of Petitioner, Supporting Reversal at 7–10, *Perry v. New Hampshire*, 132 S. Ct. 716 (2012) (No. 10-8974), 2011 WL 3439922; Brief for the Respondent at 48, 51–52, *Dist. Attorney’s Office for the Third Judicial Dist. v. Osborne*, 129 S. Ct. 2308 (2009) (No. 08-6), 2009 WL 208117.

139. 129 S. Ct. at 2315, 2323.

140. *Id.* at 2322.

141. 132 S. Ct. 716 (2012).

142. *Id.* at 730.

143. *See supra* text accompanying note 2.

to judicial reform. There are, however, ways in which Garrett's analysis might inform doctrinal innovations that could feasibly be urged even within the context of the prevailing conservatism of courts vis-à-vis legal responses to accuracy concerns raised by wrongful convictions.

Consider, for example, the doctrine of structural error, which in federal habeas proceedings exempts certain fundamental trial deficiencies from the most stringent level of harmless-error review.¹⁴⁴ Currently, these structural errors are found only in a "very limited class of cases"¹⁴⁵ reflecting "defect[s] affecting the framework within which the trial proceeds, rather than simply an error in the trial process itself."¹⁴⁶ By contrast, as to the vast majority of constitutional claims in postconviction proceedings, the state has the opportunity to establish that any error was "harmless beyond a reasonable doubt"¹⁴⁷—a test that, under any standard, is often met by courts pointing to substantively convincing evidence of guilt. Indeed, Garrett identifies harmless-error doctrine as a mechanism by which investigative contamination is currently occluded, and reliability-suppressing missteps are left unremedied in the postconviction process.¹⁴⁸ But Garrett's analysis of the systemic nature of error also points to a potential wedge to expand structural-error doctrine. The dynamics of contamination that Garrett documents and describes arguably create "defect[s] [that] affect[] the framework within which the trial proceeds,"¹⁴⁹ either categorically or in particular cases.¹⁵⁰ Garrett's analysis thus suggests, and provides an avenue for arguing to courts, that certain categories of error currently treated as potentially "harmless" might in fact be properly analyzed as structural error, thus permitting greater accuracy-based scrutiny in postconviction review.¹⁵¹

Second, and thinking more strategically than doctrinally, Garrett's account points to ways in which criminal justice reformers can, have, and should focus on the courts as part of a coordinated strategy that includes other actors who generate criminal justice policy. The history of eyewitness-identification evidence reform in New Jersey is exemplary. There, reform began internally to law enforcement, with the attorney general mandating that law enforcement agencies adopt identification-procedure policies, though specifying that its mandate "should in no way be used to imply that

144. *Fulminante v. Arizona*, 499 U.S. 279, 308–10 (1991).

145. *Johnson v. United States*, 520 U.S. 461, 468 (1997); see also *Sullivan v. Louisiana*, 508 U.S. 275, 281 (1993) (finding that an error in a reasonable-doubt instruction was a structural error).

146. *Fulminante*, 499 U.S. at 310.

147. *Id.* at 295 (quoting *Chapman v. California*, 386 U.S. 18, 24 (1967)).

148. GARRETT, *supra* note 8, at 200–02.

149. *Fulminante*, 499 U.S. at 310.

150. *Cf. Commonwealth v. Durand*, 931 N.E.2d 950, 966 (Mass. 2010) (reserving the question of whether coercion of a confession is structural error under the Massachusetts constitution).

151. Garrett himself has gestured toward this possibility in prior work. See Brandon L. Garrett, *Aggregation in Criminal Law*, 95 CALIF. L. REV. 383, 422–24 (2007) (arguing that systemic misconduct is wrongly conceptualized within individualized harmless-error frameworks).

identifications made without [the] procedures are inadmissible or otherwise in error.”¹⁵² The New Jersey Supreme Court subsequently, over objection from the State, exercised its supervisory authority to require recording of identification procedures as a condition of admissibility¹⁵³ and later to adopt model instructions cautioning jurors on the reliability of eyewitness identification.¹⁵⁴ Critically, the court noted and commended the laudable steps taken by the attorney general but concluded that its own duty and authority permitted it to supplement (and enhance) oversight in this realm.¹⁵⁵ Of course, the New Jersey Supreme Court ultimately went even further and, following the appointment of a special master, issued a decision that “established a comprehensive social science framework for regulating eyewitness identifications in the courtroom”¹⁵⁶—i.e., just the sort of fundamental reliability-based reform that Garrett would (rightly) urge be more broadly pursued. But even if New Jersey had stopped short in its more incremental steps of requiring recording and jury charges, it would have been an important judicial advancement and reinforcement of more limited and piecemeal reform undertaken by other institutions.

Interestingly, there is evidence that a similar dynamic might be at play in Texas’s still-evolving experiment with eyewitness-identification reform. Subsequent to the legislature’s adoption of an eyewitness-identification bill,¹⁵⁷ the Court of Criminal Appeals issued its first decision reversing a conviction based on the exclusion of a defense expert on eyewitness identifications—citing not only the New Jersey Supreme Court and Garrett but also, more to the point, the Texas Legislature’s then-recent enactment.¹⁵⁸ Indeed, some have argued that in light of the legislature’s failure to provide for a sanction for departments that fail to comply with the mandated eyewitness-identification policies, the Court of Criminal Appeals’ decision amounts to a critical oversight mechanism, providing a previously unavailable opportunity for defendants to present evidence concerning the importance of compliance with best practices in identification procedures.¹⁵⁹

152. Letter from John J. Farmer, Jr., Attorney Gen., State of N.J., to All County Prosecutors et al. 3 (Apr. 18, 2001), available at http://www.innocenceproject.org/docs/NJ_eyewitness.pdf.

153. *State v. Delgado*, 902 A.2d 888, 896–97 (N.J. 2006).

154. *State v. Romero*, 922 A.2d 693, 702–03 (N.J. 2007).

155. *Id.* at 702; *Delgado*, 902 A.2d at 896.

156. Brandon L. Garrett, *Trial and Error: Learning from Patterns of Mistakes*, 26 CRIM. JUST. 30, 35 (2012) (citing *State v. Henderson*, 27 A.3d 872 (2011)).

157. See *supra* notes 123–24, 130 and accompanying text.

158. *Tillman v. State*, 354 S.W.3d 425, 436 (Tex. Crim. App. 2011) (citing *Henderson*, 27 A.3d 872; GARRETT, *supra* note 8, at 8–9, 279); *id.* at 442 (citing TEX. CRIM. PROC. CODE art. 38.20, § 1).

159. See *Chiefs’ Push to Weaken Eyewitness ID Improvements Could Boost Market for Defense Experts*, GRITS FOR BREAKFAST (Jan. 30, 2012), <http://gritsforbreakfast.blogspot.com/2012/01/chiefs-wish-to-weaken-eyewitness-id.html> (arguing that *Tillman* created a penalty to incentivize police departments to conform with the best practices in identification procedures by allowing expert witnesses for defendants if they failed to comply); see also *Tillman*, 354 S.W.3d at 442

The critical insight in both the New Jersey and Texas examples is that something less than wholesale judicial reassessment of constitutional or evidentiary doctrine can, in concert with movement by other criminal justice stakeholders, build upon or reinforce other steps toward reform. In this respect, it is worth highlighting that the most effective site of focus is likely to be the oft-neglected (at least in scholarly accounts) state courts rather than their federal counterparts. Indeed, state courts have proven more receptive to broader conceptualization of structural error along precisely the dimensions described above.¹⁶⁰ But as Garrett's analysis reminds us, because the difficulties of error detection inevitably grow more intractable over the lifetime of a criminal case, accuracy-enhancing doctrine will have the greatest effect at the earliest possible stages of litigation—which, in the overwhelming majority of instances, occurs in state court.

These are just two examples, sketched at high levels of generality, of ways in which Garrett's diagnostic account might helpfully guide court-based reform efforts along nonstandard lines. Given Garrett's past scholarship exploring innovative jurisprudential reform in the arenas of criminal and civil procedure,¹⁶¹ I have no doubt that *Convicting the Innocent* and the debates it seeks to influence could have benefitted substantially from deployment of Garrett's considerable creativity and lawyerly chops to the project of moving beyond what seems increasingly to be a moribund reliability-based doctrinal-reform agenda.

Conclusion

Garrett asks, "Should we be pessimistic or optimistic about actually fixing the flaws in our criminal justice system?"¹⁶² He does not directly answer the question, and *Convicting the Innocent* is replete with well-founded ambivalence on this score. In the final analysis, Garrett adopts a distinctively hopeful tone—though one that at times may lead him to be too unqualified in presenting the compelling evidence he adduces of broad-based failings in our system of criminal adjudication and too confident that the lessons of DNA exonerations will ultimately advance the practical agenda for criminal justice reform. But these limitations do not diminish the significant and timely contribution that *Convicting the Innocent* makes. Indeed, it is *because* Garrett is correct that we are at a crossroads in charting a path forward from internalizing the potential for error in criminal adjudication,

(noting that expert witnesses would be able to testify concerning law enforcement compliance with standard eyewitness-identification procedures).

160. See, e.g., *Connecticut v. Johnson*, 460 U.S. 73, 91 (1983) (Powell, J., dissenting) ("A State, of course, may apply a more stringent *state* harmless-error rule than *Chapman* would require.").

161. See, e.g., Garrett, *supra* note 151, at 424–49 (developing models of aggregated claim adjudication in criminal law); Brandon L. Garrett, *Claiming Innocence*, 92 MINN. L. REV. 1629, 1699–715 (2008) (developing an analytical framework for a constitutional innocence claim).

162. GARRETT, *supra* note 8, at 273.

and *because* of breadth, quality, and stature of *Convicting the Innocent*, that there is at times a worrisome sense that the effort might not intervene in the march toward reform at the most relevant point. Nevertheless, Garrett's contributions are significant. His elucidation of the machinery of criminal justice error and the extent to which we are systemically compromised in correcting the mistakes it generates is illuminating, paradigm shifting, and generative of further questions that are now prominently positioned for future scholarship to probe. Ultimately, it is work that admirably carries Borchard's torch into a new era of criminal justice debate.

* * *

Constructing a Canon of Law-Related Poetry

POETRY OF THE LAW: FROM CHAUCER TO THE PRESENT. Edited by David Kader & Michael Stanford. Iowa City, Iowa: University of Iowa Press, 2010. 200 pages. \$22.00.

Reviewed by Alexandra J. Roberts*

I. Introduction

Law and poetry make a potent, if surprising, pair. Poetic language thrives on simultaneity and open-endedness, while legal language seeks resolution and closure. The law-and-literature movement, since its inception in the early 1970s,¹ has given rise to scores of articles, books, symposia, and specialized journals, and most U.S. law schools now offer at least one course on the topic.² Most of the scholarship it spawns has either focused on law *as* literature³—applying rhetorical tools to judicial opinions and legislation—or

* Litigation Associate, Ropes & Gray, LLP. The author would like to thank Ian Ayres, A. Mitchell Polinsky, Daniel B. Roberts, Judith Resnik, and David A. Skeel for their comments and suggestions on earlier drafts, Eric Goralnick for his support, and the editors at the *Texas Law Review* for their expert revisions.

1. Some scholars wrote on law and literature earlier. *E.g.*, Bond Almand, *Law, Language, and Literature*, 14 *MERCER L. REV.* 372 (1963); Henry B. Cushing & Ernest F. Roberts, *Law and Literature: The Contemporary Image of the Lawyer*, 6 *VILL. L. REV.* 451 (1961). However, most trace the birth of law and literature as a field of study to JAMES B. WHITE, *THE LEGAL IMAGINATION* (1973).

2. Jeanne Gaakeer, *Law and Literature*, *IVR ENCYCLOPEDIA JURISPRUDENCE, LEGAL THEORY & PHIL. L.*, http://ivr-enc.info/index.php?title=Law_and_Literature.

3. The law-and-literature movement in its current incarnation can be divided into several strands. The humanist or “moral uplift” approach argues that to become a good lawyer and a good citizen, a law student needs to learn “how to be a human being capable of love and imagination.” MARTHA C. NUSSBAUM, *CULTIVATING HUMANITY: A CLASSICAL DEFENSE OF REFORM IN LIBERAL EDUCATION* 13–14 (1997). Proponents of this theory argue that literature concretizes the abstract, enabling lawyers to better understand human nature, learn from others’ experiences, engage in moral decision making, and incorporate emotion and intuition with more traditional modes of legal logic. Jennifer Bard et al., *Three Ways of Looking at a Health Law and Literature Class*, 1 *DREXEL L. REV.* 512, 527–28 (2009) (citing Harvey Couch, III, *Law and Literature—A Comment*, 17 *VAND. L. REV.* 911, 914 (1964)). The second strand, the hermeneutic approach, focuses not on literature but on literary theory and criticism. *Id.* at 528. This area of inquiry is often called “law as literature” to distinguish it from the study of “law in literature” because it treats laws and judicial opinions as the texts to which literary theory can be fruitfully applied. *See, e.g.*, Martin A. Kayman, *Law-and-Literature: Questions of Jurisdiction*, in 18 *REAL YEARBOOK OF RESEARCH IN ENGLISH AND AMERICAN LITERATURE: LAW AND LITERATURE* 1, 7–8 (Brook Thomas & Gunter Narr eds., 2002) (stating that the essay title in Sanford Levinson’s *Law as Literature* posits a hermeneutical relationship between law and literature, and distinguishing the word choice from law *in* literature). A third strand of law and literature uses the narratives and stories of individuals to better understand how law functions in “real life,” to give voice to often-marginalized groups, and to glean how such stories can provide evidence or present conflicting

law in literature, considering law as portrayed in fiction, nonfiction, theater, and film.⁴ Poetry has garnered scant attention,⁵ and most discussions of law and poetry have focused exclusively on the body of work of a single lawyer-poet such as Wallace Stevens or Lawrence Joseph.⁶ Some theorists have

versions of truth. See Bard, *supra*, at 528–29 (discussing the evidentiary and epistemological purposes of the third approach).

4. See, e.g., Barbara Johnson, *Anthropomorphism in Lyric and Law*, 10 YALE J.L. & HUMAN. 549, 550 (1998) (purporting to use the concepts of “person” in poetry and law to illuminate each other); see also Elizabeth Villiers Gemmette, *Law and Literature: Joining the Class Action*, 29 VAL. U. L. REV. 665, 666, 686 (1995) (discussing the law-and-literature canon based on her survey of law schools, initially conducted in 1987 and repeated in 1994). Gemmette’s 1987 survey found 28% of responding schools offered a course in law and literature, the 1994 survey found that figure had increased to 43%, and a comparable study in 2011 would reflect further increase. *Id.* at 666; see also ASS’N AM. LAW SCH., DIRECTORY OF LAW TEACHERS: 2010–2011, at 1721–23 (2010) (listing 162 law professors reporting that they taught Law and Literature or a similarly titled course). Reviewing the results from her second survey based on responses from eighty-four professors, Gemmette points out that of the eighty-four courses on law and literature, “eighteen courses utilize fiction only[;] . . . thirty-seven courses utilize both fiction and critical works[;] . . . two courses are structured around James Boyd White’s [b]ook *The Legal Imagination*[;] one course stresses hermeneutics and utilizes, among other texts, a hermeneutic reader[;] and one course utilizes a reader on storytelling.” Gemmette, *supra*, at 670 (footnotes omitted). None of the courses appears to devote significant time to law and poetry, and the syllabi Gemmette collected include only a few poems. See *id.* app. B at 795–813 (listing in the aggregate fewer than one poem per professor in the syllabi).

5. See POETRY OF THE LAW: FROM CHAUCER TO THE PRESENT xiii (David Kader & Michael Stanford eds., 2010) [hereinafter POETRY OF THE LAW] (“[F]or all the richness of [law-and-literature] scholarship, it has focused almost entirely on fiction and drama.”). When this Review uses the term *poetry*, it excepts drama, which has typically been treated separately and more extensively in law-and-literature scholarship than has short poetry. A substantial amount of scholarship has focused on law and lawyers in several of Shakespeare’s plays. For example, dozens of articles exploring the roles of Portia and Shylock in *The Merchant of Venice* have appeared in law reviews. E.g., Anita L. Allen & Michael R. Seidl, *Cross-Cultural Commerce in Shakespeare’s The Merchant of Venice*, 10 AM. U. J. INT’L L. & POL’Y 837 (1995); Thomas C. Bilello, *Accomplished with What She Lacks: Law, Equity, and Portia’s Con*, 16 LAW & LITERATURE 11 (2004); Daniela Carpi, *Law, Discretion, Equity in The Merchant of Venice and Measure for Measure*, 26 CARDOZO L. REV. 2317 (2005); Christopher A. Colmo, *Law and Love in Shakespeare’s The Merchant of Venice*, 26 OKLA. CITY U. L. REV. 307 (2001); Edith Z. Friedler, *Shakespeare’s Contribution to the Teaching of Comparative Law—Some Reflections on The Merchant of Venice*, 60 LA. L. REV. 1087 (2000); Geoffrey Hartman, *The Trickster Word: Richard Weisberg on The Merchant of Venice*, 23 LAW & LITERATURE 71 (2011); Jonathan M. Hyman & Lela P. Love, *If Portia Were a Mediator: An Inquiry into Justice in Mediation*, 9 CLINICAL L. REV. 157 (2002); Randy Lee, *Who’s Afraid of William Shakespeare?: Confronting our Concepts of Justice and Mercy in The Merchant of Venice*, 32 U. DAYTON L. REV. 1 (2006); Daniel H. Lowenstein, *The Failure of the Act: Conceptions of Law in The Merchant of Venice*, Bleak House, Les Misérables, and Richard Weisberg’s Poethics, 15 CARDOZO L. REV. 1139 (1994); Ken Masugi, *Race, the Rule of Law, and The Merchant of Venice: From Slavery to Citizenship*, 11 NOTRE DAME J.L. ETHICS & PUB. POL’Y 197 (1997); Trisha Olson, *Pausing upon Portia*, 19 J.L. & RELIGION 299 (2004); Michael Jay Willson, *A View of Justice in Shakespeare’s The Merchant of Venice and Measure for Measure*, 70 NOTRE DAME L. REV. 695 (1995); Kenji Yoshino, *The Lawyer of Belmont*, 9 YALE J.L. & HUMAN. 183 (1997); Erin A. Cook, Comment, *Shining Lights at the Bar: Shakespeare’s Portia as a Model for Female Attorneys*, 30 CUMB. L. REV. 517 (2000).

6. See, e.g., THOMAS C. GREY, *THE WALLACE STEVENS CASE: LAW AND THE PRACTICE OF POETRY* (1991) (discussing the poetry of Wallace Stevens); ROBERT G. LAMBERT, JR., *EMILY DICKINSON’S USE OF ANGLO-AMERICAN LEGAL CONCEPTS AND VOCABULARY IN HER POETRY: MUSE AT THE BAR* (1997) (discussing the poetry of Emily Dickinson); RICHARD A. POSNER, *LAW*

blamed the paucity of scholarship about law and poetry on “the dearth of poetry about law.”⁷ In the text used most often in law school courses on law and literature, Richard Posner opines that “[r]elatively few short poems take law as their theme.”⁸

Yet, a recently published anthology of one hundred “law-related poems” entitled *Poetry of the Law: From Chaucer to the Present*, edited by law professor David Kader and attorney and poet Michael Stanford,⁹ belies Posner’s claim. This Review considers the place of poetry in legal studies and advocates incorporating poetry into both the dialogue and the curriculum of the law-and-literature movement.¹⁰ It identifies themes in the poems of Kader and Stanford’s anthology, explores the relationship of fixed-verse forms to law in several of those poems, and draws attention to those voices that have been silenced both in this collection and in the study of law and literature generally. This Review relies primarily on the process of close reading,¹¹ and in so doing, it practices law *in* literature while it models precisely the type of critical approach that would well serve those participating in the study of law *as* literature.

The editors of *Poetry of the Law* characterize as law-related “a poem with a legal setting (the courtroom, the lawyer’s office, the judge’s chambers,

AND LITERATURE 191–94 (3d ed. 2009) (discussing the poetry of Wallace Stevens); David A. Skeel, Jr., *Lawrence Joseph and Law and Literature*, 77 U. CIN. L. REV. 921 (2009) [hereinafter Skeel, *Lawrence Joseph*] (discussing the poetry of Lawrence Joseph); David A. Skeel, Jr., *Practicing Poetry, Teaching Law*, 92 MICH. L. REV. 1754 (1994) [hereinafter Skeel, *Practicing Poetry*] (reviewing LAWRENCE JOSEPH, *BEFORE OUR EYES* (1993)).

7. Skeel, *Practicing Poetry*, *supra* note 6, at 1757.

8. POSNER, *supra* note 6, at 191 n.32. David Ray Papke offers one explanation for the popularity and apparent authoritativeness of Posner’s book despite its author’s “sniping” approach to the law-and-literature movement: “Since much of Posner’s work is, in effect, highly combative tertiary scholarship, his comments on other law-and-literature scholars constitute an important annotated roster of the movement.” David Ray Papke, *Problems with an Uninvited Guest: Richard A. Posner and the Law and Literature Movement*, 69 B.U. L. REV. 1067, 1077 (1989) (reviewing RICHARD A. POSNER, *LAW & LITERATURE: A MISUNDERSTOOD RELATION* (1988)).

9. See generally POETRY OF THE LAW, *supra* note 5.

10. It bears noting that James Boyd White’s *The Legal Imagination*, a seminal text of the law-and-literature movement, draws heavily upon the work of “New Criticism,” a theoretical approach applied frequently to poetry. See Jeffrey Malkan, *Law on a Darkling Plain*, 101 HARV. L. REV. 702, 711 n.18 (1988) (book review) (discussing New Criticism); see also Michael Pantazakos, *A Humanitarian Pertinent: A Personal Reflection on the History and Purpose of the Law and Literature Movement*, 7 CARDOZO STUD. L. & LITERATURE 31, 39 (1995) (noting that most scholars concur that the publishing of *The Legal Imagination* heralded the beginning of law and literature as a distinct jurisprudential discipline). Nonetheless, poetry very rarely figures in law-and-literature analysis.

11. “Close reading” (or *explication de texte*) refers to a method of literary criticism characterized by specific, detailed interpretation of short texts or portions of texts. ROSS MURFIN & SUPRYIA M. RAY, *THE BEDFORD GLOSSARY OF CRITICAL AND LITERARY TERMS* 147 (2d ed. 2003); see also RICHARD A. POSNER, *LAW & LITERATURE* 221 (rev. and enlarged ed. 1998) (“The New Critics were committed to the close reading of works of literature . . .”). Close reading emphasizes the particular over the general, focusing on vocabulary, syntax, theme, tone, prosody, literary devices, and any other features the reader deems noteworthy. MURFIN & RAY, *supra*, at 293.

the law-school classroom) or a poem largely organized around legal issues, concepts, metaphors, or language.”¹² The poems in this anthology make a compelling case for the inclusion of poetry in the law-and-literature curriculum and demonstrate what a canon of poetry for that curriculum might include. In addition, because many law-related poems are self-contained works that offer great depth without great length, law students can read fifty poems in the time they might have read just one novel;¹³ the study of law-related poetry thus lends itself exceedingly well to the format of a single-semester law school course.¹⁴ If law-related poems are to be incorporated into the syllabi of Law and Literature courses, it behooves us to consider what the poems teach us in isolation and what patterns and tropes emerge when we read them in juxtaposition.

On the first page of their introduction, Kader and Stanford declare that *Poetry of the Law* “fill[s] a striking gap” as “the first anthology of poetry about the law to be published in the United States¹⁵ in half a century and the first *selective* anthology on the subject ever published.”¹⁶ Though their assertion is either redundant or self-canceling—an anthology is, by definition, selective—*Poetry of the Law* indeed fills a gap. The editors categorize the poems as “sort[ing] themselves into roughly six overlapping categories”: (1) poems about lawyers and judges, (2) poems about the citizen in the legal system, (3) poems about historical trials, (4) poems about punishment, (5) poems exploring legal concepts, and (6) poems applying legal metaphors to nonlegal subjects.¹⁷ The poems are not arranged by category, however, but appear chronologically by the date of birth of each poem’s author. Although not without flaws,¹⁸ the anthology contains powerful works from a

12. POETRY OF THE LAW, *supra* note 5, at xv.

13. See, e.g., Sara D. Schotland, *Justice for Undergraduates: Teaching Law and Literature in the Liberal Arts Curriculum*, CURRENTS IN TEACHING & LEARNING, Fall 2009, at 41, 46 (acknowledging that professors might “find it impractical to include *Crime and Punishment* or *Bleak House* on their syllabus because of their length”).

14. Poems also lend themselves to academic legal writing better than do long works of prose because a set of poems can be incorporated wholesale into a law review article or book that discusses them, while a novel or work of nonfiction must be summarized; consequently, the discussion of a long work of prose can prove inaccessible to those who have not read it. See, e.g., Johnson, *supra* note 4, at 553–54 & 553 n.20 (reproducing in their entirety two sonnets by Baudelaire in French and English, translated by Johnson).

15. The editors did not limit their selections to poems by American writers; the anthology includes poems by Irish poets Seamus Heaney and Paul Durcan as well as British poets Percy Bysshe Shelley and Glyn Maxwell, among other non-U.S. poets.

16. POETRY OF THE LAW, *supra* note 5, at xiii. But see Robert E. Rains, *Of Cantos and Clerihews*, 13 GREEN BAG 2d 453, 456–57 (2010) (reviewing *Poetry of the Law* and debunking both claims).

17. POETRY OF THE LAW, *supra* note 5, at xv–xx.

18. As Rains points out in his review, the anthology’s notes on poets and glosses on archaic and foreign terms are woefully inadequate: “The further one reads in *Poetry of the Law*, the less likely one is to find a note about poet or poem. It is as if the editors were writing against deadline and simply ran out of time.” Rains, *supra* note 16, at 461. The anthologized excerpts from Chaucer,

range of important poets and possesses the potential to augment and reshape the law-and-literature curriculum. At the same time, the editors' slavish adherence to a male-dominated literary tradition¹⁹ severely undermines their attempt to broaden the existing canon.

The first part of this Review identifies some of the themes that emerge from the poems anthologized, focusing on the poets' use of particular tropes and devices in their portrayal of law and lawyers. The second part focuses on the role of form in the anthology's fixed-verse poems, arguing that traditional poetic forms represent law in the poems. The third part explores the roles played by women within the poems, including the personification of Justice that appears in a number of pieces. It also laments the near-total absence of women poets in the anthology—an absence that mirrors the underrepresentation of women writers in law-and-literature syllabi generally. The Review primarily engages the technique of close reading and incorporates portions of the text discussed, rendering it accessible to those unfamiliar with the poems.

Legal and poetic language differ in the relationship of each to meaning. While poets intentionally "cultivate ambiguity and multiply meaning, legal prose aims for concision and clarity."²⁰ Despite the ostensible aim of legal prose, statutes that give rise to controversial judicial opinions "are often deeply ambiguous texts," as are poems; "[t]he specter of hopeless indeterminacy, of rampant subjectivity, hovers over the key texts of both" literature and law.²¹ One prevalent strand in the law-and-literature movement has focused on the most complicated literary works precisely because their complexity and open-endedness provide fertile ground for multiple interpretations. Given that "law, like literature, is unavoidably linguistic in nature, literary insights into the situated, contingent, often ambiguous character of language [a]re as applicable to legal texts as to literary ones."²² No other genre offers more ambiguous language than does poetry.

One lawyer writing in the late '90s declared: "It is no longer necessary to justify the use of literature to approach law, or vice versa. We live by analogy and metaphor; to the extent that exploration of one medium yields light upon another, it has value."²³ The skills required to read poetry

Spenser, and Dunbar are especially opaque, rendered inaccessible to most readers by the absence of either translation or comprehensive glossary.

19. See, e.g., Gemmette, *supra* note 4, at 686 (summarizing the results of her survey data and noting that a dozen works of fiction—*Billy Budd*, *Measure for Measure*, *The Stranger*, *The Merchant of Venice*, *Antigone*, *The Trial*, *Oresteia*, *A Jury of Her Peers*, *Bleak House*, *To Kill A Mockingbird*, *Bartleby the Scrivener*, and *Noon Wine*—were taught in at least twelve and sometimes as many as thirty of the responding teachers' courses but that only two of the twelve stories were written by women).

20. Skeel, *Lawrence Joseph*, *supra* note 6, at 930.

21. POSNER, *supra* note 6, at 273.

22. Skeel, *Practicing Poetry*, *supra* note 6, at 1757.

23. Steven M. Richman, *William Cullen Bryant and the Poetry of Natural Law*, 30 AKRON L. REV. 661, 662 (1997).

carefully and well are related but not identical to the skills lawyers exercise in interpreting statutes and case law.²⁴ Poetry requires diligent attention to its author's choices, including theme, rhetorical device, figurative language, organization, and intertextual conversations. Those same choices inform all good legal writing, and the reader who is attuned to them brings a broader understanding and a sharper skill set to bear on legal texts.²⁵ Judges, academics, students, and practitioners²⁶ thus can benefit from the study of poetry in general;²⁷ the study of "law-related" poems adds relevance and accessibility, and it can also draw attention to those voices marginalized by the judicial process.

II. Discussion: Themes, Form, and Omissions in *Poetry of the Law*

A. Themes

"[T]hat that kept the mind becalmed all winter?— . . . call that the law."²⁸

A number of the poems in Kader and Stanford's anthology draw attention to the ways in which law uses language, sometimes contrasting legal with poetic language and other times highlighting parallels between the two modes. As one scholar observed,

[P]oetry, like trial practice or legal writing, if it works well, is an art of rhythms, imagery, and the crafting of language, with the intent to have a certain effect upon the reader/listener. Poetry is a twin to the law . . .

24. See Daniel J. Kornstein, *The Law and Literature*, N.Y. ST. B.J., May/June 1994, 34, 36 (1994) ("[L]aw and literature involve some similarity of intellectual tasks . . .").

25. Posner grudgingly acknowledges that

[l]awyers might be able to derive some professional utility from studying [Wallace] Stevens's poetry simply because it is difficult Reading a poem by Stevens requires the reader not only to attend carefully to every word but also to consider the extent to which guides to meaning can appropriately be sought from sources outside the text itself To be a good lawyer one must be a careful and resourceful reader, and immersion in poetry and other difficult imaginative literature is therefore not the worst preparation for the study and practice of law.

POSNER, *supra* note 6, at 194.

26. See Daniel J. Kornstein, *A Practicing Lawyer Looks Back on Law and Literature*, 10 CARDOZO STUD. L. & LITERATURE 117, 118 (1998) (arguing that the law-and-literature movement must appeal to and reach practicing lawyers in order to deliver on its promises).

27. Judith Resnik identifies how "the conventional framings of Law and Literature . . . [may miss] that literature is of use not only in the service of law . . . [R]ather than conceptualizing either discipline as being in the service of the other, one might have considered the joint venturing of the disciplines . . ." Judith Resnik, *On the Margin: Humanities and the Law*, 10 YALE J.L. & HUMAN. 413, 418 (1998) [hereinafter Resnik, *On the Margin*]; see also Judith Resnik, *Changing the Topic*, 8 CARDOZO STUD. L. & LITERATURE 339, 350 (1996) ("I do not see literature as the 'handmaiden' . . . of law but on equal footing."). Sara D. Schotland, who teaches at Georgetown University, advocates offering Law and Literature courses to undergraduates. Schotland, *supra* note 13, at 41.

28. Robert Hass, *The Woods in New Jersey*, in POETRY OF THE LAW, *supra* note 5, at 151.

just as demanding, and ultimately as capable of being used toward a multiplicity of ends.²⁹

In some of the poems in the anthology, the alignment of law and language is explicit. In Reznikoff's³⁰ *Early History of a Writer*, the poet attends law school and finds himself thrilled by the use to which words are put: "the sharp prose, / the forthright speech of the judges," "sentences that seem[] meaty," "words / . . . as a pattern for thinking."³¹ The poet is initially "delight[ed] / . . . to use words for their daylight meaning / and not as prisms / playing with the rainbows of connotation," although he becomes disillusioned by his second year of law school.³² Emily Dickinson applies to law the grammar of literature: "I read my sentence—steadily— / Reviewed it with my eyes, / To see that I made no mistake / In its extremest clause."³³

Eavan Boland, in *The Hanging Judge*, decries the law in grammatical terms that echo Dickinson's:

Come to the country where justice is seen to be done,
 Done daily. Come to the country where
 Sentence is passed by word of mouth and raw
 Boys split like infinitives. Look, here
 We hanged our son, our only son
 And hang him still and still we call it law.³⁴

Seamus Heaney offers yet another play on "sentence" in *The Stone Verdict*, whose speaker declares, "It will be no justice if the sentence is

29. Charles Abourezk, *From a Lawyer's Heart, a Pulse of Poetry*, 41 S.D. L. REV. 624, 624 (1996).

30. Reznikoff devoted several volumes of poetry to stories drawn from case law, initially inspired by those opinions he was tasked with summarizing for the legal encyclopedia *Corpus Juris*. David Skeel, *Point-Blank Verse*, LEGAL AFF., Sept.–Oct. 2005, at 56, 56, available at http://www.legalaffairs.org/issues/September-October-2005/review_skeel_sep0ct05.msp. Skeel's article includes an extensive discussion of Reznikoff's work, including Reznikoff's two-volume found poem, *Testimony*:

With its use of judicial opinions as the raw material of poetry, *Testimony* radically undercuts the traditional assumption that the poet works in a private sphere that is somehow separate from the pressures and pulls of the public domain. Not only is the poem an object, but it is an object taken from the workaday world that poets traditionally have viewed as unsuitable for poetry. . . . In the poetic tradition inaugurated by Reznikoff's *Testimony*, . . . distinctions [between law and poetry] begin to collapse.

Id. at 56, 59.

31. Charles Reznikoff, *Early History of a Writer*, in POETRY OF THE LAW, *supra* note 5, at 93, 94–95.

32. *Id.* at 94–96.

33. Emily Dickinson, *I Read My Sentence Steadily*, in POETRY OF THE LAW, *supra* note 5, at 67, 67.

34. Eavan Boland, *The Hanging Judge*, in POETRY OF THE LAW, *supra* note 5, at 155, 155.

blabbed out.”³⁵ The theme of law as naming force recurs throughout the collection. In John Ciardi’s *E Is for Earwig*, law is cast as “the mother tongue.”³⁶ The poem cautions, “Be careful what names you allow to adhere to you. / Avoid going to law.”³⁷ To William Cowper, the role of “[d]efendant, thus, becomes a name.”³⁸

Other poems in the collection lament the impotence of both law and poetry to redress wrongs in a specific instance. Martín Espada, in *Mi Vida: Wings of Fright*, decries the immigration attorney’s, and by extension the law’s, ineffectualness: the lawyer, “like the fortune teller, / ha[s] a bookshelf of prophecy / but a cabinet empty of cures.”³⁹ The speaker in D.H. Lawrence’s *Auto-da-Fé* cries out rhetorically, “Help! Help! I am calling still in English; / is the language dead and empty of reply!”⁴⁰ In a poem about the trial of John Brown, law is portrayed as broken, incapable of adjudicating the claims against Brown fairly when he committed no moral wrong:

No one can say
That the trial was not fair. The trial was fair,
Painfully fair by every rule of law,
And that it was made not the slightest difference.
The law’s our yardstick, and it measures well
Or well enough when there are yards to measure.
Measure a wave with it, measure a fire,
Cut sorrow up in inches, weigh content.
You can weigh John Brown’s body well enough,
But how and in what balance weigh John Brown?⁴¹

The indictment of “yardstick law” is brutally poignant: the law in this poem lacks both heart and intellect. It is a crude tool to measure the heroic acts of an abolitionist.

In its quiet defiance, *John Brown’s Body* echoes Oscar Wilde’s imprecation of law in *The Ballad of Reading Gaol*. Wilde’s poem begins, “I

35. Seamus Heaney, *The Stone Verdict*, in POETRY OF THE LAW, *supra* note 5, at 147, 147.

36. John Ciardi, *E is for Earwig*, in POETRY OF THE LAW, *supra* note 5, at 120, 120.

37. *Id.*

38. William Cowper, *The Cause Won*, in POETRY OF THE LAW, *supra* note 5, at 40, 40.

39. Martín Espada, *Mi Vida: Wings of Fright*, in POETRY OF THE LAW, *supra* note 5, at 174, 174.

40. D.H. Lawrence, *Auto-da-Fé*, in POETRY OF THE LAW, *supra* note 5, at 89, 89. “Auto da fé” literally means “act of the faith”; it refers to the ceremony for pronouncing judgment by the Inquisition which was followed by the execution of the sentence after trial. MERRIAM-WEBSTER’S COLLEGIATE DICTIONARY 83 (11th ed. 2003).

41. Stephen Vincent Benét, *From John Brown’s Body*, in POETRY OF THE LAW, *supra* note 5, at 97, 99.

know not whether Laws be right, / Or whether laws be wrong,"⁴² but belies that feigned neutrality in the stanzas that follow, declaring that "every prison that men build / Is built with bricks of shame, / And bound with bars lest Christ should see / How men their brothers maim."⁴³

A poem by A.E. Housman, *Oh Who Is That Young Sinner*, follows Wilde's in the anthology—the placement is fortuitous given that most critics agree the poem was written in response to Wilde's trial for sodomy, and that the poem's eponymous "young sinner" was homosexual. Housman satirizes the sodomy laws via a tale of a man imprisoned simply because of "the color of his hair . . . the nameless and abominable color of his hair."⁴⁴

While law employs precise language, it is nonetheless disparaged in many poems as meaningless drone or indecipherable babble. George Crabbe dismisses the "perverted minds" of lawyers as using the force of law "[n]ot to protect mankind, but to annoy."⁴⁵ In Lewis Carroll's satirical *The Barrister's Dream*, the law is "puzzling" and "never . . . clearly expressed," and the lawyer holds forth for hours before anyone in the courtroom can ascertain what the defendant "was supposed to have done."⁴⁶ While the lawyer "bellow[s] on,"⁴⁷ the judge is perpetually "explaining the state of the law / In a soft under-current of sound."⁴⁸ Lawyers "drone[] along,"⁴⁹ "babbling" and "dull,"⁵⁰ or "drone on with the hollow sound / of boats rubbing a dock that they're tied to."⁵¹ To John Donne, the lawyer is a windbag, spewing "[w]ords, words, which would tear / The tender labyrinth of a soft maid's ear, / More, more, than ten Sclavonians' scolding, more / Than when winds in our ruin'd abbeys roar."⁵²

The epizeuxis of "words, words" and "more, more" and the repetition of initial and final "r" sounds in "words," "tear," "tender," "ear," "more," "our," "ruin'd," and "roar" imitate the lawyer's endless, meaningless speeches.⁵³ The reviled lawyer writes as much as he talks, "spar[ing] no length"—men have been made "Fathers of the Church for writing less."⁵⁴

42. Oscar Wilde, *The Ballad of Reading Gaol*, in *POETRY OF THE LAW*, *supra* note 5, at 74, 74.

43. *Id.*

44. A.E. Housman, *Oh Who Is That Young Sinner*, in *POETRY OF THE LAW*, *supra* note 5, at 75, 75.

45. George Crabbe, *From Professions—Law*, in *POETRY OF THE LAW*, *supra* note 5, at 44, 45.

46. Lewis Carroll, *The Barrister's Dream*, in *POETRY OF THE LAW*, *supra* note 5, at 69, 69–70.

47. *Id.* at 71.

48. *Id.* at 69.

49. Benét, *supra* note 41, at 98.

50. William Blackstone, *The Lawyer's Farewell to His Muse*, in *POETRY OF THE LAW*, *supra* note 5, at 36, 38.

51. Ted Kooser, *The Witness*, in *POETRY OF THE LAW*, *supra* note 5, at 144.

52. John Donne, *Satire 2*, in *POETRY OF THE LAW*, *supra* note 5, at 16, 17.

53. *Id.*

54. *Id.* at 18.

Lawyer-philosopher George Anastaplo has argued that “[s]ince a poem is, in the best manifestations, an imitation of nature, it lends itself to careful and productive inquiry irrespective of the intention or the explicit awareness of the poet, just as nature itself does even without being aware of herself.”⁵⁵ A poem that employs law purely as metaphor illustrates one of the dominant themes in the anthology’s portrayal of law. Robert Hass dedicates to Justice Brennan his poem *The Woods in New Jersey*, which begins as follows:

Where there was only grey, and brownish grey,
 And greyish brown against the white
 Of fallen snow at twilight in the winter woods,

Now an uncanny flamelike thing, black
 and sulphur-yellow, as if it were dreamed by Audubon,
 Is turned upside down in a delicate cascade

Of new green leaves, feeding on whatever mites
 Or small white spiders haunt underleaves at stem end.
 A magnolia warbler, to give the thing a name.⁵⁶

The poem crisply offsets nature against “the law.” The magnolia warbler, the embodiment of life and nature, is colorful, surprising, and dreamlike. The bird and the natural world precede names and explanations: we are given first “an uncanny flamelike thing,” characterized by colors and prey, hinted at by the reference to “Audubon”; it is only named later, two full stanzas after it enters the poem.⁵⁷ While the lines describing “life” are free of repetition and rhyme, those about “law” are packed with both. The first stanza rhymes “only” with “snow” and “white” with “twilight,” and uses forms of the word “grey” three times and forms of the word “brown” twice.⁵⁸ The letter “w” dominates with “where,” “was,” “white,” “twilight,” and “winter woods.”⁵⁹ The “w” reappears in the third stanza when law asserts itself over the bird, naming it a “warbler” that feeds on “white spiders” and “whatever mites.”⁶⁰ The sixth stanza too is packed with assonance, the hard “e” dominating four words in a row: “reasons, trees reaching each,” which

55. George Anastaplo, *Law & Literature and the Moderns: Explorations*, 20 N. ILL. U. L. REV. 251, 277 (2000).

56. Hass, *supra* note 28, at 151.

57. *Id.*

58. *Id.*

59. *Id.*

60. *Id.* (emphasis added).

pick up the same vowel sound in “austere” above and “deer” below.⁶¹ The poem continues:

The other name we give this overmuch of appetite
And beauty unconscious of itself is life.
And that that kept the mind becalmed all winter?—

The more austere and abstract rhythm of the trunks,
Vertical music the cold makes visible,
That holds the whole thing up and gives it form,

or strength—call that the law. It’s made,
whatever we like to think, more of interests
than of reasons, trees reaching each their own way

for the light, to make the sort of order that there is.⁶²

Bird and life are characterized as pure appetite, “beauty unconscious of itself,” so the names applied to them are externally imposed.⁶³ Law is the namer, comprising “vertical” girders “[t]hat hold[] the whole thing up and give[] it form,” “mak[ing] the sort of order that there is.”⁶⁴ Natural life preexists law, while law is “made” from the stuff of men, “more of interests / than of reasons.”⁶⁵ Life surprises, “flamelike,” while law “becalm[s]” the mind.⁶⁶ Yet in the poem’s final lines, Hass describes the union of the two:

And what of those deer threading through the woods
In a late snowfall and silent as the snow?

Look: they move among the winter trees, so much
the color of the trees, they hardly seem to move.⁶⁷

61. *Id.*

62. *Id.*

63. *Id.*

64. *Id.*

65. *Id.* Legal Realism supposes that the court system creates inconsistent results because outcomes depend too heavily on the predilections of individual judges. Hass’s description of law as “made, / whatever we like to think, more of interests / than of reasons, trees reaching each their own way / for the light” is of a piece with that philosophy. *Id.*

66. *Id.*

67. *Id.*

The deer seem to disappear into their “grey, and brownish grey, / And greyish brown” backdrop, blending with the trees until they nearly vanish.⁶⁸ They become nameless once again, returning to the anonymous status of “uncanny . . . thing.” Named “deer” in the second to last stanza, the animals are only “they” in the final couplet, paradoxically moving and seeming not to move.⁶⁹ Whether law has tamed nature or come to resemble it, the tension in the poem culminates with the apparent synthesis of opposites.

Throughout *Poetry of the Law*, law commands nature: “[T]he law would have a bull stop at red.”⁷⁰ In fact, animals play a surprisingly large role in the anthology. In just one hundred poems, we see dogs in seven different poems,⁷¹ wolves in two,⁷² pigs in two,⁷³ and oysters in two⁷⁴ as well as appearances by a calf,⁷⁵ ox,⁷⁶ ass,⁷⁷ fish,⁷⁸ bear,⁷⁹ pony,⁸⁰ crocodile,⁸¹ beaver,⁸² mouse,⁸³ mule,⁸⁴ snake,⁸⁵ antelope,⁸⁶ bass,⁸⁷ squirrel,⁸⁸ deer,⁸⁹ cow,⁹⁰

68. *Id.*

69. *Id.*

70. Stephen Dunn, *Outlaw*, in *POETRY OF THE LAW*, *supra* note 5, at 141, 141.

71. Seth Abramson, *If You Ask Your Attorney to Be Concise*, in *POETRY OF THE LAW*, *supra* note 5, at 177, 177; John Donne, *Satire 5*, in *POETRY OF THE LAW*, *supra* note 5, at 20, 22; Ralph Waldo Emerson, *Hamatreya*, in *POETRY OF THE LAW*, *supra* note 5, at 56, 56; Robert Fergusson, *The Rising of the Session*, in *POETRY OF THE LAW*, *supra* note 5, at 41, 43; Ben Jonson, *An Epigram to the Counsellor*, in *POETRY OF THE LAW*, *supra* note 5, at 14, 14; Brad Leithauser, *Law Clerk, 1979*, in *POETRY OF THE LAW*, *supra* note 5, at 168, 170; James Wright, *At the Executed Murderer's Grave*, in *POETRY OF THE LAW*, *supra* note 5, at 128, 129.

72. John Ashbery, *Ignorance of the Law Is No Excuse*, in *POETRY OF THE LAW*, *supra* note 5, at 132, 132; Jonson, *supra* note 71, at 14.

73. Carroll, *supra* note 46, at 69, 70; Fergusson, *supra* note 71, at 43.

74. Alexander Pope, *Verbatim from Boileau*, in *POETRY OF THE LAW*, *supra* note 5, at 35, 35; Mona Van Duyn, *The Poet Reconciles Herself to Politicians*, in *POETRY OF THE LAW*, *supra* note 5, at 123, 123.

75. Donne, *supra* note 52, at 17.

76. *Id.*

77. *Id.* at 18.

78. Robert Herrick, *Upon Case*, in *POETRY OF THE LAW*, *supra* note 5, at 24, 24.

79. Ashbery, *supra* note 72, at 132.

80. Fergusson, *supra* note 71, at 41.

81. Percy Bysshe Shelley, *To the Lord Chancellor*, in *POETRY OF THE LAW*, *supra* note 5, at 53, 55.

82. Carroll, *supra* note 46, at 69.

83. Carl Sandburg, *The Lawyers Know Too Much*, in *POETRY OF THE LAW*, *supra* note 5, at 85, 85.

84. Benét, *supra* note 41, at 97.

85. *Id.* at 98.

86. Miriam Waddington, *In a Corridor at Court*, in *POETRY OF THE LAW*, *supra* note 5, at 121, 121.

87. Robert Lowell, *Law*, in *POETRY OF THE LAW*, *supra* note 5, at 122, 122.

88. Philip Levine, *Possession*, in *POETRY OF THE LAW*, *supra* note 5, at 133, 133.

89. Hass, *supra* note 28, at 151.

90. Paul Durcan, *This Week the Court Is Sleeping in Loughrea*, in *POETRY OF THE LAW*, *supra* note 5, at 154, 154.

mole,⁹¹ trout,⁹² and ferret.⁹³ A remarkable number of birds fly across the pages, including sparrow,⁹⁴ chicken,⁹⁵ hawk,⁹⁶ warbler,⁹⁷ and crow,⁹⁸ and even more insects crawl, with references to slug,⁹⁹ snail,¹⁰⁰ fly,¹⁰¹ moth,¹⁰² bee,¹⁰³ and earwig.¹⁰⁴ Worms appear in four different poems¹⁰⁵ and spiders in two.¹⁰⁶ Sometimes man and beast are conflated, as in Paul Durcan's *This Week the Court Is Sleeping in Loughrea*, where "[u]p in the amphitheatre of the public gallery / An invisible mob are chewing the cud."¹⁰⁷ While judge, lawyers, and defendants snooze in the courtroom, a thin stream "carries water out to the parched fields / Where cleg-ridden cattle wait thirsty in the shadowy leas, / Their domain far away from the sleeping courtroom of human battle. / Is it any surprise that there are children who would rather be cattle?"¹⁰⁸

Mona Van Duyn, in *The Poet Reconciles Herself to Politicians*, also contrasts the living, *id*-driven world with law, using livestock to characterize the former in visceral, animalistic terms: "Wormy riches, pride's weed, / the hot mash of sex, / power's cold chickenfeed, / earned by bloody head-pecks"¹⁰⁹ that "greed gobbled raw."¹¹⁰ Something needs impose order upon the fowl as they lunge at one another's heads in the dirt, and that something is "law, / dealing by force / with the heart's monstrous maw."¹¹¹

In Reznikoff too, the language of law is both dispassionate and violent, with "all the blood—the heartache and the heartening—gone out of the words / and only, as a pattern for thinking, / the cool bones of the judge's reasoning."¹¹²

91. Rita Dove, *Twelve Chairs*, in *POETRY OF THE LAW*, *supra* note 5, at 163, 165.

92. *Id.*

93. Abramson, *supra* note 71, at 177.

94. Fergusson, *supra* note 71, at 42.

95. Van Duyn, *supra* note 74, at 123.

96. Benét, *supra* note 41, at 99.

97. Hass, *supra* note 28, at 151.

98. Dove, *supra* note 91, at 165.

99. Edna St. Vincent Millay, *Justice Denied in Massachusetts*, in *POETRY OF THE LAW*, *supra* note 5, at 91, 91.

100. Benét, *supra* note 41, at 97.

101. *Id.* at 98.

102. *Id.* at 97.

103. Ashbery, *supra* note 72, at 132.

104. Ciardi, *supra* note 36, at 120.

105. Benét, *supra* note 41, at 98; Donne, *supra* note 71, at 20; Shelley, *supra* note 81, at 53; Van Duyn, *supra* note 74, at 123.

106. Ashbery, *supra* note 72, at 132; Hass, *supra* note 28, at 151.

107. Durcan, *supra* note 90, at 154.

108. *Id.*

109. Van Duyn, *supra* note 74, at 123.

110. *Id.*

111. *Id.*

112. Reznikoff, *supra* note 31, at 95.

George Crabbe, writing in the early 1800s, applies to lawyers the central theme of Thomas Jefferson's famous letter about the head and the heart¹¹³: "How they should judge of man; his word and deed / They in their books and not their bosoms read."¹¹⁴ For Crabbe, "[t]he heart resigns all judgment to the head; / Law, law alone, for ever kept in view, / His measures guides, and rules his conscience too."¹¹⁵ Likewise, Edward Taylor asserts that God gave man law "[t]o regulate his thoughts, words, life thereby."¹¹⁶

A number of other poems present law as a force for making and keeping order, as in W.S. Merwin's *Tool*, in which "the law" has only "one truth . . . stirring in its head / order order."¹¹⁷ The double "order" mimics the judge's classic exhortation for "order in the court" while simultaneously communicating the function of law and the mandates it issues. It may also allude to Samuel Taylor Coleridge's comparison of prose, "words in their best order," to poetry, "the best words in their best order."¹¹⁸

William Matthews's *Negligence*, which mimics the structure and tone of a legal argument, culminates in a musical plea: "Ladies and gentlemen of the jury, I ask / you to vote against random pain, to vote / that suffering has cause and thus has blame, / to vote that our lives can be explained, and / to vote compensation for my client."¹¹⁹ The anaphora of "to vote" makes a politician out of the lawyer, and his final attempt to sway his audience calls on the metaphor of law as imposing order on nature: a verdict for the victim is an affirmation "that our lives can be explained."¹²⁰ Thomas Lux expands upon the tool metaphor in a poem about instruments of torture, decrying "[w]hat man has done to woman and man / and the tools he built to do it with"¹²¹ as "pure genius in its pain."¹²²

In Muriel Rukeyser's *The Trial*, the wrongly convicted "walk to a chair, to the straps and rivets / and the switch spitting death and Massachusetts' will."¹²³

113. Letter from Thomas Jefferson to Mrs. Cosway (Oct. 12, 1786), in 2 MEMOIRS, CORRESPONDENCE, AND PRIVATE PAPERS OF THOMAS JEFFERSON 46-55 (Thomas Jefferson Randolph ed., London, Henry Colburn & Richard Bentley 1829); see also POSNER, *supra* note 6, at 3 (observing that novelist E.M. Forster's novel *Howards End* reveals that Forster "associates the legal style of thinking with the failure to connect heart and mind").

114. Crabbe, *supra* note 45, at 44.

115. *Id.*

116. Edward Taylor, *Meditation 38: An Advocate with the Father*, in POETRY OF THE LAW, *supra* note 5, at 25, 25.

117. W.S. Merwin, *Tool*, in POETRY OF THE LAW, *supra* note 5, at 131, 131.

118. WILLIAM S. WALSH, *HANDY-BOOK OF LITERARY CURIOSITIES* 909 (Phila., J.B. Lippincott Co. 1893).

119. William Matthews, *Negligence*, in POETRY OF THE LAW, *supra* note 5, at 152, 153.

120. *Id.*

121. Thomas Lux, *Traveling Exhibit of Torture Instruments*, in POETRY OF THE LAW, *supra* note 5, at 157, 157.

122. *Id.*

123. Muriel Rukeyser, *The Trial*, in POETRY OF THE LAW, *supra* note 5, at 112, 113.

One key difference between law and literature is that while both rely on language, law's language is backed by the power of the police state and the threat of physical violence: "[T]he law has direct, instrumental consequences that distinguish it from poetry and other literature in important respects."¹²⁴ Per the disillusioned speaker in William Carlos Williams's *Impromptu: The Suckers*, the accused must "get the current / shot into"¹²⁵ him despite his innocence, "for the glory of the state / and the perpetuation of abstract justice."¹²⁶ To Ben Jonson, writing over three hundred years before Lux and at least two hundred before Williams, the lawyer comes: "like a chief into the court, / Armed at all pieces, as to keep a fort / Against a multitude, and (with thy style / So brightly brandished) wound'st, defend'st—the while / Thy adversaries fall, as not a word / They had, but were a reed unto thy sword."¹²⁷

Though a law is made up of words, it also brings to bear the physical punishment of imprisonment, torture, and even death. Lord Byron summarizes that distinction: "The lawyer's brief is like the surgeon's knife, / Dissecting the whole inside of a question, / And with it all the process of digestion."¹²⁸

The lines pack humor—"dirty"¹²⁹ lawyers nauseate the poet—but the trope of legal argument as knife is emblematic of the tension between law and poetry throughout the collection. While the metaphorical sword usually appears in judicial decisions to signify an offensive use of evidence or legal argument, law as the "surgeon's knife" possesses the potential to cure as well as wound. As such, Byron's metaphor reflects a more optimistic view of the role of law than do many others in the anthology.

As embodiments of the law, lawyers are portrayed in a number of poems as logical and joyless in their use of language. The protagonist in Lawrence Joseph's *Admissions Against Interest* identifies himself: "[W]hat type of animal asks after facts? / —so I'm a lawyer."¹³⁰ He observes a girl on a bus, singing, laughing, and throwing her hands in the air, her "clear fierce eyes wet / in this rain either with rain or with tears."¹³¹ The rhymes of "clear," "fierce," and "tears" and the monosyllables of the two lines contrast the legalistic five-dollar words that characterize the protagonist throughout the poem: he is "literal," "reconciled," "circumspect," and wholly transfixed

124. Skeel, *Practicing Poetry*, *supra* note 6, at 1756; *see also* Skeel, *Lawrence Joseph*, *supra* note 6, at 930 ("Unlike a novel, judicial opinions can assess damages or send a defendant to jail.").

125. William Carlos Williams, *Impromptu: The Suckers*, in *POETRY OF THE LAW*, *supra* note 5, at 86, 87.

126. *Id.*

127. Jonson, *supra* note 71, at 14–15.

128. George Gordon, Lord Byron, *from* Canto 10, *Don Juan*, in *POETRY OF THE LAW*, *supra* note 5, at 52, 52.

129. *Id.*

130. Lawrence Joseph, *Admissions Against Interest*, in *POETRY OF THE LAW*, *supra* note 5, at 159, 160.

131. *Id.* at 161.

by the free spirit on the bus.¹³² “[I]t makes you want to / shout”,¹³³ she “shouts,” throwing both her hands up in an act of abandon the lawyer envies. He laments: “I could never act / as if I didn’t think. My best cogitations / dwell in air so thick it weighs / on the skin, a solid complex, constrained.”¹³⁴

This subpart has attempted to highlight some of the ways in which law and poetry are portrayed as both kin and counterparts to one another and provide a glimpse of the range, accessibility, and beauty of the poems in the anthology. The themes and interpretations discussed, however, represent just a few of the myriad available to readers of the collection. My close readings are not definitive but merely serve as examples. Law students grappling with questions of justice, public policy, and the role of lawyers will find other avenues of ingress to these poems and draw different connections and parallels between their legal studies and the poems.

B. Form

*“screaming in this box forever”*¹³⁵

If, as Hass writes, law “holds the whole thing up and gives it form,”¹³⁶ the fixed-verse poems in *Poetry of the Law* embody that theory. Form in these poems stands in for the law, erecting a framework and set of rules to which the poet must adhere in accordance with tradition. Literary critic Barbara Johnson dubbed lyric poetry “the more law-abiding or rule-bound of the genres.”¹³⁷ Wordsworth has compared the sonnet form to a prison,¹³⁸ while Keats described it as “chain[ing]” and “fetter[ing]” language.¹³⁹ Edna St. Vincent Millay, in a sonnet that opens, “I will put Chaos into fourteen lines,” imprisons the figure of Chaos within “the strict confines” of her sonnet, “this sweet Order . . . where, in pious rape, / I hold his essence and amorphous shape, / Till he with Order mingles and combines.”¹⁴⁰ Fixed-verse forms thus lend themselves elegantly to the subject, where “the strict confines” of the poems themselves mirror the confines imposed by law.

Poetry of the Law includes a range of fixed-verse poems, from Ben Jonson’s epigram in heroic couplets¹⁴¹ and Robert Fergusson’s use of the

132. *Id.* at 159–61.

133. *Id.* at 161.

134. *Id.*

135. Weldon Kees, *After the Trial*, in *POETRY OF THE LAW*, *supra* note 5, at 117, 117.

136. Hass, *supra* note 28, at 151.

137. Johnson, *supra* note 4, at 550.

138. WILLIAM WORDSWORTH, *Nuns Fret Not at Their Convent’s Narrow Room*, in *SELECTED POETRY AND PROSE OF WORDSWORTH* 169, 169 (Geoffrey H. Hartman ed., 1980).

139. JOHN KEATS, *On the Sonnet*, in *SELECTED POETRY OF KEATS* 264 (Paul de Man ed., 1966).

140. EDNA ST. VINCENT MILLAY, *I Will Put Chaos into Fourteen Lines*, in *COLLECTED POEMS OF EDNA ST. VINCENT MILLAY* 728 (Norma Millay ed., Harper Perennial 1981) (1917).

141. Jonson, *supra* note 71, at 14–15.

standard Habbie¹⁴² to Weldon Kees's sestina,¹⁴³ accompanied by sonnets Shakespearean,¹⁴⁴ Petrarchan,¹⁴⁵ Wordsworthian,¹⁴⁶ and Cyhydedd Fer.¹⁴⁷ John Hollander, asking "How can a punishment fit a crime?," titled his poem *Tailor-Made* and crafted a tailor-made form for it, following two Petrarchan quatrains with one Shakespearean one.¹⁴⁸

William Cowper's *The Cause Won* begins as follows:

Two neighbours furiously dispute,
A field the subject of the suit;
Trivial the spot—yet such the rage
With which the combatants engage,
'Twere hard to tell who covets most
The prize, at whatsoever cost.¹⁴⁹

Published in the eighteenth century, the poem exhibits a number of the themes prevalent in the anthology, including the revelation of law as mere "words," the metaphorical role of law as the process of naming, the often futile nature of litigation, and lawyers as blowhards:

The pleadings swell. Words still suffice;
No single word but has its price;
No term but yields some fair pretence
For novel and increased expence.¹⁵⁰

142. Fergusson, *supra* note 71, at 41–43. The standard Habbie, also called the Burns stanza, was popular among lowland Scots poets like Fergusson and Robert Burns. Robert Crawford, *Robert Fergusson's Robert Burns*, in ROBERT BURNS AND CULTURAL AUTHORITY 1, 4 (Robert Crawford ed., 1997). Each stanza is six lines in length and rhymes AAABAB, with tetrameter (four-foot) A lines and dimeter (two-foot) B lines. Douglas Dunn, "A Very Scottish Kind of Dash": *Burns's Native Metric*, in ROBERT BURNS AND CULTURAL AUTHORITY, *supra*, at 58, 60.

143. Kees, *supra* note 135, at 117–18.

144. William Shakespeare, *Sonnet 35*, in POETRY OF THE LAW, *supra* note 5, at 10, 10; William Shakespeare, *Sonnet 49*, in POETRY OF THE LAW, *supra* note 5, at 11, 11; William Shakespeare, *Sonnet 134*, in POETRY OF THE LAW, *supra* note 5, at 12, 12; Bartholomew Griffin, *Arraigned, Poor Captive*, in POETRY OF THE LAW, *supra* note 5, at 23, 23.

145. Sir John Davies, *Into the Middle Temple of My Heart*, in POETRY OF THE LAW, *supra* note 5, at 13, 13.

146. William Wordsworth, *from Sonnets upon the Punishment of Death*, in POETRY OF THE LAW, *supra* note 5, at 51, 51. Wordsworth's sonnets often employed a rhyme scheme of ABBA, ACCA, DEED, FF.

147. Cowper, *supra* note 38, at 40.

148. John Hollander, *Tailor-Made*, in POETRY OF THE LAW, *supra* note 5, at 134, 134.

149. Cowper, *supra* note 38, at 40.

150. *Id.*

The poem appears to be a Cyhydedd Fer sonnet, a Welsh form not often seen among the works of British poets like Cowper. It consists of seven rhyming couplets in which the rhyme falls on the eighth syllable,¹⁵¹ with a resolution in the final couplet. Unlike most sonnet forms, the Cyhydedd Fer imposes no specific metrical requirements, and the lines are typically self-contained rather than enjambed; Cowper thus may have selected it instead of its English equivalent—the Shakespearean sonnet in iambic pentameter—to underscore the internal din of litigation. Here, each line ends on a masculine rhyme until the final couplet, which is feminine:

Defendant, thus, becomes a name
Which he that bore it may disclaim,
Since both, in one description blended,
Are plaintiffs when the suit is ended.¹⁵²

The neighbors, squabbling over a mere “[t]rivial . . . spot,” inhabit every one of the first ten lines like fighters in a ring, “disput[ing],” “rag[ing],” “engag[ing],” “pleading[],” “swell[ing],” and never “yield[ing].”¹⁵³ In the final couplet, they are united and “blended,” “both . . . plaintiffs when the suit is ended.”¹⁵⁴ “[P]laintiff[]” here refers not only to the instigator of a lawsuit, but to a complainer; it also connotes the root word *plaintive*, meaning melancholy or pathetic. The “furious[] . . . rage” of the first stanza thus gives way to the boredom and plaintiveness of the second.¹⁵⁵

Unsurprisingly, a greater proportion of the twentieth-century poems in the anthology than of the earlier works are written in free verse, but *Poetry of the Law* contains notable exceptions. In his sestina *After the Trial*, Weldon Kees uses form to capture the trapped feeling that plagues the imprisoned speaker and the repetitive thoughts that torture him. It begins:

Hearing the judges’ well-considered sentence,
The prisoner saw long plateaus of guilt,
And thought of all the dismal furnished rooms
The past assembled, the eyes of parents
Staring through walls as though forever

151. The rhyme pattern is AABCCDDEEFFGG.

152. Cowper, *supra* note 38, at 40. Masculine rhymes are those in which the stress occurs on the final syllable of the words, such as Cowper’s “dispute”/“suit”; “rage”/“engage”; “name”/“disclaim.” Feminine rhymes are those in which the stress is placed on the penultimate syllable of the words, as in Cowper’s “blended”/“ended.”

153. *Id.*

154. *Id.*

155. *Id.*

To condemn and wound his innocence.

And if I raise my voice, protest my innocence,
 The judges won't revoke their sentence.
 I could stand screaming in this box forever,
 Leaving them deaf to everything but guilt;
 All the machinery of law despised by parents
 Could not be stopped though fire swept the rooms.¹⁵⁶

The *sestina*, among the “most complicated” of verse forms,¹⁵⁷ is a thirty-nine line poem that adheres to a very strict pattern.¹⁵⁸ It ends each line with one of six words—in Kees’s poem, they are “sentence,” “guilt,” “rooms,” “parents,” “forever,” and “innocence.”¹⁵⁹ The six words continually rotate position in each six-line stanza according to a specific scheme, and all six are incorporated in the final three-line “envoy” in which the poem culminates.¹⁶⁰

Here, the recurrence of the same six words in the poem communicates the repetitiveness of the speaker’s experience: the nagging “guilt” and the feeling of being imprisoned “forever,” reliving the hellish moment of “sentence” *ad infinitum*, are portrayed elegantly by the restrictive poetic form.¹⁶¹ Yet even as the images and emotions repeat, they evolve throughout the poem, changing with each appearance. The judges and parents of the first stanza become enmeshed with each other so that by the third stanza, it is parents who “speak the hideous sentence,” and in the fourth stanza the speaker hears in the voices of the judges “the believing voice of parents.”¹⁶²

Whenever my thoughts move to all those rooms
 I sat alone in, capable of innocence,
 I know now I was not alone, that parents
 Always were there to speak the hideous sentence:
 “You are our son; be good; we know your guilt;
 We stare through walls and see your thoughts forever.”

156. Kees, *supra* note 135, at 117.

157. Alex Preminger & Clive Scott, *Sestina*, in *THE NEW PRINCETON ENCYCLOPEDIA OF POETRY AND POETICS* 1146, 1146 (Alex Preminger & T.V.F. Brogan eds., 1993).

158. The typical rhyme scheme for a *sestina* is ABCDEF, FAEBDC, CFDAEB, ECBFAD, DEACFB, BDFECA, ECA or ACE. *Id.* Kees adheres to the scheme for the first thirty-six lines but uses BDC rather than ECA for the envoy.

159. Kees, *supra* note 135, at 117.

160. *Id.* at 118.

161. *Id.* at 117.

162. *Id.*

Sometimes I wished to go away forever;
 I dreamt of strangers and of stranger rooms
 Where every corner held the light of guilt.
 Why do the judges stare? I saw no innocence
 In them when they pronounced the sentence;
 I heard instead the believing voice of parents.¹⁶³

The first stanza's purported innocence is also seen through a kaleidoscope. The speaker first considers "protest[ing his] innocence" paradoxical because we can read it either as asserting the speaker's innocence or protesting the label "innocence."¹⁶⁴ In the next stanza, Kees uses enjambment to engage the same double meaning: the speaker is "capable of innocence," but the phrase follows "rooms / I sat alone in," so that the resulting juxtaposition creates "in, capable of innocence"—simultaneously both capable and *incapable*.¹⁶⁵ In the fourth stanza, the speaker sees "no innocence" in the judges, and in the sixth, none in his parents, until every figure in the poem has been colored both guilty and innocent.¹⁶⁶

Their eyes burn. How can I deny my guilt
 When I am guilty in the sight of parents?
 I cannot think that even they were innocent.
 At least I shall not have to wait forever
 To be escorted to the silent rooms
 Where darkness promises a final sentence.¹⁶⁷

The rooms too change over the course of the poem, evolving from the "dismal furnished rooms" of the past, to "strange[] rooms," to the rooms "[w]here darkness promises a final sentence."¹⁶⁸ The "silent rooms" of the second-to-last stanza belie the speaker's lament in the final tercet that he is doomed "[n]ever to enter innocent and quiet rooms," but it seems evident that the "innocent and quiet rooms" have come to represent the unattainable quiet mind, no longer tortured by "sentences and eyes of parents".¹⁶⁹

163. *Id.*

164. *Id.* at 117.

165. *Id.*

166. *Id.* at 117–18.

167. *Id.* at 118.

168. *Id.* at 117–18.

169. *Id.* at 118.

We walk forever to the doors of guilt,
 Pursued by our own sentences and eyes of parents,
 Never to enter innocent and quiet rooms.¹⁷⁰

The sestina form also magnifies the roles that vision, sound, and voice play. The judges “stare” and the prisoner “s[ees]” in both the first and fourth stanzas, but it is the parents to whom most of the vision verbs belong: it is “the eyes of parents,” “staring through walls,” “see[ing the speaker’s] thoughts,” and gazing through “watchful eyes,” with “eyes burn[ing].”¹⁷¹ The speaker is “guilty in the sight of parents,” “[p]ursued by . . . [the] eyes of parents,” so that the parents, who arguably stand in for the state, rotate to a different position in each stanza, constantly surveilling the speaker.¹⁷² Only the parents speak out loud in the poem. Their statements are quoted in the third and fifth stanzas, while the speaker is utterly voiceless: he “could stand screaming in this box forever,” but his audience would remain deaf and his sentence, unaltered.¹⁷³

Kees’s sestina lends itself well to analysis under the first strand of the law-and-literature scholarship.¹⁷⁴ To humanists, the poem’s protagonist provides a glimpse into the deeply troubled mind of a young person who stands accused of some wrong. His story might enable us to better empathize with a person—rightfully or wrongfully sentenced—on the receiving end of “[a]ll the machinery of law.”¹⁷⁵ Here, the sestina form with its strict requirements imprisons the poem’s protagonist, but his experience of prison is not static. As the six end words migrate position and change connotations, they illustrate the shifting and subjective nature of truth.

C. Omissions

*“This woman’s bruised heart / is evidence.”*¹⁷⁶

Few poems in the collection reflect female voices: only eight of the hundred poems the editors selected were written by women.¹⁷⁷ A fifty–fifty

170. *Id.*

171. *Id.*

172. *Id.* at 118.

173. *Id.* at 117.

174. See *supra* note 3 and accompanying text.

175. Kees, *supra* note 135, at 117.

176. Matthews, *supra* note 119, at 153.

177. The women whose poetry is included in *Poetry of the Law* are as follows: Emily Dickinson (two poems); Edna St. Vincent Millay; Muriel Rukeyser; Miriam Waddington; Mona Van Duyn; Eavan Boland; and Rita Dove. While the discussion in this Review is limited to poems included in Kader and Stanford’s 2010 anthology, I take this opportunity to recommend a few examples of the many stellar “law-related” poems by women for the professor assembling a law-and-literature reader. MARGARET ATWOOD, *Marrying the Hangman*, in *SELECTED POEMS II*, at 17 (1987); GWENDOLYN BROOKS, *The Boy Died in My Alley*, in *TO DISEMBARK* 49 (1981); LUCILLE

split would be inorganic, given that we have access to far fewer poems by women than by men written in the time of Chaucer or Wordsworth. But the anthology's vast gender imbalance, especially among modern poets, is highly offensive: we are left to assume that the editors considered poems written by women and deemed almost all unworthy of their project. Interestingly, gay poets are better represented in the anthology than are women.¹⁷⁸ The omission of so many important female poets is noteworthy¹⁷⁹ in this anthology not because it is unusual, but precisely because the vast majority of writers appearing in law-and-literature curricula are also male,¹⁸⁰ even though many of its scholars and teachers are women. As Judith Resnik has observed, "[i]n the emerging 'canon' of what falls within late-twentieth-century Law and Literature in law schools, women remain barely visible."¹⁸¹ That invisibility parallels women's invisibility in the law itself.¹⁸² Literature and law "have worked—separately and together, via canonised texts and legal rules, to suppress and make silent much of the world inhabited and understood by women. . . . Until quite recently, women were the objects of the discussion, as property, as victims, as defendants, but not the authors [or] speakers"¹⁸³

CLIFTON, *At the Cemetery, Walnut Grove Plantation, South Carolina, 1989*, in *QUILTING: POEMS, 1987–1990*, at 11 (1991); DENISE DUHAMEL, *What Happened This Week*, in *QUEEN FOR A DAY 11* (2001); JORIE GRAHAM, *Of Forced Sights and Trusty Ferefulness*, in *THE DREAM OF THE UNIFIED FIELD 95* (1995); JUNE JORDAN, *Letter to the Local Police*, in *DIRECTED BY DESIRE 267* (Jan Heller Levi & Sara Miles eds., 2005); AUDRE LORDE, *Power*, in *THE COLLECTED POEMS OF AUDRE LORDE 215* (1997); THYLIAS MOSS, *Interpretation of a Poem by Frost*, in *RAINBOW REMNANTS IN ROCK BOTTOM GHETTO SKY 44* (Charles Simic ed., 1991); ELISE PASCHEN, *Voir Dire*, in *INFIDELITIES 31* (1996); MARGE PIERCY, *The Grey Flannel Sexual Harassment Suit*, in *WHAT ARE BIG GIRLS MADE OF? 41* (1997); MINNIE BRUCE PRATT, *Crime Against Nature*, in *CRIME AGAINST NATURE 111* (1990); MURIEL RUKEYSER, *The Bill*, in *THE COLLECTED POEMS OF MURIEL RUKEYSER 100* (1978); JEAN VALENTINE, "Actuarial File," in *DOOR IN THE MOUNTAIN 157* (2004); and Averill Curdy, *To the Voice of the Retired Warden of Huntsville Prison (Texas Death Chamber)*, *POETRY*, June 2009, at 194.

178. The anthology includes poems from at least ten poets believed to be gay or bisexual: John Ashbery, W.H. Auden, Lord Byron, Thom Gunn, A.E. Housman, Langston Hughes, D.H. Lawrence, Edna St. Vincent Millay, Walt Whitman, and Oscar Wilde. At least two of the poems, Wilde's *The Ballad of Reading Gaol* and Housman's *Oh Who Is That Young Sinner*, concern the treatment of homosexual men under the law. See *supra* text accompanying notes 42–44.

179. In a short review, poet and blogger Ron Slate calls the collection "markedly testosteroneish." Ron Slate, *On Four New Poetry Anthologies*, *ON THE SEAWALL* (June 16, 2010), http://www.ronslate.com/four_new_poetry_anthologies.

180. See Gemmette, *supra* note 4, at 671 n.46 (finding that among all law schools offering courses in law and literature, twenty-two works of nonfiction were assigned three times or more; of these top twenty-two works, only one was written by a female critic).

181. Resnik, *On the Margin*, *supra* note 27, at 418–19.

182. See Robin L. West, *The Literary Lawyer*, 27 *PAC. L.J.* 1187, 1197 (1996) ("Outsider's voices have historically been censored from the language of literature and high culture at least as relentlessly as they have been banned from the language and courts of law.").

183. Judith Resnik, *Changing the Topic*, 7 *AUSTRALIAN FEMINIST L.J.* 95, 108 (1996).

Few women appear in *Poetry of the Law*'s hundred poems, and those that do typically fill archetypal roles as virgin, mother, or whore.¹⁸⁴ They are seldom cast as lawyers or judges, appearing instead to bear witness, play victim, or receive punishment.¹⁸⁵ In *The Witness* by Ted Kooser, the speaker addresses in second person a woman awaiting her turn to testify in a divorce trial:

and you wait at the back of the courtroom
 as still as a flag on its stand, your best dress
 falling in smooth, even folds that begin now
 to gather the dust of white bouquets
 which like a veil of lace is lifting
 away from the kiss of the sunlit windows.

In your lap, where you left them, your hands
 Lie fallen apart like the rinds of a fruit.
 Whatever they cupped has been eaten away.¹⁸⁶

Bridal imagery dominates the first stanza, from the white bouquets to the veil of lace to the sunlit kiss, but the flowers have turned to dust and so perhaps has the bride, reminiscent of Dickens's Miss Havisham waiting in wedding garb for eternity.¹⁸⁷ Still and obedient as an American flag, the woman awaits her turn to take the (or *a*) stand.¹⁸⁸ Kooser deliberately leaves open the question of whether the subject of the poem is testifying at her own divorce or that of another, but the effect is to distance her further from the action: she is merely an anonymous witness at trial, waiting in the courthouse while around her "voices drone on" without meaning or content.¹⁸⁹ The poem begins, "The divorce judge has asked for a witness" and ends with the single-line stanza, "You know what to say when they call you."¹⁹⁰ The use

184. See, e.g., Donne, *supra* note 52, at 17 (referring to a woman as the "soft maid" and including the line "imbrothel'd strumpets prostitute"); Seamus Heaney, *Punishment*, in *POETRY OF THE LAW*, *supra* note 5, at 145, 145 (identifying a woman as a "[l]ittle adulteress").

185. See, e.g., W.H. Davies, *The Inquest*, in *POETRY OF THE LAW*, *supra* note 5, at 82 (depicting a situation in which a purportedly objective speaker visits a mother and the corpse of her child). "It was a love-child, she explained. / And smiled[.]" *Id.* The dead baby girl's refrain, delivered with "laugh[ter]" and "glee": "What caused my death you'll never know— / Perhaps my mother murdered me." *Id.*

186. Kooser, *supra* note 51, at 144.

187. *Id.*; CHARLES DICKENS, *GREAT EXPECTATIONS* 87–90 (Roger D. Sell ed., MacMillan 1984) (1861).

188. Kooser, *supra* note 51, at 144.

189. *Id.*

190. *Id.*

of “you know what to say” rather than “you know what you will say” or “you know what happened,” coupled with the passivity of being “called upon” and “asked for,” implies that the words the witness waits to speak are not her own, but a predetermined speech authored by a lawyer.

Elsewhere in the anthology, hands serve as synecdoche for power and autonomy; hands commit crimes and dole out justice. For example, in Dylan Thomas’s *The Hand that Signed the Paper*, a hand is the subject of every stanza and nearly every sentence: the hand “holds dominion over / Man by a scribbled name.”¹⁹¹

But for the woman in *The Witness*, hands that once held nourishment lie immobile where they were left; “fallen apart” and “eaten away” characterize the woman¹⁹² as much as the position in which her hands lie in her lap. In that, they resemble the hands of the woman in William Matthews’s *Negligence*. Matthews’s poem reads like a closing argument, describing the plight of a mother who “opens a parcel with no / return address,” and finds herself wrist-deep “in her son’s [a]shes—not, by the way, like silt or dust, / but nubble and grit, boneshards and half-burnt / burls of cartilage, cinders and nuggets.”¹⁹³ Excoriating the crematorium that failed to adequately label the ashes, the speaker plays to the jurors’ fears and sympathies. He implores them to put themselves in the victim’s position, “to glove her hands with yours / and sieve the rubble of your beloved / only son.”¹⁹⁴

The assonance of “glove,” “beloved,” “rubble,” and “son” and repetition of “yours . . . your” juxtaposes juror with victim, making the woman of Matthews’s poem as passive a receptacle as that of Kooser’s.¹⁹⁵ In guiding the jurors to a measure of damages, the lawyer in *Negligence* asks how the jury might “run this cruel film backwards, / . . . lift this woman’s hands from the cinders / of her son and wind them back to her slack / lap.”¹⁹⁶

Matthews sets off the hard tongue-twisting consonance of “nubble” with “rubble,” “burnt” with “burls,” and “nuggets” with “grit” with the easy assonance of “back,” “slack,” “lap,” and “balm” and the consonance of “cinders,” “son,” and “slack,” persuading the jurors to agree on a generous damages award by demonstrating that it is within their power to return the woman to her previous calm state.¹⁹⁷ He characterizes her plight three times as simply “grief.”¹⁹⁸ The victim’s body parts, her hands and lap, thus

191. Dylan Thomas, *The Hand that Signed the Paper*, in *POETRY OF THE LAW*, *supra* note 5, at 115, 115.

192. Kooser, *supra* note 51, at 144; *see also* Wordsworth, *supra* note 146, at 51 (“Ah, think how one compelled for life to abide / Locked in a dungeon needs must eat the heart / Out of his own humanity[.]”).

193. Matthews, *supra* note 119, at 152.

194. *Id.*

195. *Id.*

196. *Id.*

197. *Id.*

198. *Id.*

become evidence in the case against her tortfeasors, just as her “bruised heart / is evidence.”¹⁹⁹

While female victims abound, the most prominent female presence in *Poetry of the Law* is the personified figure of Justice.²⁰⁰ She is typically blind, “a beautiful woman with bandaged eyes / Standing on the steps of a marble temple,” a sword in one hand and a scale in the other.²⁰¹ In Muriel Rukeyser’s depiction of the trial of the Scottsboro boys, while “[a] blinded statue attends before the courthouse, / bronze and black men lie on the grass, waiting.”²⁰² The descriptor “bronze” links the two lines together, seeming to characterize both the statue (or perhaps, statute) and the men, and in so doing, aligning them. Yet, “all the people’s anger finds its vortex here / as the mythic lips of justice open, and speak.”²⁰³ The poem offers, as the anthology’s editors claim, a “sweeping evocation of the history of legal injustice,”²⁰⁴ tying the defendants in the titular trial with John Brown and Sacco and Vanzetti, parading forth “all our celebrated shambles.”²⁰⁵ Justice here is blinded, not blind, and the poem’s audience is blinded too: “[T]he air is populous beyond our vision.”²⁰⁶ The players in the poem are described in terms of their vision or lack thereof: the lynched men have “eyes showing a wild iris,” the jury members’ “eyes like hardware,” and the judge’s “eye-sockets” are “dark and immutably secret.”²⁰⁷

Sir William Blackstone’s representation of Justice, on the other hand, is not blind but hidden from view: she is the queenly guardian of law, “from vulgar sight retired.”²⁰⁸ The speaker longs to “pierce the secret shade / Where dwells the venerable maid!”²⁰⁹ Alexander Pope’s Dame Justice appears with “scale in hand,” but she is neither blind nor silent.²¹⁰ Instead, encountering two “hungry . . . trav’lers” fighting over an oyster, she performs a Solomon-like split, taking the spoils for herself.²¹¹ After each side has “plead[ed] the laws” and argued his case for the oyster,

199. *Id.* at 153.

200. Justice is usually, but not always, gendered female in the poems in Kader & Stanford’s anthology. See *infra* notes 201–14 and accompanying text. For a fascinating discussion of the figure of Justice throughout history, see JUDITH RESNIK & DENNIS CURTIS, REPRESENTING JUSTICE: INVENTION, CONTROVERSY, AND RIGHTS IN CITY-STATES AND DEMOCRATIC COURTROOMS (2011).

201. Edgar Lee Masters, *Carl Hamblin*, in *POETRY OF THE LAW*, *supra* note 5, at 81, 81.

202. Rukeyser, *supra* note 123, at 112.

203. *Id.*

204. *POETRY OF THE LAW*, *supra* note 5, at xviii.

205. Rukeyser, *supra* note 123, at 113.

206. *Id.* at 112.

207. *Id.* at 113.

208. Blackstone, *supra* note 50, at 37–38.

209. *Id.* at 38.

210. Pope, *supra* note 74, at 35.

211. *Id.*

Dame Justice, weighing long the doubtful right,
 Takes, opens, swallows it, before their sight.
 The cause of strife remov'd so rarely well,
 "There take" (says Justice) "take ye each a shell
 We thrive at Westminster on fools like you:
 'Twas a fat oyster—live in peace—Adieu."²¹²

Oysters were less a delicacy than a standard meal in the public houses and taverns of the eighteenth century. Justice here likens the squabblers to the sustenance, linking fools with fat oysters and thriving on both.²¹³ Pope's Dame Justice thus cautions about the futility of legal squabbles, but she is more teasing trickster than solemn embodiment of an ideal. She is also the only true speaker in the poem: where the travelers "dispute," "explain," "clamour" and "plead,"²¹⁴ only Dame Justice's words appear between quotation marks, making her one of the few female voices in the anthology.

Though this Review focuses on poetry and thus falls within the rubric of law *in* literature rather than law *as* literature, some of the questions raised in this subpart could be applied as fruitfully to legal texts such as court opinions, treatises, textbooks, and statutes. A professor might, for example, ask his or her class to consider what roles women play in those texts and who else seems to be absent from them or subordinated within them. Students should be encouraged to consider what or who has been omitted from a law or rule as written and how the omissions affect our reading, applying, and extracting precedents from it. On a more abstract level, students might consider whether the concept of Justice is portrayed in a given text as blind, blinded, or all-seeing. Those types of questions have the potential to open the law up for some types of students and help them tease out problems and paradoxes within the law in order to better understand it.

III. Conclusion

In the words of Richard Weisberg, a key figure in the law-and-literature movement, the "poetic method provokes us, as customary learning does not, to highlight the linguistic, sensory aspects of every part of our craft."²¹⁵ *Poetry of the Law* proposes that law-related poetry can benefit lifelong students both of law and of literature. Despite a number of overarching criticisms of the law-and-literature movement, the steady climb in scholarship that either practices law and literature or critiques it reflects growing interest in the ways in which the two fields of study interconnect. A number of

212. *Id.*

213. *Id.*

214. *Id.*

215. Kornstein, *supra* note 24, at 36.

prominent voices in the academy have challenged the canon of texts, both primary and secondary, that compose most Law and Literature syllabi, typically because those syllabi so often omit works from many marginalized groups. Of those critics, few have objected on the ground that the traditional Law and Literature course excludes poetry. In applying theoretical frameworks from literary criticism to legal texts, the standard Law and Literature course syllabus also declines to borrow from critical approaches to the study of poetry. Those exclusions are to the detriment of students and scholars alike. A careful reading of this new anthology reveals the value to lawyers, law students, and theorists of poetry and of the particular methods with which literary critics effectively engage with poems. Elizabeth Villiers Gemmette gently chides the would-be anthologist, “the Law and Literature Canon never ‘is’[;] rather[,] it is in a perpetual state of ‘becoming.’”²¹⁶ As the canon continues to evolve and “become,” law-related poems like those collected in *Poetry of the Law* deserve a place of honor within it.

216. See Gemmette, *supra* note 4, at 690.

Notes

Monetary Damages and the (b)(2) Class Action: A Closer Look at *Wal-Mart v. Dukes**

Introduction

Last term the Supreme Court decided the highly controversial case of *Wal-Mart v. Dukes*.¹ The case represented the largest employment discrimination class action in recent U.S. history² and involved a class of female employees suing Wal-Mart for alleged gender discrimination in its hiring and promotion practices.³ The Supreme Court faced two issues related to class certification: first, whether the plaintiff class fulfilled the threshold requirements of Rule 23(a),⁴ and second, the circumstances, if any, under which a plaintiff class could recover monetary damages in a Rule 23(b)(2) class action.⁵ This Note focuses on the latter issue.

The advisory committee's note to Rule 23(b)(2) indicates that the subdivision is not meant to apply to cases where the final relief sought "relates exclusively or predominantly to money damages."⁶ Much of the existing legal scholarship on the issue of monetary damages in (b)(2) class actions tends to focus on competing interpretations of this "predominance" language in the advisory committee's note.⁷ This Note argues that such

* I would like to thank Professor Robert Bone for his invaluable advice and guidance throughout the writing and editing process. Learning from and working with him was and continues to be a pleasure and incredible honor. I would also like to thank my family for their continued support of me throughout all of my academic endeavors. Finally, I am grateful to my colleagues on the *Texas Law Review*—particularly, the Volume 90 Notes Editors and Daniel Clemons—for their excellent editing efforts on this Note.

1. *Wal-Mart Stores, Inc. v. Dukes*, 131 S. Ct. 2541 (2011).

2. Linda S. Mullenix, *Attention Female Workers: Will Wal-Mart Roll Back the Largest Employment Discrimination Class Action Ever?*, 38 PREVIEW U.S. SUP. CT. CAS. 249, 249 (2011).

3. *Dukes*, 131 S. Ct. at 2547.

4. *Id.* at 2550.

5. *Id.* at 2557.

6. FED. R. CIV. P. 23 advisory committee's note (1966 Amendment).

7. See, e.g., Jeffrey H. Dasteel & Ronda McKaig, *What's Money Got to Do with It?: How Subjective, Ad Hoc Standards for Permitting Money Damages in Rule 23(b)(2) Injunctive Relief Classes Undermine Rule 23's Analytical Framework*, 80 TUL. L. REV. 1881, 1883 (2006) ("This Article questions the use of the subjective, ad hoc predominance standard to permit nonincidental damages to be included in a Rule 23(b)(2) class and concludes that the subjective, ad hoc standard destroys the analytical framework of Rule 23."); Suzette M. Malveaux, *Fighting to Keep Employment Discrimination Class Actions Alive: How Allison v. Citgo's Predominance Requirement Threatens to Undermine Title VII Enforcement*, 26 BERKELEY J. EMP. & LAB. L. 405, 408 (2005) ("[I]t is imperative that the predominance approach—taken by the majority of circuits that have ruled on [the issue of certifying a 23(b)(2) class seeking monetary damages]—be abandoned in favor of the more equitable ad hoc balancing approach . . ."); Linda S. Mullenix, *Nine Lives: The Punitive Damages Class*, 58 U. KAN. L. REV. 845, 860–61 (2010) (discussing the predominance approach as it relates to classes seeking punitive damages).

scholarship is misconceived because it overlooks the normative policies that underlie Rule 23(b)(2). Accordingly, this Note analyzes the damages question in (b)(2) class actions by first identifying the policies that inform the Rule and then by considering whether particular damage remedies are appropriate provided that they comport with the relative weight given to each of these policies.

Part I begins by examining the procedural requirements of Rule 23 and then briefly reviews the Supreme Court's recent decision as well as existing case law and scholarship on the issue of monetary damages in (b)(2) class actions. Part II identifies the competing policies at stake in (b)(2) class actions—the right to individual participation and the need for remedial efficacy—and contrasts them with the competing policies that underlie (b)(3) class actions. Part III considers whether particular damage remedies are appropriate in (b)(2) class actions in light of these competing policies. Part IV concludes.

I. Rule 23 and the Supreme Court's Decision

Rule 23 of the Federal Rules of Civil Procedure governs the adjudication of class action lawsuits.⁸ For a class to be certified under Rule 23, it must first satisfy the threshold requirements of Rule 23(a): numerosity, commonality, typicality, and adequacy of representation.⁹ In addition, it must fit within one of the pigeonholes in Rule 23(b).¹⁰

The first threshold requirement of Rule 23(a) is numerosity, which requires that the class be so numerous as to make joinder of all of its members impracticable.¹¹ The second requirement is commonality, which requires the existence of questions of law and fact common to the entire class.¹² The third requirement is that the claims or defenses of the representative parties be "typical" of the claims or defenses of the class.¹³ Finally, the Rule addresses adequacy of representation and requires the representative parties to "fairly and adequately protect the interests of the class."¹⁴

In addition to satisfying the threshold requirements of Rule 23(a), the proposed plaintiff class must also meet the requirements for one of the Rule 23(b) provisions. Courts have historically certified employment discrimination class actions under Rule 23(b)(2) or Rule 23(b)(3).¹⁵ Class certification under Rule 23(b)(2) has typically been reserved for cases where declaratory

8. FED. R. CIV. P. 23.

9. FED. R. CIV. P. 23(a).

10. See *infra* notes 15–22 and accompanying text.

11. FED. R. CIV. P. 23(a)(1).

12. FED. R. CIV. P. 23(a)(2).

13. FED. R. CIV. P. 23(a)(3).

14. FED. R. CIV. P. 23(a)(4).

15. Mullenix, *supra* note 2, at 251.

or injunctive relief is the primary remedy being sought.¹⁶ That being said, various courts have authorized monetary damages in (b)(2) class actions where monetary relief is not the exclusive or predominant remedy being sought.¹⁷ The (b)(2) class action is known as the mandatory class action because it binds class members to the final judgment and does not provide them with notice or opt-out rights.¹⁸ The (b)(3) class, also known as the damage class action,¹⁹ can be certified provided that a court finds that (1) common questions of law and fact predominate (predominance), and (2) the class action is superior to other methods for adjudicating the controversy (superiority).²⁰ Moreover, unlike Rule 23(b)(2), Rule 23(b)(3) provides class members with notice and opt-out rights.²¹ Since Rule 23(b)(3) is more restrictive than Rule 23(b)(2), plaintiffs often choose to seek certification under the latter provision.²²

In reaching its decision in *Dukes*, the Supreme Court reversed the Ninth Circuit's ruling in favor of the plaintiffs, by holding that (1) the plaintiffs could not fulfill Rule 23's commonality requirement, and (2) individualized monetary claims, like backpay, could not be certified under Rule 23(b)(2).²³ Specifically, Justice Scalia, writing for the majority, noted with respect to the issue of monetary damages in (b)(2) class actions (which is the subject of this Note) that the Rule "does not authorize class certification when each class member would be entitled to an individualized award of monetary damages."²⁴ Although the Court declined to engage the broader question of whether Rule 23(b)(2) can ever authorize the class certification of monetary claims at all, it noted that "Wal-Mart is entitled to individualized determinations of each employee's eligibility for backpay"; therefore, backpay was not a permissible remedy under Rule 23(b)(2).²⁵ The broader question of the extent to which any monetary claims can be authorized under this provision of the Rule will be examined here in further detail.

16. FED. R. CIV. P. 23 advisory committee's note (1966 Amendment).

17. See, e.g., *Molski v. Gleich*, 318 F.3d 937, 947 (9th Cir. 2003) ("[I]n order to permit certification under [Rule 23(b)(2)], the claim for monetary damages must be secondary to the primary claim for injunctive or declaratory relief." (citing *Probe v. State Teachers' Ret. Sys.*, 780 F.2d 776, 780 (9th Cir. 1986))); *Robinson v. Metro-North Commuter R.R. Co.*, 267 F.3d 147, 164 (2d Cir. 2001) (rejecting a bright-line rule that would bar all claims for monetary damages under 23(b)(2) and instead adopting an ad hoc approach).

18. See *Mullenix*, *supra* note 2, at 251 ("The Rule 23(b)(2) class is for declaratory or injunctive relief, and it is mandatory and does not permit class members to opt-out.").

19. *Id.*

20. FED. R. CIV. P. 23(b)(3).

21. FED. R. CIV. P. 23(c)(2)(B).

22. *Mullenix*, *supra* note 2, at 251.

23. *Wal-Mart Stores, Inc. v. Dukes*, 131 S. Ct. 2541, 2556–57 (2011).

24. *Id.* at 2557.

25. *Id.* at 2560.

A. Existing Case Law

Both Wal-Mart and the plaintiffs relied on the advisory committee note's language in advancing their respective positions. The note states in relevant part, "The subdivision [(b)(2)] does not extend to cases in which the appropriate final relief relates exclusively or predominantly to money damages."²⁶ Courts have offered starkly differing interpretations of this language, specifically as to the meaning of the word *predominantly*. The two most prominent cases reflecting this split are *Allison v. Citgo Petroleum Corp.*²⁷ and *Robinson v. Metro-North Commuter Railroad*.²⁸

In *Allison*, the Fifth Circuit adopted a restrictive approach to interpreting the language of the advisory committee's note. In that case, a number of African-American employees and prospective employees of the Citgo Petroleum Corporation brought a class action lawsuit alleging race-based employment discrimination in a number of areas, including in hiring and in promotion decisions.²⁹ Plaintiffs sought class certification, requesting injunctive relief as well as compensatory and punitive damages.³⁰ The Fifth Circuit relied on the language of the advisory committee's note in reaching its decision.³¹ The court interpreted the note's language to mean that "monetary relief predominates in (b)(2) class actions unless it is *incidental* to requested injunctive or declaratory relief."³² The court then defined *incidental damages* as "damages that flow directly from liability to the class as a whole."³³ Thus, while the Fifth Circuit in *Allison* authorized monetary damages in (b)(2) class actions, it limited the circumstances under which this could happen to those involving damages "in the nature of a group remedy" and excluded those involving "complex individualized determinations."³⁴ Since *Allison*, the Third, Seventh, and Eleventh Circuits have adopted the Fifth Circuit's incidental-damages approach.³⁵

The Second Circuit in *Robinson* adopted a much less restrictive interpretation. *Robinson* involved a class action brought by a number of present and former African-American employees of the Metro-North Commuter Railroad who alleged employment discrimination with respect to

26. FED. R. CIV. P. 23 advisory committee's note (1966 Amendment).

27. 151 F.3d 402 (5th Cir. 1998).

28. 267 F.3d 147 (2d Cir. 2001).

29. *Allison*, 151 F.3d at 407.

30. *Id.*

31. *Id.* at 411.

32. *Id.* at 415 (emphasis added).

33. *Id.* (citing FED. R. CIV. P. 23(b)(2)).

34. *Id.*

35. *Barabin v. Aramark Corp.*, No. 02-8057, 2003 WL 355417, at *1 (3d Cir. Jan. 24, 2003) (citing *James v. City of Dallas*, 254 F.3d 551, 571 (5th Cir. 2001)); *Jefferson v. Ingersoll Int'l, Inc.*, 195 F.3d 894, 898 (7th Cir. 1999) (citing *Allison*, 151 F.3d at 411-16); *Murray v. Auslander*, 244 F.3d 807, 812 (11th Cir. 2001) (citing *Allison*, 151 F.3d at 411).

promotion and discipline in violation of Title VII of the Civil Rights Act.³⁶ The plaintiffs sought injunctive relief for all members of the class along with compensatory damages for class members alleging individual acts of discrimination.³⁷

In reaching its decision, the Second Circuit declined to adopt the incidental-damages test set forth by the Fifth Circuit in *Allison*.³⁸ Instead, it adopted a more pragmatic approach, which required district courts to “consider[] the evidence presented at a class certification hearing and the arguments of counsel,” and then assess whether (b)(2) certification is appropriate in light of “the relative importance of the remedies sought, given all of the facts and circumstances of the case.”³⁹ Specifically, the Second Circuit indicated that district courts should permit (b)(2) certification if: “(1) ‘the positive weight or value [to the plaintiffs] of the injunctive or declaratory relief sought is predominant even though compensatory or punitive damages are also claimed,’ and (2) class treatment would be efficient and manageable, thereby achieving an appreciable measure of judicial economy.”⁴⁰

The Ninth Circuit, in *Molski v. Gleich*,⁴¹ also refused to adopt the Fifth Circuit’s incidental-damages test. In relevant part, the Ninth Circuit noted, as the Second Circuit did in *Robinson*, that the “adoption of a bright-line rule distinguishing between incidental and nonincidental damages for the purposes of determining predominance would nullify the discretion vested in the district courts through Rule 23.”⁴² Instead of adopting a bright-line rule, then, the Ninth Circuit in *Molski* looked to “the specific facts and circumstances of each case” and then assessed whether certification was appropriate in light of those circumstances.⁴³

B. Existing Legal Scholarship

I begin by reviewing the pre-*Allison* literature on monetary damages in mandatory class actions. David Rosenberg, in his article *Class Actions for Mass Torts: Doing Individual Justice by Collective Means*, argues that “bureaucratic justice”—a mode of decision making that focuses on the aggregation of interests of affected individuals in pursuit of collective benefits⁴⁴—provides better opportunities for achieving individual justice than

36. *Robinson v. Metro-North Commuter R.R. Co.*, 267 F.3d 147, 155 (2d Cir. 2001).

37. *Id.*

38. *Id.* at 164.

39. *Id.* (alteration in original) (quoting *Hoffman v. Honda of Am. Mfg., Inc.*, 191 F.R.D. 530, 536 (S.D. Ohio 1999)).

40. *Id.* (alteration in original) (quoting *Allison*, 151 F.3d at 430 (Dennis, J., dissenting)).

41. 318 F.3d 937 (9th Cir. 2003).

42. *Id.* at 950.

43. *Id.*

44. See David Rosenberg, *Class Actions for Mass Torts: Doing Individual Justice by Collective Means*, 62 IND. L.J. 561, 562 (1987) (“[B]ureaucratic justice . . . legitimates the aggregation and averaging of circumstances and interests of affected individuals in pursuit of the collective benefits

private, disaggregative processes.⁴⁵ Rosenberg also suggests that rights-based objections to the bureaucratic-justice model lack merit because they problematically equate individual trial outcomes with individual justice.⁴⁶ Building on Rosenberg's thesis, David Shapiro argues that the class action should be viewed as an "entity" for determining the nature of the lawsuit and its component parts, instead of as an "aggregation" of individuals.⁴⁷ He concludes that "the notion of the class as an entity should prevail over more individually oriented notions of aggregate litigation," even though "substantial institutional problems remain when it comes to implementation."⁴⁸ Finally, Robert Bone, in his review of Steve Yeazell's book *From Medieval Group Litigation to the Modern Class Action*, considers why notice and opt-out rights are provided in (b)(3) class actions but not (b)(2) suits.⁴⁹ Bone proposes that we can make sense of Rule 23's notice and opt-out requirements if we "assume[] that the Advisory Committee approached the res judicata problem in a way that shared much in common with the personal-impersonal dichotomy that dominated late nineteenth and early twentieth century representative suit law."⁵⁰ The Committee may have been sensitive to the "homogeneity" of the class in deciding whether to provide notice and opt-out rights because the case for notice was stronger when the "solidarity of the class" was called into question.⁵¹ Bone suggests that when Committee members contemplated homogeneity, they were most certainly thinking about "whether the adjudication focused on the impersonal class as an aggregate or on class members as individuals."⁵² He argues that the language of Rule 23(b)(2) assumes that if the party opposing the class "deals with the class as an impersonal status rather than with class members as individuals," then the remedy must target that impersonal class and not individual class members.⁵³ By contrast, he indicates that the remedial focus of the (b)(3) class action is on adjudicating the individual entitlements of class members.⁵⁴ Since the (b)(3) judgment has res judicata effect on all

from process efficiency, outcome consistency, and the maximum production of substantive goods.").

45. *Id.* at 567.

46. *Id.*

47. David L. Shapiro, *Class Actions: The Class as Party and Client*, 73 NOTRE DAME L. REV. 913, 917 (1998).

48. *Id.* at 917-18.

49. Robert G. Bone, *Personal and Impersonal Litigative Forms: Reconceiving the History of Adjudicative Representation*, 70 B.U. L. REV. 213, 294-98 (1990) (reviewing STEPHEN C. YEAZELL, *FROM MEDIEVAL GROUP LITIGATION TO THE MODERN CLASS ACTION* (1987)).

50. *Id.* at 296.

51. *Id.* at 296-97 (internal quotation marks omitted).

52. *Id.* at 297.

53. *Id.* at 298.

54. *Id.*

class members, “the opt-out right limits the res judicata impact by giving absentees a choice whether or not to be bound.”⁵⁵

The legal scholarship on this issue following *Allison* and *Robinson* has mostly focused on critically evaluating the two different approaches taken by the circuit courts. For instance, Suzette Malveaux has argued that the restrictive formulation laid out by the Fifth Circuit in *Allison* “threatens to undermine the enforcement of civil rights.”⁵⁶ Specifically, she argues that the incidental-damages test makes it much more difficult for plaintiffs seeking such damages to get a class certified under Rule 23(b)(2) and that the heightened standard for class certification under Rule 23(b)(2) forces plaintiffs to seek certification under Rule 23(b)(3), which imposes greater costs and burdens on them.⁵⁷ Conversely, Jeffrey Dasteel and Ronda McKaig applaud the incidental-damages standard from *Allison* because it ensures that unmanageable individualized damages issues do not become part of Rule 23(b)(2) classes.⁵⁸ They argue that the more pragmatic approach laid out by the Second Circuit in *Robinson* undermines the analytical framework of Rule 23.⁵⁹

The more recent literature focusing on *Dukes* also discusses the viability of obtaining monetary damages in (b)(2) class actions. Linda Mullenix briefly considers the viability of a Rule 23(b)(2) punitive-damage class in her article *Nine Lives: The Punitive Damage Class*.⁶⁰ She predicts that the approach of “shoe-horning” the punitive-damage class action into the Rule 23(b)(2) provision is unlikely to be well received by the Supreme Court, given the Court’s political leanings.⁶¹ Mark Perry and Rachel Brass note that plaintiffs alleging employment discrimination have “made aggressive use of Rule 23(b)(2).”⁶² They argue, however, that proponents of an expansive reading of Rule 23(b)(2) have failed to heed the Supreme Court’s decision in *Ortiz v. Fibreboard Corp.*,⁶³ which, if applied to (b)(2) class actions, strongly suggests that employment discrimination cases like *Dukes*—where significant compensatory and punitive damages are sought in addition to injunctive relief—cannot be certified under Rule 23(b)(2).⁶⁴ Thus, employment discrimination actions like *Dukes*, according to Perry and Brass, must be

55. *Id.*

56. Malveaux, *supra* note 7, at 407.

57. *Id.*

58. Dasteel & McKaig, *supra* note 7, at 1883.

59. *Id.* at 1900–02.

60. See generally Mullenix, *supra* note 7.

61. *Id.* at 886–87.

62. Mark A. Perry & Rachel S. Brass, *Rule 23(b)(2) Certification of Employment Class Actions: A Return to First Principles*, 65 N.Y.U. ANN. SURV. AM. L. 681, 681 (2010).

63. 527 U.S. 815 (1999).

64. Perry & Brass, *supra* note 62, at 700–04.

certified under Rule 23(b)(3), where defendants and absent class members are afforded greater protections.⁶⁵

The existing literature on the issue of damages in mandatory class actions, though valuable, is largely underdeveloped. An effective analysis of the issue must first examine the normative policies that inform Rule 23(b)(2) and then determine whether particular damage remedies comport with the weight attached to those respective policies.

II. The Competing Normative Policies

The most fundamental debate among class action scholars is between advocates of individual autonomy in litigation, on the one hand, and proponents of collective justice, on the other.⁶⁶ Those in the former camp tend to argue for notice and opt-out rights in most class action lawsuits where damages are being sought by the plaintiff class.⁶⁷ The proponents of collective justice, however, are less concerned with providing notice and opt-out rights to litigants, except where explicitly required (in Rule 23(b)(3) class actions).⁶⁸ While many scholars understand these two policies as being in contradistinction to one another,⁶⁹ others, like David Rosenberg, argue that

65. *Id.* at 703.

66. Shapiro, *supra* note 47, at 916. Those who lean on the individual-autonomy side of the debate include Richard Epstein and Roger Transgrud. See Richard A. Epstein, *The Consolidation of Complex Litigation: A Critical Evaluation of the ALI Proposal*, 10 J.L. & COM. 1, 3-4 (1990) (criticizing a proposal to modify the mandatory-consolidation rules in light of the low threshold for forced consolidation it foists on plaintiffs); Roger H. Transgrud, *Mass Trials in Mass Tort Cases: A Dissent*, 1989 U. ILL. L. REV. 69, 69-70 (criticizing the common practice of cutting procedural and substantive corners to expedite mass tort litigation); see also Patricia Anne Solomon, Note, *Are Mandatory Class Actions Unconstitutional?*, 72 NOTRE DAME L. REV. 1627, 1629 (1997) (arguing that mandatory class actions are unconstitutional because they violate an individual's due process right to choose an individual remedy rather than collective action). Those who are on the side of the collective-justice approach include Jack Weinstein, Robert Bone, Bruce Hay, and David Rosenberg. See JACK B. WEINSTEIN, *INDIVIDUAL JUSTICE IN MASS TORT LITIGATION: THE EFFECT OF CLASS ACTIONS, CONSOLIDATIONS, AND OTHER MULTIPARTY DEVICES* (1995) (providing various mass tort case studies and making recommendations to courts that attempt to balance considerations of efficiency and cost for a class of plaintiffs against the need for individually effective remedies); Bone, *supra* note 49, at 294 (noting that class members have the same goals for the suit, making notice and opt-out requirements for 23(b)(3) actions superfluous); Bruce L. Hay, *Asymmetric Rewards: Why Class Actions (May) Settle for Too Little*, 48 HASTINGS L.J. 479, 480-81 (1997) (proposing that although class actions often settle for too little, the solution to this problem is not to curtail the use of the class action altogether but rather to adjust class counsels' incentives to settle); Rosenberg, *supra* note 44, at 586-93 (contending that the economic benefits of class action suits outweigh fairness concerns, especially when the concern for individual autonomy threatens the economic feasibility of pursuing a class action in the first place).

67. See Shapiro, *supra* note 47, at 918 ("Under this view, . . . the individual retains his own counsel, retains the right to leave the group before, during, and after the litigation, and can insist on playing a significant role in the operations of the group so long as he chooses to remain a part of that group.").

68. *Id.* at 937-38.

69. See, e.g., Bone, *supra* note 49, at 215 ("The premise of adjudicative representation—that persons can do the litigating work for one another—seems at odds with a belief that individual litigants ought to control their own lawsuits."); Shapiro, *supra* note 47, at 918-19 (contrasting the

this understanding actually presents a false dichotomy and that collective justice (what he calls “bureaucratic justice”) actually achieves individual autonomy objectives.⁷⁰

This part argues that these two policies are necessarily distinct and that they are implicated to varying degrees depending on the kind of class action lawsuit being brought. In (b)(2) class actions, where the remedy being sought is ordinarily group injunctive relief, less emphasis is placed on the right to individual participation. By contrast, in (b)(3) class actions, where individual damages are the primary relief sought by the class, litigant autonomy is implicated to a much greater extent. Complications arise when, as in *Dukes*, the plaintiff class seeks certification under Rule 23(b)(2) but also seeks sizeable damages in addition to group-wide injunctive relief. This complication will be explored in further detail in Part III.

Before examining the normative policies that underlie (b)(2) class actions, it is useful to know a bit about the history and purpose behind Rule 23(b)(2). The language of the advisory committee’s note to Rule 23(b)(2) is particularly helpful in understanding these matters. First, the note clarifies that the drafters of Rule 23 intended for the (b)(2) provision to “reach situations where a party has taken action or refused to take action with respect to a class” and, as such, where injunctive or declaratory relief is appropriate.⁷¹ The note expands on this notion by stating that the subdivision does not apply to “cases in which the appropriate final relief relates exclusively or predominantly to money damages.”⁷² As noted earlier, this language has served as the tipping point for the debate over the extent to which monetary damages can ever be authorized in (b)(2) class actions.⁷³

The advisory committee’s note also suggests that civil rights cases were at the forefront of the Committee’s mind when it drafted this particular subdivision.⁷⁴ The note lists a number of civil rights cases that are illustrative of the class-wide discrimination that the (b)(2) subdivision was designed to remedy.⁷⁵ That being said, the note also indicates that (b)(2) is not limited

“aggregation” model, where litigants simply aggregate their preexisting rights with no corresponding sacrifice or binding agreement, with the “entity” model, where the class action lawsuit itself is the autonomous party binding the individual plaintiffs).

70. See Rosenberg, *supra* note 44, at 567 (arguing that bureaucratic justice can provide better individual outcomes by more efficiently dealing with the problems of mass tort litigation).

71. FED. R. CIV. P. 23 advisory committee’s note (1966 Amendment).

72. *Id.*

73. See *supra* Part I.

74. See FED. R. CIV. P. 23 advisory committee’s note (1966 Amendment) (“Illustrative [of the types of classes that might fall under (b)(2)] are various actions in the civil-rights field where a party is charged with discriminating unlawfully against a class, usually one whose members are incapable of specific enumeration.”).

75. *Id.*

to civil rights cases and goes on to list a number of alternative instances in which a class action could be brought successfully under Rule 23(b)(2).⁷⁶

The advisory committee's note evinces a clear intent on the part of the drafters to fashion group remedies under Rule 23(b)(2). The drafters did not intend for the (b)(2) subdivision to provide individual remedies for individual plaintiffs; rather, it was designed to provide a remedy to an entire plaintiff class that could root out a complex legal wrong at its source. This point is aptly demonstrated by the example of school desegregation in the South. In *Potts v. Flax*,⁷⁷ one of the civil rights cases listed in the advisory committee's note, the Fifth Circuit held that the suit brought on behalf of all African-American children in Fort Worth, Texas, was in fact a class action.⁷⁸ The defendant school-board officials argued that even if the decree granting desegregation relief applied to some individual plaintiffs, it could not be applied to all similarly situated plaintiffs.⁷⁹ However, the court disagreed, explaining that "the purpose of the suit was not to achieve specific assignment of specific children to any specific grade or school" but rather to obliterate a "policy of system-wide racial discrimination."⁸⁰ This case reveals that courts sought to desegregate schools in order to root out a complex wrong—namely, racial discrimination—at its source. Rather than providing individual remedies for individual plaintiffs (e.g., allowing an African-American plaintiff to attend an all-white school), courts fashioned impersonal group remedies designed to benefit African-Americans as a group by terminating racially discriminatory segregation practices.

The advisory committee's note illuminates the tension between the competing normative policies that inform Rule 23(b)(2). In what follows in this part, I describe these two policies—individual participation and remedial efficacy—in more detail and then contrast them with the competing policies underlying (b)(3) class actions. In particular, I explain why remedial efficacy weighs more heavily in (b)(2) class actions than in (b)(3) class actions. In (b)(3) class actions, the need for remedial efficacy takes a back seat to the more important goal of achieving judicial-economy gains.

A. Individual Participation

The fundamental basis for individual participation is the value we place on every plaintiff being entitled to her personal day in court. The procedural rules that govern our system of adjudication are designed to embrace this

76. *Id.* (suggesting that 23(b)(2) could encompass actions by "a numerous class of purchasers, say retailers of a given description, against a seller alleged to have [overcharged] that class" or by "a numerous group of purchasers or licensees [of a patent] . . . to test the legality of [a] 'tying' condition" that required them to "also purchase or obtain licenses to use an ancillary unpatented machine").

77. 313 F.2d 284 (5th Cir. 1963).

78. *Id.* at 289.

79. *Id.* at 288.

80. *Id.* at 288–89.

participatory norm.⁸¹ Before delving into the way individual participation functions in the class action context, it is helpful to distinguish between two competing theories of procedural rights. The first is an outcome-based approach. Under this instrumentalist approach, an individual's right to participate is privileged only if that form of participation "is likely to further a judge's ability to make good law."⁸² Many class action scholars only approach an individual's participation right from an outcome-based perspective.⁸³ Those who do tend to ignore a more process-based approach to procedural rights. Under such an approach, scholars, relying on Kantian principles, argue that the right to participate is required in order to respect the dignity of those bound by a decision.⁸⁴ Others have argued that the right to participate is essential to the legitimacy of adjudication as a source of binding judgments, just as participation is essential to the legitimacy of legislation.⁸⁵ Distinguishing between these two approaches is helpful because it informs our analysis with respect to authorizing particular damage remedies under the various pigeonholes of Rule 23.⁸⁶

In class actions, the participatory safeguards take the form of notice and opt-out rights. This right to individual participation, however, must be balanced against the need for remedial efficacy. Striking such a balance is especially important in (b)(2) class actions, where the interests of individual class members take a back seat to the interests of the class as a whole. The issue here is the extent to which this participation norm is implicated by (b)(2) class actions.

I commence this analysis by reiterating one of the most important features of Rule 23(b)(2)—that it facilitates a class-wide remedy designed to root out a complex wrong at its source. The *Dukes* case might help illustrate the importance of this feature. The plaintiffs' complaint in *Dukes* alleged gender discrimination by Wal-Mart in its hiring and promotion practices.⁸⁷

81. See Robert G. Bone, *Procedure, Participation, Rights*, 90 B.U. L. REV. 1011, 1014 (2010) ("There is no question that . . . participation rights in particular figure prominently in current modes of justification for rules and practices."). Examples include pleading standards and summary judgment standards. See *id.* at 1011–12 ("[S]ome critics object to stricter pleading standards on the ground that strict standards impede the right of access to court . . .").

82. *Id.* at 1025.

83. See, e.g., Rosenberg, *supra* note 44, at 567 ("[There are] important intersections in the mass tort context where the ends of individual justice are better served by collective, rather than disaggregative, processes."); Shapiro, *supra* note 47, at 919 ("[I]n the situations in which class action treatment is warranted, the individual who is a member of the class . . . must tie his fortunes to those of the group with respect to the litigation, its progress, and its outcome. Of course, even this entity model does not deny the class member the opportunity to seek private advice, or to contribute in some way to the progress of the litigation, but it severely limits such aspects of individual autonomy as the range of choice to move in or out of the class or to be represented before the court by counsel entirely of one's own selection.").

84. Bone, *supra* note 81, at 1027 & n.62.

85. E.g., Lawrence B. Solum, *Procedural Justice*, 78 S. CAL. L. REV. 181, 275–77 (2004).

86. See *supra* notes 15–22 and accompanying text.

87. See *supra* note 3 and accompanying text.

On its face, the complaint resembled a paradigmatic (b)(2) class action because it functionally asked the court to put an end to an alleged company-wide gender discrimination policy. If a district court were to find in favor of the plaintiff class on the substantive merits, it would be likely to issue an injunction against Wal-Mart barring all discriminatory practices. It is unlikely that a court would issue an injunction that barred discriminatory practices against Betty Dukes in particular or against one of the other named plaintiffs. Rather, the remedy would take the form of an injunction that eliminated all discriminatory practices, thereby benefiting all female employees at Wal-Mart. Since the lawsuit focuses on the group, it is reasonable to treat participation as a group right instead of an individual right.

But what happens when the injunctive relief sought by plaintiffs is at odds with the interests of individual class members? Derrick Bell develops this idea in his famous essay *Serving Two Masters: Integration Ideals and Client Interests in School Desegregation Litigation*.⁸⁸ He argues that the interests of many African-Americans in the South were at odds with the interests of NAACP lawyers who were litigating their cases.⁸⁹ Specifically, Bell suggests that African-Americans would have been better off if the NAACP lawyers had worked to rigorously enforce the “equal” portion of the notorious “separate but equal” doctrine of *Plessy v. Ferguson*⁹⁰ rather than the *Brown v. Board of Education*⁹¹ approach of pursuing integrationist policies aimed at achieving racial balance.⁹² Bell’s essay underscores the lawyer’s obligation to represent his clients in class actions where injunctive relief is the final form of relief sought by the plaintiff class.⁹³

While Bell’s essay offers an insightful critique of the class action mechanism, it also presents an opportunity to distinguish between a plaintiff’s rights and a plaintiff’s preferences. The conflicting interests Bell describes refer to the conflicting *preferences* of individual litigants represented in a class action lawsuit. These preferences are distinguishable from (procedural) *rights*, which inhere from our body of procedural law. In the desegregation litigation discussed above, individual class members had conflicting preferences, not conflicting rights.

David Marcus further explores the relationship between procedural rights and remedial choice in the context of desegregation.⁹⁴ He suggests

88. Derrick A. Bell, Jr., *Serving Two Masters: Integration Ideals and Client Interests in School Desegregation Litigation*, 85 YALE L.J. 470 (1976).

89. *Id.* at 471–72, 512.

90. 163 U.S. 537 (1896).

91. 347 U.S. 483 (1954).

92. Bell, *supra* note 88, at 487–88.

93. See Bell, *supra* note 88, at 512 (“[S]ome civil rights lawyers . . . are making decisions . . . that should be determined by their clients . . .”).

94. See generally David Marcus, *Flawed but Noble: Desegregation Litigation and Its Implications for the Modern Class Action*, 63 FLA. L. REV. 657 (2011).

that the modern class action does not “account[] for conflicting interests coherently [or] comprehensively.”⁹⁵ While Rule 23(b)(3), according to Marcus, has “elegantly” avoided the dilemma of conflicting preferences by providing absent class members with notice and opt-out rights, Rule 23(b)(2) has not done the same.⁹⁶ When class members in (b)(2) suits have conflicting preferences with respect to a remedy, the adequacy-of-representation requirement is implicated.⁹⁷ Marcus ultimately concludes that “[a] single threshold for the adequacy of representation requirement presumes a unified basis for Rule 23 that does not fit its original design.”⁹⁸

Marcus, however, fails to draw the distinction between an outcome-based approach to individual participation and one that is more process-based. From an outcome-based perspective, accounting for conflicting preferences is less significant if doing so does not further a judge’s ability to make good law. That being said, from a process-based perspective, taking into account conflicting preferences might be more important in order to preserve the legitimacy of the process and the dignity of those bound by the decision. The focus on remedial efficacy, however, trumps any process-based participatory concerns. Accounting for conflicting preferences (presumably through notice and opt-out rights) would stunt a court’s ability to fashion a remedy designed to root out a complex wrong at its source and thereby benefit a group of claimants.

The recent amendments to Rule 23 also signal an attempt to deal with the problem of conflicting preferences among class members. In 2003, Rule 23 was amended to give courts the discretionary power to give all class members notice so that class members would have an opportunity to object to representation.⁹⁹ This additional procedural safeguard tempers problems associated with class members having disparate interests and downplays the importance of individual participation in (b)(2) suits.

B. Remedial Efficacy

The basis for remedial efficacy lies in providing the most effective remedy possible. In class actions, remedial efficacy serves the goal of rooting out at its source wrongful conduct that produces actionable harms to individual litigants. As such, in (b)(2) class actions, individuals are freed from discriminatory practices indirectly as a consequence of eliminating those discriminatory practices at their sources. In addition to serving this primary goal, the (b)(2) class action also achieves judicial-economy gains, albeit differently from (b)(3) class actions. Whereas in (b)(3) suits individual claims are adjudicated collectively, in (b)(2) suits that is not the case at all.

95. *Id.* at 712.

96. *Id.* at 712–13.

97. FED. R. CIV. P. 23(a)(4); Marcus, *supra* note 94, at 713.

98. Marcus, *supra* note 94, at 714.

99. FED. R. CIV. P. 23(c)(2); FED. R. CIV. P. 23 advisory committee’s note (2003 Amendment).

In (b)(2) class actions, judicial-economy gains are achieved by transforming individual suits into a unitary group legal challenge. If plaintiffs are successful in bringing this challenge, the source of discrimination is addressed prospectively, and individuals suffering from discriminatory treatment are thereby relieved from suing individually for prospective relief.

The collective remedy tends to be implicated to a greater degree in (b)(2) class actions because the remedies are typically aimed at a group rather than at individual plaintiffs. Since declaratory and injunctive relief are both essentially “group” remedies, it makes more sense to place a larger premium on the normative policy that is rooted in providing the most effective remedy and a smaller premium on one that embraces individual participation. Therefore, putting the issue of damages aside once again, the need for remedial efficacy is enhanced in Rule 23(b)(2) class actions, through which claimants primarily seek injunctive or declaratory relief.

C. Rule 23(b)(3)

It is useful as a heuristic device to compare the policies that underlie (b)(2) class actions with the policies that underlie (b)(3) class actions. Such a comparison allows for a more complete understanding of the circumstances in which we might value individual participation to a greater extent. Rule 23(b)(3) imposes two additional requirements—predominance and superiority—on those seeking certification under the subdivision.¹⁰⁰ Rule 23 also requires courts to provide notice and opt-out rights to (b)(3) class members.¹⁰¹ Provided that class claimants are able to fulfill the predominance and superiority requirements, they are free to seek damages as the only form of relief. These requirements represent the essential differences between Rule 23(b)(3) and Rule 23(b)(2).

Remedial efficacy tends not to be the centerpiece of (b)(3) class actions, where damages, rather than injunctive or declaratory relief, are the primary remedy being sought. When damages are the primary form of relief being sought by the class, a greater emphasis is placed on the right to individual participation. In contrast to injunctive relief, damages tend to be more individualized in nature because they typically compensate individual claimants for individual harms they suffered.¹⁰² The notice and opt-out rights afforded to all (b)(3) class members reflect the premium we place on individual participation in (b)(3) class actions. When damages are being sought, it is imperative that class members be given notice and opt-out rights to ensure that they have the ability to re-litigate their individual damage claims in a separate legal forum. This stands in contrast to (b)(2) class actions, where the remedy most commonly sought—injunctive relief—is directed primarily

100. See *supra* note 20 and accompanying text.

101. See *supra* note 21 and accompanying text.

102. Of course, this depends on the type of damages being sought, but I will explore this issue in more detail in the next part.

at a group, not at an individual. Thus, it seems apt to provide these kinds of procedural safeguards when the remedy being sought is individualized as opposed to when it is directed primarily toward a group.

As opposed to (b)(2) class actions, the primary function of (b)(3) class actions is to achieve judicial-economy gains by adjudicating individual suits collectively. While (b)(2) class actions also achieve judicial-economy gains (as described above), the primary goal of (b)(2) class actions lies in providing the most effective remedy possible. In (b)(3) class actions, where the plaintiff seeks damages, the focus is on adjudicating similar disputes with overlapping issues. In (b)(2) class actions, where declaratory or injunctive relief is primarily sought, the primary purpose is to design a remedy that roots out a complex wrong at its source. As such, the need for notice and opt-out provisions in (b)(2) class actions is less significant than in (b)(3) class actions, where the remedy is aimed at compensating individuals for injuries they suffered.

III. Damage Remedies in (b)(2) Class Actions

Thus far, the analysis paints a fairly neat picture with respect to the class action device. The two normative policies at stake in a (b)(2) class action are the right to individual participation and the need for remedial efficacy.¹⁰³ In (b)(2) class actions, where the primary relief sought is declaratory or injunctive, remedial efficacy is privileged to a greater degree than individual participation because declaratory and injunctive relief are both group remedies. Conversely, in (b)(3) class actions, where the primary relief sought takes the form of monetary damages, individual participation is privileged to a greater extent than remedial efficacy because damages tend to be individualized in nature. As such, the (b)(3) subdivision provides notice and opt-out rights for all class members as a way of protecting individual claimants and embracing individual participation.

Once damages enter the (b)(2) scene, however, the picture becomes murkier. The policy balance articulated above is disrupted when (b)(2) claimants seek damages in addition to the injunctive and declaratory relief that Rule 23(b)(2) typically provides. This is precisely the issue that the courts have confronted in *Allison* and *Robinson*, and one that the Supreme Court recently addressed in *Dukes*.

The issue of (b)(2) damages raises a number of interesting analytical questions, some of which I attempt to answer in this Note. First and foremost, what is the effect of allowing class members to seek damages under Rule 23(b)(2)? Does allowing plaintiffs to seek damages comport with the normative policies that underlie the Rule? If doing so tends to embrace individual participation more so than when injunctive relief was primarily sought, should we provide (b)(2) class members with notice and opt-out

103. See *supra* Part II.

rights, like we do under Rule 23(b)(3)? But then, what would be the difference between a (b)(2) class action and a (b)(3) class action? After all, if plaintiffs can seek damages under Rule 23(b)(2), are they not more likely to seek certification under this subdivision since certification under Rule 23(b)(3) requires fulfilling the predominance and superiority requirements? Finally, how should we decide whether to allow plaintiffs to seek damages; does it make sense to adopt one of the tests articulated by the circuit courts or is there a better method?

Scholars who support the approach taken by the *Allison* court have argued that allowing plaintiffs to seek damages under Rule 23(b)(2) (beyond what the incidental-damages test permits) effectively serves as an end run around the Rule 23(b)(3) requirements of predominance and superiority; specifically, they argue that if plaintiffs' attorneys have the option of certifying a class under Rule 23(b)(2) or Rule 23(b)(3), they will often select the former option because it obviates the need to fulfill the predominance and superiority requirements of Rule 23(b)(3).¹⁰⁴ The result is problematic in two respects: first, it exposes defendants to tremendous potential liability without requiring the putative class to fulfill the (b)(3) threshold requirements; second, it compromises the interests of class members at the cost of satiating plaintiffs' attorneys because class members are no longer afforded notice and opt-out rights.

In the ensuing subparts, I address a number of these concerns. I argue that the most logical approach to analyzing the issue of monetary damages in mandatory class actions is to ask if allowing damages comports with the normative policies that inform the Rule. I attempt to make this determination in the context of plaintiffs seeking backpay, compensatory damages, and punitive damages.¹⁰⁵ Since each of the three damage remedies have different purposes, each should be analyzed in accordance with the individual purposes it is meant to serve.

A. *Backpay*

Backpay is one form of damages commonly sought by plaintiffs. It is determined by calculating the difference between what an employee was paid and what he or she should have been paid.¹⁰⁶ There is very little debate in the courts and even among legal scholars over whether class members can seek backpay under Rule 23(b)(2). Courts have traditionally authorized backpay

104. *E.g.*, Mullenix, *supra* note 2, at 251.

105. Of course, the three categories are not mutually exclusive. Plaintiffs can seek one, two, or all three. This, however, does not affect my analysis.

106. *See* BLACK'S LAW DICTIONARY 158–59 (9th ed. 2009) (defining *backpay* as “[t]he wages or salary that an employee should have received but did not because of an employer’s unlawful action in setting or paying the wages or salary”).

because it has been viewed as secondary to injunctive relief typically being sought by (b)(2) class members.¹⁰⁷

Authorizing backpay comports with the relative weight given to the normative policies underlying (b)(2) class actions. Much like declaratory or injunctive relief, backpay is fundamentally a group remedy. It is aimed at uniformly compensating all class members for what they should have been paid in the absence of the challenged discriminatory policy. Since it tends not to require individualized determinations, backpay does not implicate the right to individual participation in such a way as to upset the balance of the competing policies that inform (b)(2) class actions.

The *Dukes* case, however, questions whether backpay is in fact a group remedy. Wal-Mart argued that because backpay awards are discretionary, courts must exercise discretion and conduct an individual analysis with respect to each plaintiff's claim.¹⁰⁸ As it would take a trier of fact years to conduct additional proceedings for each plaintiff's claim, backpay was not "incidental" to the injunctive relief being sought.¹⁰⁹ Finally, Wal-Mart argued that per Title VII and the Due Process Clause, it had a right to litigate each of the plaintiffs' backpay claims.¹¹⁰

The Supreme Court's holding, which disallowed backpay, is problematic for two reasons. First, as the plaintiffs correctly argued, backpay does not qualify as "monetary damages" but rather constitutes an equitable remedy under Title VII.¹¹¹ As such, backpay is distinguishable from both compensatory and punitive damages, which are both understood to be forms of compensatory relief under Title VII. Second, as the Supreme Court noted in *Albemarle Paper Co. v. Moody*,¹¹² backpay is critical to Title VII's make-whole remedial scheme.¹¹³ Injunctive relief by itself will not deter employers from engaging in arguably discriminatory practices, but the threat of backpay would make deterrence more effective.¹¹⁴

The answer is much less clear with respect to Wal-Mart's claim that it is entitled to present individual defenses for each of the backpay claims according to Title VII. To deny Wal-Mart such a right, as the district court did, would be to strip away a fundamental right that it has been conferred by

107. See, e.g., *Allison v. Citgo Petroleum Corp.*, 151 F.3d 402, 415 (5th Cir. 1998) ("We construe[] (b)(2) to permit monetary relief when it [is] an equitable remedy, and the defendant's conduct ma[kes] equitable remedies appropriate. Back pay, of course, ha[s] long been recognized as an equitable remedy under Title VII." (citations omitted)).

108. Brief for Petitioner at 53–55, *Wal-Mart Stores, Inc. v. Dukes*, 131 S. Ct. 2541 (2011) (No. 10-277).

109. *Id.* at 55; see also *supra* notes 31–33 and accompanying text.

110. Brief for Petitioner, *supra* note 108, at 42–43.

111. Brief for Respondents at 57, *Dukes*, 131 S. Ct. 2541 (No. 10-277) (citing 42 U.S.C. §§ 1981a(b)(2), 2000e-5(g)).

112. 422 U.S. 405 (1975).

113. *Id.* at 419–21.

114. *Id.* at 417–18.

statute. That being said, it is also important to consider the impracticality of allowing Wal-Mart to exercise this right under these circumstances. Doing so would undermine the paramount goal of remedial efficacy at the cost of privileging Wal-Mart's individual right to participation. If we believe that allowing Wal-Mart to litigate individual defenses would result in better outcomes, then perhaps it makes more sense to afford them such a right. However, if this right is being preserved purely on process-based grounds (i.e., to maintain dignity, legitimacy, etc.), then it makes more sense to forego Wal-Mart's statutory right in lieu of fashioning a more effective group remedy.

In sum, backpay fits the remedial focus of (b)(2) class actions because it effectively serves as a counterpart to the injunctive relief typically aimed at rooting out a complex wrong at its source. Because backpay is essentially a group remedy as it is classified as equitable per Title VII, and because allowing Wal-Mart to exercise its statutory right per Title VII would not necessarily result in a better outcome, the Court incorrectly decided the backpay issue in *Dukes*.

B. Compensatory Damages

In addition to backpay, plaintiffs also often seek compensatory damages, which are intended to cover actual injury or economic loss.¹¹⁵ The plaintiffs in *Dukes* did not seek compensatory damages. Courts are much less uniform on how they treat compensatory-damage relief under Rule 23(b)(2). Courts that tend to follow the Fifth Circuit's approach in *Allison* typically disallow compensatory damages,¹¹⁶ whereas courts that tend to follow the Second Circuit's more liberal ad hoc approach in *Robinson* are more amenable to allowing compensatory damages in (b)(2) class actions.¹¹⁷ The focus of this split is on the "predominance" language in the advisory committee's note to Rule 23.

If, instead of relying on the language of the advisory committee's note, we focus on the normative policies underlying the Rule, the analytical framework is different. First, we must ask how compensatory damages should be characterized—that is to say, do compensatory damages tend to be

115. See BLACK'S LAW DICTIONARY 445 (9th ed. 2009) (defining *compensatory damages* as "[d]amages sufficient in amount to indemnify the injured person for the loss suffered").

116. See, e.g., *Barabin v. Aramark Corp.*, No. 02-8057, 2003 WL 355417, at *1-3 (3d Cir. Jan. 24, 2003) (denying class certification under 23(b)(2) after finding that the compensatory and punitive damages requested by the class were not incidental to the injunctive relief sought); *Murray v. Auslander*, 244 F.3d 807, 812 (11th Cir. 2001) (holding that a district court abused its discretion by not excluding individual compensatory damages claims from class treatment where the class was certified under 23(b)(2)); *Jefferson v. Ingersoll Int'l, Inc.*, 195 F.3d 894, 898 (7th Cir. 1999) (holding that non-incidental compensatory damages cannot be sought in a (b)(2) class because of the lack of notice and an opportunity to opt out).

117. See, e.g., *Molski v. Gleich*, 318 F.3d 937, 949-50 (9th Cir. 2003) (rejecting the Fifth Circuit's bright-line dichotomy of incidental and non-incidental money damages in favor of a looser, more fact-intensive approach to determining predominance).

more individual in nature or are they more like a group remedy? Compensatory damages are in fact more individualized. They attempt to make *individual* plaintiffs whole for *individual* wrongs perpetrated against them. If we accept that compensatory damages are inherently individual in nature, then allowing plaintiffs to seek compensatory damages in (b)(2) actions would implicate the right to individual participation to a greater degree.

Then, one might ask, how can we remedy this balancing problem? One solution might be to afford (b)(2) plaintiffs who seek compensatory damages notice and opt-out rights. Notice and opt-out rights are problematic, however, because affording (b)(2) plaintiffs such procedural protections would frustrate the purpose behind Rule 23(b)(3). Allowing for notice and opt-out rights in (b)(2) class actions would essentially serve as an end run around the predominance and superiority requirements of the (b)(3) subdivision.¹¹⁸ Consequently, defendants would be exposed to potentially tremendous liability without plaintiffs even fulfilling the threshold requirements of Rule 23(b)(3).

Perhaps, then, it makes less sense to allow plaintiffs to pursue compensatory damages in a (b)(2) class action. To the extent that the compensatory damages being sought are individual in nature, it makes more sense to require plaintiffs to pursue those damages under Rule 23(b)(3), where plaintiffs have to meet the requirements of predominance and superiority, and plaintiffs are afforded both notice and opt-out rights. If, however, the compensatory damages sought by plaintiffs focus on the group, then perhaps such relief should be permitted under Rule 23(b)(2).

It is important to note that the approach I propose is significantly different from the one taken by most circuit courts. Rather than attempting to define *predominance* in the context of the advisory committee's note, we should direct our attention to the remedial focus of the damages being sought. If the damages tend to single out individual plaintiffs for individual wrongs perpetrated against them, then it makes more sense to disallow those compensatory damages under Rule 23(b)(2). Conversely, if the damages sought are designed to remedy a wrong committed against a group, then those damages should be permitted under Rule 23(b)(2).

C. Punitive Damages

Whether punitive damages can be authorized under Rule 23(b)(2) is perhaps the most controversial issue. Punitive damages are different from compensatory damages in that their purpose is to punish a particular individual or entity.¹¹⁹ In *Dukes*, the plaintiffs sought punitive damages in addition

118. See *supra* note 104 and accompanying text.

119. See BLACK'S LAW DICTIONARY 448 (9th ed. 2009) (defining *punitive damages* as "[d]amages awarded in addition to actual damages when the defendant acted with recklessness,

to traditional backpay.¹²⁰ In its brief, Wal-Mart pointed out that no court of appeals had ever authorized punitive damages in a Rule 23(b)(2) class action and that if the Supreme Court were to authorize punitive damages in this case, it would be straying from precedent.¹²¹ The plaintiffs disputed this contention but then reminded the Court that the Ninth Circuit remanded the issue of whether punitive damages could be sought under Rule 23(b)(2) down to the district court and that, therefore, the Supreme Court did not need to confront the issue.¹²² For this reason, the Supreme Court did not reach this important issue in *Dukes*.

We might ask why courts have historically been so averse to awarding punitive damages in (b)(2) class actions. We can only speculate, but one might assume that this is because punitive damages are typically awarded *in addition to* compensatory damages.¹²³ If courts are reluctant to allow plaintiffs to pursue compensatory damages under Rule 23(b)(2), one might surmise that punitive damages would then surely be outside the realm of possibility. I find this logic to be unsound.

If instead we look back to the normative policies underlying the Rule, we find that the punitive-damage remedy actually comports with that policy balance. Punitive damages are more like a group remedy than an individual one; their purpose is to deter wrongful behavior by said entities or similar entities.¹²⁴ Moreover, their focus is not on compensating the individual for any harm done to him or her—that purpose is served by compensatory damages.¹²⁵

Thus, if we conclude that the focus of punitive damages is on punishing the entity who perpetrated the wrong as opposed to compensating individuals who were wronged by the entity, then we reach a different result than the courts. Since punitive damages do not implicate the individual-participation norm, the relative weight it is afforded in (b)(2) class actions remains the same. As such, it makes more sense for courts to allow plaintiffs to seek punitive damages under Rule 23(b)(2). In fact, it makes more sense for courts to allow plaintiffs to seek punitive damages than to allow them to seek compensatory damages.

malice, or deceit; specifically], damages assessed by way of penalizing the wrongdoer or making an example to others”).

120. *Wal-Mart Stores, Inc. v. Dukes*, 131 S. Ct. 2541, 2548 (2011).

121. Brief for Petitioner, *supra* note 108, at 55–56.

122. Brief for Respondents, *supra* note 111, at 64.

123. See Dan Markel, *Retributive Damages: A Theory of Punitive Damages as Intermediate Sanction*, 94 CORNELL L. REV. 239, 241–43 (2009) (defining punitive damages as “extra-compensatory” damages and explaining that such extra-compensatory damages “best calibrated in reference to a defendant’s likelihood of evading payment of full compensatory damages”).

124. See RESTATEMENT (SECOND) OF TORTS § 908(1) (1979) (“Punitive damages are damages, other than compensatory or nominal damages, awarded against a person to punish him for his outrageous conduct and to deter him and others like him from similar conduct in the future.”).

125. See *id.* § 903 (“‘Compensatory damages’ are damages awarded to a person as compensation, indemnity or restitution for harm sustained by him.”).

IV. Conclusion

The debate over whether to allow monetary damages in Rule 23(b)(2) class actions has been contentious, to say the least. In this Note, I have attempted to reassess the way we normally think about damages in class action lawsuits. Instead of focusing on interpreting the language of the advisory committee's note, as most courts have, I first identified the normative policies that underlie (b)(2) class actions. Those two policies are the right to individual participation and the need for remedial efficacy. I then showed how these two policies are weighed depending on the particular subdivision of Rule 23. Specifically, I contrasted the relative weight they are afforded in (b)(2) class actions with the weight they are given in (b)(3) class actions. While individual participation is implicated to a greater degree in (b)(3) class actions, that is not the case in (b)(2) class actions, where the need for remedial efficacy is of primary importance.

I then asked whether allowing for certain damage remedies under Rule 23(b)(2) comports with the relative weight afforded to the normative policies underlying the Rule. I concluded that while it might be problematic for courts to authorize compensatory damages in (b)(2) class actions, courts should be more willing to authorize backpay and punitive damages. While compensatory damages are more individualized by nature, punitive damages and backpay are both inherently group remedies. They are aimed less at compensating individual plaintiffs and more at deterring defendants' wrongful behavior.

—Neil K. Gehlawat

* * *

Would You Want William Hung[†] as Your Trier of Fact? The Case for a Specialized Musicology Tribunal^{*}

I. Introduction

Few analyses of copyright law do not begin with the premise of constitutional mandate as espoused in Article I, Section Eight of the United States Constitution. Never punk enough to rebel, this author will capitulate. Congress is explicitly granted the power to “promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.”¹ The English Statute of Anne, widely revered as the foundation of modern copyright law, was billed as “[a]n Act for the Encouragement of Learning, by Vesting the Copies of Printed Books in the Authors or Purchasers of such Copies, during the Times therein mentioned.”² The bridge from providing remuneration for the bookmakers’ “very great Detriment . . . too often to the Ruin of them and their Families”³ to the Sonny Bono Copyright Term Extension Act’s (CTEA) addition of twenty years of protection⁴ to the arguably eternal duration of copyright⁵ seems anything but logical. Surely, some would explain away these different approaches as the inevitable result of technological innovation. Such a construction, while truthful to a fault, belies an integral

† The author is not blind to the possibility that this reference will be dated before it has even been made. (Indeed, that would be both appropriate and ideal.) William Hung is the Chinese-American University of California, Berkeley civil-engineering student whose off-key rendition of Ricky Martin’s “She Bangs” during the third season of *American Idol* made him a household name. He has since parlayed the “fame” (or perhaps his particular phenomenon would more appropriately be described as infamy) into numerous television, film, and concert appearances, and even a record deal with Koch Entertainment. *Biography Page*, WILLIAMHUNG.NET, <http://www.williamhung.net/bio.html>.

* I want to express my eternal gratitude and zealous affection for the following persons and entities for their invaluable assistance with this Note: Chris Harrison, Professor Oren Bracha, Professor David Sokolow, Kristin Malone, Jamie France, Darius Dills, Zac Padgett, Amelia Rendeiro, the underappreciated members and the indefatigable editors of Volumes 89 and 90 of the *Texas Law Review*, Thai Kitchen on Guadalupe, Torchy’s, Shaindela, McWhitey, the *Lev Lev*, Lunarcy Pictures, and Robert Zimmerman. Finally, I wish to thank my parents Bob and Harriet, my sister Rachel, and my grandparents Leo, Noima, Irving, and Hilda for their everlasting love and support.

1. U.S. CONST. art. I, § 8, cl. 8.

2. Statute of Anne, 1710, 8 Ann., c. 19 (Eng.), available at <http://www.copyrighthistory.com/anne.html>.

3. *Id.*

4. Sonny Bono Copyright Term Extension Act, Pub. L. No. 105-298, § 102(b), 112 Stat. 2827, 2827 (1998) (codified as amended at 17 U.S.C. § 302 (2006)).

5. See *Eldred v. Ashcroft*, 537 U.S. 186, 221–22 (2003) (dismissing a public-domain enthusiast’s First Amendment challenges to the CTEA’s expansion of copyright durations).

fallacy of perspective. The Statute of Anne was structured “for the Encouragement of Learned Men to Compose and Write” and to reverse the “Ruin of [Authors or Proprietors] and their Families.”⁶ By contrast, the CTEA worked to prolong existing financial enterprises.⁷ Somewhere along the way, the statutory motives for enacting copyright legislation moved dramatically away from incentivization.⁸

This Note suggests a complete reform of the present judicial treatment of musical-copyright-infringement cases, the history and pertinent details of which will be outlined in Part II. The larger issue of the varying motivations of American copyright law is beyond this Note’s scope. Nevertheless, this lesson in perspective is necessarily applicable to the numerous calls for improvement in the enforcement of the musical copyrights. If the existing American copyright regime is to have any legitimacy in the eyes of the public and the musicians most in need of its consistent functionality, it must alter its perspective.

Tim Westergren, founder of Pandora,⁹ relayed a story in a recent *New York Times Magazine* article:

[There was a] Pandora user who wrote in to complain that he started a station based on the music of Sarah McLachlan, and the service served up a Celine Dion song. [Westergren] “wrote back and said, ‘Was the music just wrong?’ Because we sometimes have data errors,” [Westergren] recounts. “He said, ‘Well, no, it was the right sort of thing—but it was Celine Dion.’ [Westergren] said, ‘Well, was it the set, did it not flow in the set?’ He said, ‘No, it kind of worked—but it’s Celine Dion.’ [Westergren and the user] had a couple more back-and-forths, and finally [the user’s] last e-mail to [Westergren] was: ‘Oh, my God, I like Celine Dion.’”¹⁰

As this humorous anecdote evinces, Westergren’s popular website has unearthed an important and revelatory realization: music is, at its core, an objectively classifiable and mathematical art form.¹¹ As such, any test for

6. Statute of Anne, 1710, 8 Ann., c. 19 (Eng.).

7. See Brief for Intellectual Property Law Professors as Amici Curiae Supporting Petitioners at 19, *Eldred*, 537 U.S. 186 (No. 01-618), 2002 WL 1041866 (“The only effect of these amendments is to give a windfall to current proprietors of works first published or registered during the 55 years between 1923 and 1978 . . .”).

8. See Neil Weinstock Netanel, *Copyright and a Democratic Civil Society*, 106 YALE L.J. 283, 353 (1996) (“Prior to the first modern copyright statutes in the eighteenth century, writers and artists were heavily dependent on royal, feudal, and church patronage for their livelihoods. This dependency undermined expressive autonomy and thwarted the development of a vital, freethinking intelligentsia.” (footnotes omitted)).

9. Pandora is an Internet radio service that utilizes an online database that helps create “a much more personalized radio experience—stations that play music you’ll love—and nothing else.” *About the Music Genome Project*, PANDORA, <http://www.pandora.com/corporate/mgp>.

10. Rob Walker, *The Song Decoders*, N.Y. TIMES, Oct. 18, 2009, § 6 (Magazine), at 48.

11. See *id.* (“What Pandora’s system largely ignores is, in a word, taste.”). This runs contrary to the way similar sites work. See *id.* at 50 (“[T]he idea is that the taste of your cool friends, your

musical infringement that analyzes potentially illicit activity solely based on a lay, subjective understanding of music is inherently problematic. Musical copyrights, unique in their duality, deserve a uniform infringement doctrine grounded in a sound, logical policy rationale specifically derived from the exclusively scientific art form of music.

This Note advocates the formation of academically constituted music tribunals to create stability and credibility. Specifically, Part III will address why musical copyrights are worthy of special analysis, separate and distinct from other enumerated works. Part IV will delve into the academic, logistical, and procedural failures of the present legal treatment of musical copyrights. Part V will outline a solution and the rationales for its immediate implementation. Part VI will conclude the Note. There will be no encore.¹²

II. History of Music Copyright Law in the United States

A. Statutory Formulation

The Statute of Anne provided the impetus for the new nation under President George Washington to pass the first American copyright act in 1790.¹³ Over time, the 1790 Act grew in scope, eventually extending past “books, maps, and charts” to include “prints, . . . dramatic works, photographs, graphic works, and sculpture.”¹⁴ In 1831, Congress explicitly added musical compositions and further expanded protection via a performance right in 1897.¹⁵ In 1909, Congress passed the Act’s most “significant overhaul . . . since its founding,” expanding the duration and scope of protection.¹⁶ Congress enacted a law providing separate copyright protection for sound recordings in 1971.¹⁷ This Note is only concerned with sound recordings of musical performances,¹⁸ most of which are “owned by five

peers, the traditional music critics, big-label talent scouts and the latest influential music blog are all equally irrelevant. That’s all cultural information, not musical information.”).

12. No matter how many lighters are held up for the author, or how many times readers chant the author’s name and arguments in support, the proverbial house lights and PA music will come on, and the Note will be, quite simply, finished.

13. See ROBERT P. MERGES ET AL., *INTELLECTUAL PROPERTY IN THE NEW TECHNOLOGICAL AGE* 385–86 (rev. 4th ed. 2007) (noting that most of the American states passed copyright laws modeled on the Statute of Anne shortly after independence and that the federal Copyright Act of 1790 was very similar to the Statute of Anne).

14. *Id.* at 386.

15. *Id.* at 440.

16. *Id.* at 386.

17. *Id.* at 387.

18. Although there are a variety of nonmusical performance types, there has occasionally been controversy in the application of overlap. Rarely has contentiousness been as prevalent as in the suit between avant-garde composer John Cage and British songwriter Mike Batt. See Dennis Kurzon, *Peters Edition v. Batt: The Intertextuality of Silence*, 20 INT’L J. FOR SEMIOTICS L. 285, 285–88 (2007) (explaining how Batt was forced via an out-of-court settlement to pay for the use of “silence” in an audio recording); see also Mark Steyn, *You Have the Right to Remain Silent, but If You Do You May Be Liable to Considerable Damages*, TELEGRAPH (U.K.) (July 20, 2002),

major record labels—Universal Music Group, Sony Music Entertainment, Warner Brothers Music, BMG Entertainment, and EMI Group.”¹⁹ 1976 saw the passage of a new Copyright Act, “which continues to serve as the principal framework for copyright protection in the United States.”²⁰

Section 102(a) of the 1976 Act states that “[c]opyright protection subsists . . . in original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device.”²¹ Subsections 102(a)(2) and 102(a)(7), respectively, plainly list “musical works, including any accompanying words”²² and “sound recordings.”²³ The Act encourages real innovation and creation, clarifying that “[i]n no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery.”²⁴

The last two decades have seen several legislative calls for reform. The Digital Performance Right in Sound Recordings Act of 1995 was aimed solely at the sound recording copyright.²⁵ For the first time, Congress provided a public-performance entitlement to owners of sound recordings, albeit subject to various limitations and exemptions.²⁶ The aforementioned CTEA added twenty years to the duration of copyrights. The Digital Millennium Copyright Act attempted to modernize the 1976 legislation via the provision of anticircumvention techniques and the proffering of a “safe harbor” for online service providers adhering to certain requirements.²⁷ All three were the results of costly and lengthy lobbying campaigns from content owners.²⁸

<http://www.telegraph.co.uk/comment/personal-view/3579362/You-have-the-right-to-remain-silent-but-if-you-do-you-may-be-liable-to-considerable-damages.html> (describing a fictitious legal action involving the notable silences of Mike Batt, John Cage, comedian Jack Benny, the British Unknown Soldier, and Osama Bin Laden).

19. W. Jonathan Cardi, *Über-Middleman: Reshaping the Broken Landscape of Music Copyright*, 92 IOWA L. REV. 835, 848 (2007).

20. MERGES ET AL., *supra* note 13, at 387.

21. 17 U.S.C. § 102(a) (2006).

22. *Id.* § 102(a)(2).

23. *Id.* § 102(a)(7).

24. *Id.* § 102(b).

25. Digital Performance Right in Sound Recordings Act of 1995, Pub. L. No. 104-39, 109 Stat. 336 (codified as amended at 17 U.S.C. §§ 106, 114–15 (2006)).

26. *Id.* § 3, 109 Stat. at 336–44.

27. MERGES ET AL., *supra* note 13, at 585, 589–90.

28. See, e.g., *Arista Records, LLC v. Launch Media, Inc.*, 578 F.3d 148, 152–53 (2d Cir. 2009) (attributing the enactment of the Digital Performance Right in Sound Recordings Act of 1995 to “continued lobbying by the recording industry”); MERGES ET AL., *supra* note 13, at 585 (describing the pressure exerted by content owners seeking increased anticircumvention prohibitions that resulted in Congressional enactment of the DMCA); Jonathan P. Decker, *Of Mice and (Congress)Men*, FORTUNE, Nov. 23, 1998, available at http://money.cnn.com/magazines/fortune/fortune_archive/1998/11/23/251415/index.htm (reporting extensive lobbying by Disney and motion-picture companies for the passage of CTEA, including campaign contributions to the Act’s sponsors).

A successful copyright infringement suit may entitle the plaintiff to several remedies including injunctions, actual damages, a percentage of the defendant's profits from the infringing work, statutory damages, and costs and attorneys' fees.²⁹

B. *Judicial Formulation*

In 1991, the Supreme Court clarified the two elements required for a *prima facie* copyright infringement action.³⁰ The plaintiff must prove both "ownership of a valid copyright" and "copying of constituent elements of the work that are original."³¹ With the first prong, a "copyright registration certificate constitutes *prima facie* evidence in favor of the plaintiff."³² In the absence of formal registration with the United States Copyright Office, independent authorship of a protectable work will suffice.³³ Given the relaxation of formality requirements in the 1976 Act, it is rare in the modern era for the first prong to be specifically contested.³⁴

In *Arnstein v. Porter*,³⁵ the Second Circuit held the prong of "copying" to constitute "two separate elements . . . : (a) that defendant copied from plaintiff's copyrighted work and (b) that the copying (assuming it to be proved) went so far as to constitute improper appropriation."³⁶ There are several methods of demonstrating the first prong of "copying." Where possible, courts look favorably upon "direct proof of copying—e.g., eyewitness testimony, records indicating that one author obtained the work from another, videotape of direct copying, or distinctive flaws (such as unusual errors that are common to the two works)."³⁷ More likely,

29. 17 U.S.C. §§ 502–505 (2006).

30. *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 361 (1991).

31. *Id.*

32. 4 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 13.01[A] (2011) [hereinafter NIMMER ON COPYRIGHT].

33. See MERGES ET AL., *supra* note 13, at 405 ("U.S. law has progressed from a regime in which failure to adhere to certain technical requirements resulted in forfeiture of copyright protection to the current regime in which formalities are largely voluntary and failure to comply does not risk forfeiture."). A common—albeit legally ineffective—practice among musicians unwilling or unable to pay the Copyright Office's required fees for registration is to mail a copy of one's work to oneself. See *Copyright in General*, U.S. COPYRIGHT OFFICE (July 12, 2006), <http://www.copyright.gov/help/faq/faq-general.html> ("The practice of sending a copy of your own work to yourself is sometimes called a 'poor man's copyright.' There is no provision in the copyright law regarding any such type of protection, and it is not a substitute for registration.")

34. *But see generally* *Williams v. Broadus*, 60 U.S.P.Q.2d (BNA) 1051 (S.D.N.Y. 2001) (exemplifying the difficulties presented by the rap genre's practice of sampling, whereby summary judgment for defendants was deemed inappropriate even though the plaintiffs' copyright was itself of questionable validity given its unlicensed usage of Otis Redding's "Hard to Handle").

35. 154 F.2d 464 (2d Cir. 1946).

36. *Id.* at 468.

37. MERGES ET AL., *supra* note 13, at 475; *see also* *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 344 (1991) (utilizing "fictitious listings that [plaintiff] had inserted into its [phone] directory to detect copying").

“circumstantial evidence—usually evidence of access—from which the trier of the facts may reasonably infer copying” will be utilized.³⁸

Most frequently, a demonstration of “access plus substantial similarity” is used by plaintiffs to satisfy the prong.³⁹ As the *Arnstein* citation implies, “evidence of striking similarity sometimes permits a finding of copying without proof of access. Thus, the stronger the proof of similarity, the less the proof of access that is required.”⁴⁰ Though the Second Circuit has expressly stated otherwise, this standard would seem to indicate that less similarity is required with a finding of access.⁴¹ From a circular standpoint, much confusion exists because “[c]ourts have inconsistently defined the parameters of ‘substantial similarity’ for copyright infringement purposes.”⁴² As one court pointed out, “[s]triking similarity’ is not merely a function of the number of identical notes that appear in both compositions. . . . An important factor in analyzing the degree of similarity of two compositions is the uniqueness of the sections which are asserted to be similar.”⁴³ Furthermore, regardless of how striking the similarity, a provision of evidence showing that the creation was done independently can overcome this inference.⁴⁴ Also, the copying need not even be deliberate in any sort of conscious sense. In 1976, Beatle George Harrison lost an infringement case where a jury held his “My Sweet Lord” to have infringed upon Ronnie Mack’s “He’s So Fine.”⁴⁵ The court held this prong satisfied where Harrison’s “subconscious knew it already had worked in a song his conscious mind did not remember.”⁴⁶

38. *Arnstein*, 154 F.2d at 468. The court proposed a spectrum scale, whereby it seems apparent that this prong cannot be satisfied by mere “access” under any circumstances. *See id.* (“[I]f there are no similarities, no amount of evidence of access will suffice to prove copying. If there is evidence of access and similarities exist, then the trier of the facts must determine whether the similarities are sufficient to prove copying. . . . If evidence of access is absent, the similarities must be so striking as to preclude the possibility that plaintiff and defendant independently arrived at the same result.”).

39. 4 NIMMER ON COPYRIGHT, *supra* note 32, at § 13.02[A].

40. *Id.* § 13.01[D] (footnote omitted).

41. *See Arc Music Corp. v. Lee*, 296 F.2d 186, 187 (2d Cir. 1961) (dismissing “the asserted ‘Inverse Ratio’ Rule” as an “ingeniously fabricated principle of law”).

42. John R. Autry, Note, *Toward a Definition of Striking Similarity in Infringement Actions for Copyrighted Musical Works*, 37 J. INTELL. PROP. L. 113, 114 (2002).

43. *Selle v. Gibb*, 741 F.2d 896, 903–04 (7th Cir. 1984) (citations omitted).

44. *See id.* at 901 (“[T]wo works may be identical in every detail, but, if the alleged infringer created the accused work independently or both works were copied from a common source in the public domain, then there is no infringement.”).

45. *Bright Tunes Music Corp. v. Harrisongs Music, Ltd.*, 420 F. Supp. 177, 180–81 (S.D.N.Y. 1976).

46. *Id.* at 180; *see also Three Boys Music Corp. v. Bolton*, 212 F.3d 477, 483 (9th Cir. 2000) (“The Isley Brothers’ access argument was based on a theory of widespread dissemination and subconscious copying. They presented evidence supporting four principal ways that [the defendant songwriters, including Michael Bolton,] could have had access to the Isley Brothers’ ‘Love is a Wonderful Thing’ . . .”). If this subconscious hullabaloo sounds a bit too out there for the judiciary, that might be because it is. Or, to borrow a phrase from Harrison’s hit-making foursome, “Turn off your mind, relax and float downstream Lay down all thought / Surrender to the

The second prong, pertaining to whether or not the defendant's copying is actionable, requires that there be a "substantial similarity" between the protected portions of the plaintiff's copyrighted work and the defendant's allegedly infringing work.⁴⁷ One legal scholar has said that this "presents one of the most difficult questions in copyright law, and one that is the least susceptible of helpful generalizations."⁴⁸ The *Arnstein* court stated:

The proper criterion . . . is not an analytic or other comparison of the respective musical compositions as they appear on paper or in the judgment of trained musicians. . . . The question, therefore, is whether defendant took from plaintiff's works so much of what is pleasing to the ears of lay listeners, who comprise the audience for whom such popular music is composed, that defendant wrongfully appropriated something which belongs to the plaintiff.⁴⁹

One commentator has suggested that a strict reading of *Arnstein* would render its "what is pleasing" language as permissive of an "extensive literal use of relatively dull material . . . so long as the 'valuable' portions were not appropriated."⁵⁰

More than three decades after *Arnstein*, the Ninth Circuit faced *Sid & Marty Krofft Television Productions, Inc. v. McDonald's Corp.*⁵¹ The court criticized *Arnstein*, arguing that "an analysis of [past] cases suggests that [supposed standards] frequently serve merely as boilerplate to copyright opinions."⁵² Instead, the court divided the copying element into two new prongs that approach the problem from a seemingly altogether different perspective consisting of two categories: an "extrinsic test" and an "intrinsic test."⁵³ The extrinsic test is a "determination of whether there is substantial similarity in ideas" and the intrinsic test is a "determination of [whether] there is substantial similarity between the forms of expression."⁵⁴ The extrinsic prong is capable of being "decided as a matter of law," and expert testimony is appropriate.⁵⁵ The *Krofft* intrinsic test seems similar to the *Arnstein* lay test, "depending on the response of the ordinary reasonable person" and inquiring "whether there is substantial similarity in the expressions of the ideas so as to constitute infringement."⁵⁶ Furthermore,

void." THE BEATLES, *Tomorrow Never Knows*, on REVOLVER (Capitol Records 1966). Indeed, perhaps only psychedelic rock music can explain this theory of "liability-by-plausibility." 4 NIMMER ON COPYRIGHT, *supra* note 32, at § 13.02[A].

47. 4 NIMMER ON COPYRIGHT, *supra* note 32, at § 13.03[A].

48. *Id.* (footnotes omitted).

49. *Arnstein v. Porter*, 154 F.2d 464, 473 (2d Cir. 1946) (footnotes omitted).

50. Aaron Keyt, Comment, *An Improved Framework for Music Plagiarism Litigation*, 76 CALIF. L. REV. 421, 440 (1988).

51. 562 F.2d 1157 (9th Cir. 1977).

52. *Id.* at 1162.

53. *Id.* at 1164.

54. *Id.*

55. *Id.*

56. *Id.*

“[b]ecause [it] is an intrinsic test, analytic dissection and expert testimony are not appropriate.”⁵⁷ Although *Krofft* dealt with the alleged infringement of a television show’s characters and substance, its altered formulation was specifically applied to music cases in *Baxter v. MCA, Inc.*⁵⁸ in 1987. There, the Ninth Circuit reversed the district court’s granting of summary judgment for composer John Williams and remanded for trial on grounds that the court could not affirmatively say that the two compositions at issue were “so dissimilar that reasonable minds could not differ as to a lack of substantial similarity between them.”⁵⁹ Thus, most believe *Krofft* was right in the sense that its extrinsic/intrinsic formulation is not all that different from the *Arnstein* formulation, at least with regard to the lay jury’s role.

In 1990, the Fourth Circuit essentially confirmed this belief, concluding that the difference in application between *Arnstein* and *Krofft* is minimal.⁶⁰ The court complicated matters further with yet another articulated formulation: “The plaintiff must establish substantial similarity of both the ideas of the two works and of the expression of those ideas.”⁶¹ Additionally, the audience test was clarified “to require that where the intended audience is significantly more specialized than the pool of lay listeners, the reaction of the intended audience would be the relevant inquiry.”⁶²

Recently, courts have been tasked with making significant determinations regarding the application of 1970s legislation (and terminology) to increasingly unforeseen sound recording methods. The practice of sampling, now over three decades old, is “a common practice in the music industry, but until recently, it was unclear whether a license was required to sample a small portion of a protected work.”⁶³ In *Bridgeport Music, Inc. v. Dimension Films*,⁶⁴ the Sixth Circuit held that there is no *de minimis* exception for sound recordings, in contrast to what courts had previously held for musical compositions.⁶⁵ The court rationalized its creation of a new rule and its likely effect with characteristic nonchalance: “Get a license or do not sample.”⁶⁶ Yet most of the cases initially filed by the *Bridgeport* plaintiffs were dismissed by other circuits. Thus, while a number

57. *Id.*

58. 812 F.2d 421 (9th Cir. 1987).

59. *Id.* at 425.

60. See *Dawson v. Hinshaw Music Inc.*, 905 F.2d 731, 732 n.1 (4th Cir. 1990) (asserting that the “difference in labels” is meaningless “because the apparent consensus as to the nature of the tests applicable to each prong . . . smooths over, as a practical matter, underlying differences in the inquiry’s two characterizations”).

61. *Id.* at 732.

62. *Id.* at 734.

63. M. Leah Somoano, Note, *Bridgeport Music, Inc. v. Dimension Films: Has Unlicensed Digital Sampling of Copyrighted Sound Recordings Come to an End?*, 21 BERKELEY TECH. L.J. 289, 289 (2006).

64. 410 F.3d 792 (6th Cir. 2005).

65. *Id.* at 798.

66. *Id.* at 801.

of academic articles have sprung up offering suggestions ranging from a compulsory licensing system⁶⁷ to outright rejection of the Sixth Circuit's decision,⁶⁸ it is unclear what form any long-term solution, should one arise, could or would take.

III. Why Is Music Worthy of Different Treatment?⁶⁹

As evidenced by the perpetual legislative alterations and persistent academic and trade-industry calls for reform, there are already many who believe that music is worthy of different treatment. In fact, in addition to the DMCA and DPRA, there have been several other smaller modifications. In 2010, the Performance Rights Act was approved by the Senate and House Judiciary Committees but never advanced to a floor vote.⁷⁰ The Act would have effectively equalized the public-performance entitlements for sound recordings and musical compositions, further blurring the distinction between the two.⁷¹ Although the legislation is outside this Note's scope, it speaks to the unrelenting controversy of musical copyrights. Even though "copyright protection was extended to music later than to other forms of expression,"⁷² its provisions continue to be tweaked. Indeed, it seems as though we are continually experimenting in an effort to find the correct balance between creativity and commerce in the face of widespread technological innovation and rapidly changing consumer habits.

It is decidedly much more difficult to find consensus over whether music is worthy of different treatment than to articulate why. Nevertheless, this part of this Note will attempt to do just that, via four separate arguments: (A) music as mathematics, (B) statutory formulation, (C) pervasive innateness, and (D) difficulties of objectification. To be fair, these considerations are interrelated, and they all primarily derive from the fact that music itself is a fundamentally different manifestation of the creative arts than the other copyrightable works. This might seem obvious, given that each recognizable art form has its own distinctive qualities. But music's

67. Joshua Crum, Comment, *The Day the (Digital) Music Died: Bridgeport, Sampling Infringement, and a Proposed Middle Ground*, 2008 BYU L. REV. 943, 947.

68. John Schietinger, Note, *Bridgeport Music, Inc. v. Dimension Films: How the Sixth Circuit Missed a Beat on Digital Music Sampling*, 55 DEPAUL L. REV. 209, 210 (2005).

69. Although the other enumerated § 102 works are outside this Note's scope, the author is not necessarily solely in favor of reform for musical copyrights. Rather, it is often the case where one aspect of a larger doctrine must undergo reform in an experiential sense in order to inform observers and scholars of the possibility of successful future reform of the doctrine as a whole. Thus, this Note does not intend to foreclose this larger prospect but does intend to explicate the reasons why the existing musical-copyright doctrine is desperately in need of reform.

70. Mary LaFrance, *U.S. Performance Rights in Sound Recordings*, BERKLEE C. MUSIC BUS. J., October 2011, at 1, 1, available at <http://www.thembj.org/2011/10/u-s-performance-rights-in-sound-recordings/>. It is believed the legislation will be reintroduced in the 112th Congress. *Id.*

71. *Id.* at 3.

72. Michael W. Carroll, *Whose Music Is It Anyway?: How We Came to View Musical Expression as a Form of Property*, 72 U. CIN. L. REV. 1405, 1414 (2004).

unique qualities warrant different treatment in the judiciary to match the different treatment music already receives in other settings.

A. *Music as Mathematics*

It is no coincidence that children who are gifted in mathematics often also excel at a musical instrument.⁷³ The two disciplines share many frames of reference. For example, a time signature, which looks like a fraction, signals the particular rhythm and tempo of a song by indicating the number of beats per measure and, relative to this determination, the length of one beat.⁷⁴ Music and mathematics are similar in their shared structure, organization, and emphasis on quantification.⁷⁵ Many aspects of music theory can be and often are expressed via some numerical marker.⁷⁶ The

73. See, e.g., V. Dion Haynes, *Music Helps L.A. Pupils Get Into Groove in Math*, CHI. TRIB., March 16, 1999, at N1 (“[A] study released Monday in the journal *Neurological Research* shows that after learning eighth notes, quarter notes, half notes and whole notes, the 2nd and 3rd graders scored 100 percent higher than their peers at the school who were taught fractions using traditional methods.”); Jodie Lynn, *Remedies May Help Solve Music, Math Problems on a Happy Note*, ST. LOUIS POST-DISPATCH, Oct. 2, 2005, available at Factiva, Doc. No. SLMO000020051003e1a20008g (“Studies have shown a definitive link between math and music. Many of the world’s greatest mathematicians have also been musicians. Studies have shown that the single strongest predictor of high math-test scores is in early exposure to music lessons. The greatest growth in brain development occurs between birth and 5 years old, but even at a later age, there will still be benefits to good music instruction.”).

74. MICHAEL MILLER, *THE COMPLETE IDIOT’S GUIDE TO MUSIC THEORY* 65 (2002). In noting the role of mathematics in measuring beats, Miller observes:

Written music uses something called a *time signature* to signify how many beats are in a measure and what kind of note is used for the basic beat. A time signature looks kind of like a fraction, with one number sitting on top of another number. The top number indicates how many beats are in a measure; the bottom number indicates the note value of the basic beat.

Id.

75. See Edward Rothstein, *Deciphering the Grammar of Mind, Music and Math*, N.Y. TIMES, June 19, 2004, at B9. In explaining the relationship between mathematics and music, Rothstein explains:

More than 20 years ago . . . , I argued that there was more to musical meaning than was evident in influential political and programmatic interpretations. For all of music’s deep connections to human experience and social life, music was also similar to mathematics in important respects, as ancient philosophical and musical texts insist.

Harmony and counterpoint, after all, are sonic reflections of ratio and number. Musical languages seem to possess their own premises and laws. And a coherent and elegantly phrased composition can display the beauty and inevitability of a mathematical proof. Mathematicians and musicians have long had reciprocal interests.

Id.

76. See, e.g., MILLER, *supra* note 74, at 6 (“An easier way to designate tones is to number each individual pitch. . . . In the Western World, we divide our scales into seven main notes—eight if you count the first note, which is repeated at the end of the scale. Because there are seven notes, it’s easy to number them—one through seven.”). Miller further illustrates the role of mathematics in note values:

Every time you sing or play a tone, you’re also singing or playing a note value. There are different types of note values, with each note value signifying a specific length of time—as measured by parts of a measure. To better explain, we have to get into a little

fundamental numerology of music is further evident in the circle of fifths, in which one seeking the fifth note in a key that proceeds to the fifth of that note and the fifth of that note and so on and so forth will eventually return back to the original note.⁷⁷ Going backwards through this cycle will lead one down the circle of fourths.⁷⁸ This is just a small illustration, but it is indicative of the many interrelated musical methods of analysis which bring to mind similar cyclical and interconnected methods of mathematical analysis. To this point, in formulating her own altered judicial musical-infringement framework, Yvette Joy Liebesman has argued that music's inherent tension between art and mathematical science makes it unique and that each aspect requires its own method of analysis.⁷⁹

Similar to mathematics,⁸⁰ most popular music is derived from clearly defined and prescribed rules and guidelines.⁸¹ For example, there are seven notes which make up a musical scale in any given key.⁸² Continuing along, the eighth note, like the first note of the scale, is the note of the key.⁸³ Thus, the first and the eighth comprise what is called an octave, or, to use well-known parlance, every scale "will bring us back to Do."⁸⁴ The key of C-major has no sharps or flats, so for ease of expression, it is often used to

math. . . . [E]ach note value lasts a specific duration, and each duration reflects a ratio to duration.

Id. at 56.

77. *See id.* at 47 ("Starting with the key of C, for every perfect fifth you move up, you add a sharp. So the key of G (a perfect fifth up from C) has one sharp. The key of D (a perfect fifth up from G) has two sharps . . . and so on." (alteration in original)).

78. *See id.* (noting that the circle of fifths has a function in reverse as well because "[f]or every perfect fifth you move down from C, you add a flat").

79. Yvette Joy Liebesman, *Using Innovative Technologies to Analyze for Similarity Between Musical Works in Copyright Infringement Disputes*, 35 AIPLA Q.J. 331, 345 (2007).

80. Alex Donnelly, *The Connections Between Music and Mathematics: Revised and Better than Ever*, SCIENCE RAY (May 25, 2008), <http://sciencerey.com/mathematics/the-connections-between-music-and-mathematics-revised-and-better-than-ever> (explicating the "elegant connection between math and music" by analyzing the way tones, tempos, patterns, and frequencies are seen to parallel one another across the two disciplines).

81. *See, e.g.*, RIKKY ROOKSBY, *HOW TO WRITE SONGS ON GUITAR* 30 (2000) ("[T]here is a simple formula for working out which chords fit together best. Let's look at it in a guitar-friendly key, C major. The notes of the C major scale are: C D E F G A B. They are separated by a set pattern of intervals: tone, tone, semitone, tone, tone, semitone (full step, full step, half-step, full step, full step, full step, half-step). In frets, the pattern is: 2 2 1 2 2 2 1. This pattern governs all major scales, regardless of the starting note. You can test this by choosing any note on any string and going up the string playing the notes in the 2 2 1 2 2 2 1 pattern. You will always get a major scale.").

82. *Id.*

83. *See, e.g.*, MILLER, *supra* note 74, at 27 ("[E]very C scale starts on C and ends on C . . ."); *id.* at 43 ("When a piece of music is based on a particular musical scale, we say that music is in the 'key' of that scale. For example, a song based around the C Major scale is in the key of C Major.").

84. *See* JULIE ANDREWS ET AL., *Do-Re-Mi, on THE SOUND OF MUSIC* (RCA 2005) (1965) ("Now children, Do Re Mi Fa So and so on are only the tools we use to build a song. Once you have these notes in your heads, you can sing a million different tunes by mixing them up.").

teach music.⁸⁵ Its notes are C, D, E, F, G, A, B, and then back again to C.⁸⁶ Certain notes of a scale played in different combinations make up a chord. For example, played simultaneously, the first, third, and fifth notes of a scale make up the major root, or I,⁸⁷ chord of that key. This pattern, by which chords in a key are formed, is indicative of an organizational structure existing in both music and mathematics, which permits a formula to be applied successfully in varying circumstances.⁸⁸ The rule-based nature of music renders it necessarily unlike the other artistic works, for which boundaries and guidelines beyond the purely logistical are virtually nonexistent.

A chord progression, where chords are played in succession with repetition,⁸⁹ is perhaps the most easily cognizable building block of popular music.⁹⁰ Although songs are rarely just comprised of chords, they can almost always be condensed into, and recreated via, these progressions.⁹¹ The chord progression to one of the most well-known popular music songs in history, Bob Marley's "No Woman, No Cry," is C → A-minor → F → G.⁹² In the key of C, this translates to I (the notes C, E, and G played simultaneously) → V (the sixth, first (root), and third notes of the scale: A, C, and E played simultaneously) → VI (the fourth, sixth, and first notes of the scale: F, A, and C played simultaneously) → IV (the fifth, seventh, and second notes of the scale: G, B, and D played simultaneously).⁹³ Similarly, this is also the chord progression for a popular pop-punk song by Blink-182 entitled

85. See MILLER, *supra* note 74, at 29 (beginning instruction of scales with C Major, noting its lack of sharps and flats); ROOKSBY, *supra* note 81, at 33 ("C major [has] no sharps or flats.").

86. ROOKSBY, *supra* note 81, at 30.

87. Roman numerals are commonly used in musical theory in place of Arabic numerals. See generally *id.* at 31–32.

88. See, e.g., *id.* at 31 ("There is a handy way of notating and thinking about chord progressions without specifying the key. This is possible because the *internal* chord relationships of all major keys are identical. The same is true of all minor keys as long as they are based on one type of minor scale. The system has been used for hundreds of years in classical music.").

89. It is important to note that chords can be and are used, even in popular music, that have notes outside of the scale of the key in which a song is written, but this oversimplification should suffice for our purposes. See *id.* at 30 ("Sometimes a melody or a riff will suggest an unusual chord change, and you may later discover that you have included a chord that is not strictly 'in key.' If it works in the context of the song, that's fine. The history of music is littered with broken rules.").

90. Autry, *supra* note 42, at 134 n.133 ("Chord progressions are the connections from one chord to the next, with an eye toward the cadence (or end) of the progression and how the harmony changes along the path.").

91. Cf. ROOKSBY, *supra* note 81, at 13 ("Harmony is supplied by chords One thing that makes the guitar popular as an accompaniment instrument is that it is relatively easy to learn some chord shapes, strum them and sing. Many professional performers have got by for years with little more than this.").

92. BOB MARLEY & THE WAILERS, *No Woman, No Cry*, on NATTY DREAD (Island Records 2001) (1974).

93. ROOKSBY, *supra* note 81, at 15, 57; see also *id.* at 32 ("[W]e can write a chord sequence purely in these Roman numerals: I VI IV V. Regardless of the major key in which we play this it will always comprise a major chord followed by a minor and then two majors. Their relationship to each other is constant whatever the key.").

“Dammit.”⁹⁴ Furthermore, both songs contain a refrain built around the notion that “everything is going to be all right.” Whether or not infringement exists—were a hypothetical suit brought and a jury trial commenced—would be determined by lay jurors who lack this basic contextual understanding of these songs’ musical structures.

Chord progressions demonstrate the extent to which music’s mathematical origins cause unique problems for its dual copyrights. Of greater significance, however, the proffered example only scratches the surface of the immense degree to which music as an art form is capable of academic comparison from an impartial, purely fact-based perspective. Chords merely exemplify the ways in which music’s intrinsically precise foundation enables us to compare and contrast certain objective elements of songs. Professionals can say with mathematical precision that beats per minute, tones of melody (pitch of note patterns), and rhythms of melody (length of note patterns), for example, are identical or divergent and why.⁹⁵ Professionals are also aided by digital-audio-wave software, which can analyze and compare songs with computerized precision.⁹⁶ These assessments are not opinions. Rather, they are objectively correct. And yet they likely would not and could not impact a purely lay comparison. By contrast, there is nothing mathematically predetermined about any of the other artistic copyrightable works.

The ability to quantify similarity is so fundamental to the art form of music that to analyze without, at the very least, contemplating its significance, when most professionals would begin their assessment from such a perspective, would be tantamount to ignoring the extralegal, oft policy-based arguments frequently presented in *amici curiae* briefs as though

94. BLINK-182, *Dammit*, on DUDE RANCH (MCA Records 1997); see also PHISH, *Farmhouse*, on FARMHOUSE (Elektra Records 2000) (utilizing the same precise chord progression and lyrical motif); QUIET COMPANY, *Never Tell Me the Odds (This Is the Worst Crazy Sect I’ve Ever Been In)*, on WE ARE ALL WHERE WE BELONG (Quiet Company 2011) (expressing simultaneous confidence and doubt in the author’s newly atheistic leanings through a subtle yet dramatic alteration to the lyric, “Everybody’s probably gonna be all right”).

95. See, e.g., Walker, *supra* note 10, at 50 (“Some elements that these musicologists (who, really, are musicians with day jobs) codify are technical, like beats per minute, or the presence of parallel octaves or block chords. Someone taking apart Gnarl Barkley’s ‘Crazy’ documents the prevalence of harmony, chordal patterning, swung 16ths and the like.”); *supra* text accompanying note 10.

96. Avery Li-Chun Wang, *An Industrial-Strength Audio Search Algorithm*, INT’L SYMPOSIUM ON MUSIC INFO. RETRIEVAL (Oct. 2003), available at <http://www.ee.columbia.edu/~dpwe/papers/Wang03-shazam.pdf>; see also, e.g., Lee Bergquist, *Phone App Identifies Birds’ Tunes*, MILWAUKEE WIS. J. SENTINEL (Oct. 12, 2011), <http://www.jsonline.com/sports/outdoors/phone-app-identifies-birds-tunes-131686693.html> (“In many cases, the software is good enough to discern among songs of different individuals of the same species.”); Farhad Manjoo, *That Tune, Named: How Does the Music-Identifying App Shazam Work Its Magic?*, SLATE (Oct. 19, 2009), http://www.slate.com/articles/technology/technology/2009/10/that_tune_named.html (“Shazam creates a spectrogram for each song in its database—a graph that plots three dimensions of music: frequency vs. amplitude vs. time. The algorithm then picks out just those points that represent the peaks of the graph—notes that contain ‘higher energy content’ than all the other notes around it . . .”).

legal realism never had its impact.⁹⁷ But present judicial treatment most frequently does just this, ignoring professional objectivity in favor of the calculated unawareness of a blissfully ignorant hypothesized lay ideal. Music and mathematics are intrinsically linked in a manner that complicates musical copyrights, fundamentally distinguishes music from the other artistic works, and commands exceptional treatment.

B. *Statutory Formulation*

Perhaps the most obvious reason why music is worthy of different treatment is implicitly codified in the statute via the inclusion of two separate types of works in the form of sound recordings and musical compositions.⁹⁸ These two works are treated as entirely separate in the United States Copyright Act and often, as *Bridgeport* demonstrated, in the judiciary.⁹⁹ This separation belies a complexity that sets the musical copyrights apart from the other copyrightable forms of art. The duality of musical copyrights demonstrates the extent to which music as an art form truly encompasses the creation of two separate works. A particular song is simultaneously the musical composition of the songwriter(s), while its fixation in an audio medium can be the sound recording of any combination of the performer(s), producer(s), and engineer(s).¹⁰⁰ This duality is even more significant because of the law's indistinguishable treatment of a song's music and its lyrics. Lyrics are part of the musical composition.¹⁰¹ Thus, a songwriter's lyrics are protected from infringement by the musical-composition copyright. But also, an artist's specific vocal performance is protected from infringement by the sound-recording copyright.¹⁰² This separation of elements could only conceivably be compared to the audio of a film, which is encompassed in the motion picture's copyright.¹⁰³ Yet few would argue that that audio is itself

97. See generally, e.g., *Function and Role of Amicus Briefs in Public Health Litigation*, PUBLIC HEALTH LAW CENTER, <http://publichealthlawcenter.org/sites/default/files/resources/phlc-fs-amicus.pdf> (discussing the focus on policy that is prevalent in amicus briefs due to amici's indirect relationship to the case at hand and the recognized value of these policy-based perspectives).

98. 17 U.S.C. § 102(a)(2), (7) (2006).

99. See *supra* notes 63–68 and accompanying text.

100. See 1 NIMMER ON COPYRIGHT, *supra* note 32, § 2.10[A][2] (“A sound recording must be distinguished from . . . the underlying musical composition . . . that is recorded and transposed into aural form by the sound recording.”).

101. See 2 NIMMER ON COPYRIGHT, *supra* note 32, § 30.02[A] (referring to the music and lyrics as the “two elements” of a musical composition).

102. See U.S. GOV'T ACCOUNTABILITY OFFICE, GAO-10-826, TELECOMMUNICATIONS: THE PROPOSED PERFORMANCE RIGHTS ACT WOULD RESULT IN ADDITIONAL COSTS FOR BROADCAST RADIO STATIONS AND ADDITIONAL REVENUE FOR RECORD COMPANIES, MUSICIANS, AND PERFORMERS 5 (2010) (“[T]he song, ‘I Will Always Love You,’ was part of the soundtrack for the movie, *The Bodyguard*, in 1992. The copyright holder of the musical work is the songwriter, Dolly Parton, who owns both the words and music. However, the copyright holder of the sound recording, as performed by Whitney Houston, is the record company, Sony Music, to whom the soundtrack is registered.”).

103. 1 NIMMER ON COPYRIGHT, *supra* note 32, at § 2.09[E][1].

an art form in the same way as song lyrics, which are frequently favorably compared to and accepted as poetry. As indicated by the provision of two different copyrights, popular music is structurally different from the types of art protected elsewhere in the Copyright Act. The musical copyrights generate immense revenue for their authors from a diverse array of potential projects including but not limited to: licensing for use in a film or for performance onstage, radio airplay, sampling within another song, and inclusion in a sheet-music book or a lyric passage in a book of poetry or quotations.¹⁰⁴ This list of nonexhaustive uses demonstrates why music's statutory formulation exists in its present dichotomy and why it necessitates, by virtue of its complexity, different treatment in the judiciary.

Another noteworthy statutory variation between music and other works is evidenced by the United States Copyright Act's regulations for interactive music services like Pandora.¹⁰⁵ A full discussion is outside this Note's scope, but it is enough to note that there are several limitations on how such services may operate, including restrictions on the number of times songs by an artist or from an album can be played within a given time frame.¹⁰⁶ This type of regulation is revelatory because it demonstrates the extent to which persons are often drawn to and enjoy most the music they already know. This is not the case with most of the other copyrightable works, for which the newness of the experience is usually prized.¹⁰⁷ "I've already heard it" is not an oft-heard excuse with music the way "I've already seen that" might be with the moving image or other fine art. Rather, music's allure is quite the opposite, indicating a different human approach to music as an art form.

Lastly, another statutory peculiarity with music is the provision concerning compulsory licenses for "nondramatic musical works."¹⁰⁸ Although inapplicable to sound recordings, the Copyright Act enables musicians to record cover versions of another artist's musical composition,

104. See Robert R. Carter, Jr., *Sources of Publishing Income*, MUSE'S MUSE, <http://www.musesmuse.com/robcarter-art.html> (explicating the four main sources of revenue for musical-copyright holders: mechanicals, public performance, synchronization, and print).

105. See 17 U.S.C. § 114(j)(7) (2006) (including the features of Pandora in the definition of *interactive service*, thus allowing for the regulation of the popular music provider).

106. See *id.* § 114(j)(13) ("The 'sound recording performance complement' is the transmission during any 3-hour period . . . of no more than—(A) 3 different selections of sound recordings from any one phonorecord lawfully distributed for public performance or sale in the United States, if no more than 2 such selections are transmitted consecutively; or (B) 4 different selections of sound recordings—(i) by the same featured recording artist; or (ii) from any set or compilation of phonorecords lawfully distributed together as a unit for public performance or sale in the United States . . .").

107. See, e.g., Leslie Bennetts, *One Day in the Video Rental Whirl*, N.Y. TIMES, Aug. 13, 1986, at C17 ("I call it VCR fever—you want to watch everything you've never seen before."); Neal Justin, *If You Haven't Seen It, It's New to You*, STAR TRIB. (Minneapolis–St. Paul), June 26, 2011, available at Factiva, Doc. No. MSP000020110629e76q00001 (suggesting that viewers "[s]ick of watching reruns" during the summer should watch television shows that they have not already seen).

108. 17 U.S.C. § 115(c)(3)(A) (2006).

provided that the new recording does not “change the basic melody or fundamental character of the work.”¹⁰⁹ Music is the only copyrightable work of art where covers are not only an entirely acceptable, credible aspect of the art form¹¹⁰ but are also so much a part of how the musical-composition copyright functions that so long as the work has been published in the United States, a songwriter is virtually without recourse to prevent them.¹¹¹ Granted, the covered artist receives a statutorily prescribed royalty,¹¹² but it is significant that “copying” a previously created and copyrighted work without changing its “fundamental character” would be blatant infringement in nearly all of the other copyrightable art forms.¹¹³ Yet, with music, the art form is so cognizant of and comfortable with the repetition of form and substance that existing doctrine literally seems to encourage legal copying.

The underlying reasons for why music is treated specially statutorily, as evidenced by the complex musical-copyright framework, are equally applicable to why alleged musical infringement should be treated specially from a judicial standpoint. Music is different enough from the other works that it is treated differently in a multitude of ways, yet its infringement analysis has remained identical to that of the other works. If music is explicitly treated dissimilarly within the statute, why would we not extend that same special treatment to the judicial context? The fact that there are two creative copyrightable works for music demonstrates that one was not enough to sufficiently protect the art form. Similarly, the existing infringement regime is not enough to account for the ways in which music exists differently in the artistic sphere of the human brain. The rationales behind the unique provisions in the Copyright Act for interactive music and cover

109. *Id.* § 115(a)(2).

110. *But see* Sandie Angulo Chen, *The 25 Best Movie Remakes of All Time*, MOVIEFONE (Apr. 15, 2010, 2:00 PM), <http://blog.moviefone.com/2010/04/15/best-movie-remakes/> (“Let’s face it, Hollywood is full of unoriginal ideas; just look at the disproportionate number of films based on books, plays, TV shows and even Broadway musicals that were *themselves* based on movies. But every now and then, that most unoriginal of green-lit ideas, the remake of a popular or critically acclaimed movie—like current releases ‘Clash of the Titans,’ ‘Death at a Funeral’ and the upcoming ‘Nightmare on Elm Street’ reload—can yield a film that equals, or in rare cases, surpasses the original.”).

111. *See* 2 NIMMER ON COPYRIGHT, *supra* note 32, § 8.04[C] (“[T]he Section 115 compulsory license is not triggered until a phonorecord has been publicly distributed under the authority of the copyright owner.”). The compulsory license derives from a vastly different era of public consumption of musical compositions and was intended to enable a song’s usage on a phonograph record or piano roll. *Id.* § 8.04[A]. The 1909 Congress that enacted the Copyright Act that gave rise to this “compulsory license” expressed concern that complete exclusivity in musical compositions would give rise to “a great music monopoly.” H.R. REP. NO. 60-2222, at 6 (1909).

112. 2 NIMMER ON COPYRIGHT, *supra* note 32, § 8.04[H][1].

113. *See id.* § 8.01[A] (“[T]he rights of a copyright owner are not infringed if a subsequent work, although substantially similar, has been independently created without reference to the prior work.”); *id.* § 8.04[E] (“The compulsory license permits the licensee to assemble his own musicians, singers, sound recording engineers, and equipment for the purpose of making a sound recording of the musical work that is the subject of the license.”). *But see id.* §§ 8.03, .05–.08 (articulating several limitations on the reproduction right).

songs reveal how lawmakers have tried to legislate the altered approach to music taken by consumers, content providers, and future artists. Surely, the judiciary should be able to take these same widely acknowledged differences and effect change, creating a unique infringement procedure adapted to the inimitable qualities of the musical copyrights.

C. *Pervasive Innateness*

No less a legal authority than the great Oliver Wendell Holmes once said, “Take a music-bath once or twice a week for a few seasons, and you will find that it is to the soul what the water-bath is to the body.”¹¹⁴ Holmes’s belief in music as medicine for the human psyche touches on an important distinction between music and the other copyrightable works. Law Professor J. Michael Keyes remarked that “[a]lthough it is readily admitted by those who study music that we still know relatively little about why music evokes such [intense physiological and emotional] effects within us, there is no question that music enjoys a unique place among artistic endeavors and the human experience associated therewith.”¹¹⁵ Certainly, music is not alone in its length of existence, as scientists continue to unveil impressive specimens demonstrating originality and imagination across all of antiquity, most of which would be at least theoretically copyrightable and would cover nearly every one of the presently enumerated works.¹¹⁶ Yet there is something about human interaction with music that is innate. Citing to a popular study by Don Campbell, Keyes noted that “[e]ven before birth, humans respond to music. Shortly thereafter, music moves babies to relax, clap, sway with the beat, and even sing No other artistic stimulus enjoys a response of this nature from such a young and uninitiated group.”¹¹⁷ If, even in the absence of conscious apprehension of its impact, music is capable of such influence, surely its sway over humans is distinct from that of the other copyrightable works. Humanity has perceived music and undoubtedly has regarded it differently than the other art forms since ancient times.¹¹⁸

The instinctive nature of the human connection to music is further amplified by, or perhaps is a cause of (or both), the degree to which music pervades virtually every aspect of modern life. From the moment we take our first breath, we are inundated by music in advertising and popular

114. OLIVER WENDELL HOLMES, *OVER THE TEACUPS* 97 (Boston, Houghton, Mifflin & Co. 1891).

115. J. Michael Keyes, *Musical Musings: The Case for Rethinking Music Copyright Protection*, 10 MICH. TELECOMM. & TECH. L. REV. 407, 423 (2004) (footnote omitted).

116. See, e.g., Stephen Mithen, *Introduction: The Archeological Study of Human Creativity*, in CREATIVITY IN HUMAN EVOLUTION AND PREHISTORY 1, 1 (Stephen Mithen ed., 1998) (marveling at the bounds of human ingenuity evidenced by “new discoveries, such as the 30,000-year-old paintings from Chauvet cave, or the 400,000-year-old wooden javelins from Schöningen” (citations omitted)).

117. Keyes, *supra* note 115, at 421 (footnotes omitted).

118. *Id.* at 421 n.76.

culture. Certainly this is the case with some of the other copyrightable works, but music's effect is different. Consider how many times a song has gotten "stuck in [your] head."¹¹⁹ This phenomenon has no parallel in any of the other copyrightable works.¹²⁰ To date, no American copyright case outside of the musical context has found evidence of solely copying via an author's subconscious.¹²¹ If courts have determined it to be possible that humans can intuitively plagiarize music to the exclusion of the other works of art, surely there is something special about music. Although there is much variance in the individual significance that persons assign to music, there can be no doubt that it is universally omnipresent. Though fine art and film are also, to some extent, pervasive aspects of modern life, as Keyes noted, their effect is not nearly as great and human appreciation for them is not innate and thus is necessarily acquired.¹²² Humanity's seemingly eternal bond with music and its relentless presence in both the human brain and daily life render it inherently different from the other copyrightable works, warranting special treatment by the judiciary.

D. *The Difficulty of Objectification*

Regardless of the mathematical analysis, music is also capable of objectification in other manners. But its objectification is complicated by the unique interaction between music and the human brain. Lay objectification is fundamentally overwhelmed by personal subjectivity. Thus, musical copyrights merit different treatment in the context of judicial infringement analysis by virtue of their divergence from the other copyrightable works.

Stephanie Jones has criticized the musical-copyright doctrine, arguing that music has "inherently distinctive features which dictate a different inquiry."¹²³ She believes that "a musical idea consists of more than only melody, harmony and rhythm," but rather, many factors, including "tone, pitch, tempo, spatial organization, consonance, dissonance, phrasing, accents, note choice, combinations, interplay of instruments, [and] bass lines," are

119. See, e.g., *WAYNE'S WORLD* (Paramount Pictures 1992) (making light of the commonality of the subconscious musical connection with Toni Basil's "Hey Mickey"); see also Rihit Chatterjee, *Earworms: Why Songs Get Stuck in Our Heads*, BBC NEWS (Mar. 6, 2012) <http://www.bbc.co.uk/news/magazine-17105759> (analyzing the factors that can lead to the phenomenon that music psychologist Dr. Vicky Williamson has dubbed "earworms").

120. Perhaps video games are comparable in this regard, but it is simply not the same thing for individuals to see Tetris cubes or Snood characters in their heads as it is for them to continually whistle, hum, or sing along with a song they have heard or perhaps merely thought about a significant time earlier.

121. See *supra* notes 45–46 and accompanying text.

122. See Keyes, *supra* note 115, at 422–23 ("[Music] evokes a vast array of emotional responses from the recipients of it. It inspires, consoles, motivates, awakens, and energizes us unlike other artistic endeavors.").

123. Stephanie J. Jones, *Music Copyright in Theory and Practice: An Improved Approach for Determining Substantial Similarity*, 31 DUQ. L. REV. 277, 278 (1993).

determinative.¹²⁴ From the standpoint of classification, music fans and critics are constantly arriving at new methods of description. These descriptions are usually a series of words. But words are just words.¹²⁵

The descriptions are not necessarily—as Liebesman proposed in her Mega-element Analysis (MEA)—related to “further subdivi[sion]” of the “broad, artistic categories of harmony, melody, structure, and rhythm.”¹²⁶ Rather, they often pertain to emotion, even in the absence of lyrics conveying so explicitly. As Richard Posner has said, “‘interpretation’ is a chameleon.”¹²⁷ Justice Scalia has cautioned that, inevitably, “what judges believe Congress ‘meant’” will become “whatever judges think Congress *must* have meant, *i.e.*, should have meant.”¹²⁸ Thus, necessarily, one interpreting—or, for our sake, analyzing—a musical composition or sound recording for similarity will doubtless impart one’s own preconceptions into the process. Liebesman even acknowledges as much: “A problem with the MEA is the use of trained musicians whose analyses contain some level of subjectivity.”¹²⁹ Famed novelist Leo Tolstoy once said, “Music is the shorthand of emotion.”¹³⁰ Certainly, broad categorization is possible across all the works of art, but the collective sentimentalism of music makes this process exceptionally difficult for laypersons with respect to music.¹³¹ What is distinctive about music in this regard is its personal connection in the human brain with a particular time, person, or life event.¹³² Though the notion is tough to articulate without inevitably veering towards syrupy cliché, music is indeed the “soundtrack of our lives.”¹³³ We are equally impacted by songs, or rather, genres, which do not appear on these personal soundtracks.¹³⁴ Thus, laypersons hearing a song from a completely detached, uninitiated perspective will find it difficult to overcome their predisposed

124. *Id.* at 294–95.

125. *Cf.* Church of the Holy Trinity v. United States, 143 U.S. 457, 465 (1892) (stressing, albeit not without controversy in its time and still today, the importance of purpose above literal, ordinary meaning: “no purpose of action against religion can be imputed to any legislation, state or national, because this is a religious people”).

126. Liebesman, *supra* note 79, at 345.

127. RICHARD A. POSNER, THE PROBLEMS OF JURISPRUDENCE 271 (1990).

128. Zuni Pub. Sch. Dist. No. 89 v. Dep’t of Educ., 550 U.S. 81, 117 (2007) (Scalia, J., dissenting).

129. Liebesman, *supra* note 79, at 349.

130. ROBERT I. FITZHENRY, THE HARPER BOOK OF QUOTATIONS 314 (3d ed. 1993).

131. *See infra* note 135.

132. *See, e.g.*, NICK HORNBY, HIGH FIDELITY 45 (1995) (describing The Beatles as “bubblegum cards and *Help* at the Saturday morning cinema and toy plastic guitars and singing ‘Yellow Submarine’ at the top of my voice in the back row of the coach on school trips”).

133. Chris Carlson, *Dick Clark Makes Rare Appearance at Emmys*, MSNBC (Aug. 28, 2006), <http://www.msnbc.msn.com/id/14548386/ns/entertainment-television/>.

134. *See, e.g.*, Michaeleen Doucleff, *Anatomy of a Tear-Jerker*, WSJ.COM (Feb. 11, 2012), <http://online.wsj.com/article/SB10001424052970203646004577213010291701378.html> (explaining musicologically why Adele’s “Someone Like You” is both the “perfect tear-jerker” and evidence of “a formula for commercial success”).

sentiments and tastes in attempting to remain objective. To prove the point, ask an elderly person to listen to music to which they are not accustomed—i.e., made by and for a younger generation—and see how long it takes for them to either suggest it all sounds the same or begin to begrudgingly compare it to that which they prefer—i.e., “it has no melody.” This degree of taste influence is unparalleled in scope in the creative arts. Lay music objectification is plagued by implicit subjectivity¹³⁵ in a manner unique to music as compared to the other enumerated works.

It is a relatively frequent occurrence when listeners to a song will inquire, often rhetorically, “what does this sound like?” It is rarely the case when people are at a loss for songs or artists to serve as a comparison. Rather, different people will inevitably offer up different answers. Some would ascribe this to the delusional déjà vu invocations of a select group of disinterested, uninitiated lay listeners. This dismissal might hold sway but for the fact that a lack of comparison would seem more likely to result than an abundance. Indeed, every layperson hears music through a personal prism of interpretation. We are so opinionated about music and what it means to us that our individual musical coherence is inherently shaped by past experience and understanding.¹³⁶ This is part of what makes music and the human connection to it so great, but it also makes laypersons far less equipped to pass judgment on the alleged similarities of songs.¹³⁷ This quality is not entirely unique to music, but the degree to which it manifests itself in the greater population is decidedly not present in any of the other copyrightable works and necessitates different treatment of the musical copyrights.

Viewing music in the absence of its larger context inhibits full understanding in a manner not accurately demonstrative of a work’s scope, significance, and, of particular importance for an infringement analysis, originality and creativity. Music inherently means different things to different people.¹³⁸ These differences are extremely difficult for laypersons to

135. For example, Bob Dylan wrote the song “Only a Pawn in Their Game” about the tragic death of civil rights leader Medger Evers. BOB DYLAN, *Only a Pawn in Their Game*, on THE TIMES THEY ARE A-CHANGIN’ (Columbia Records 1964). According to Dylan, he never intended to aver some grandiose, broadly sweeping “We Shall Overcome”-like sentiment, but rather, more simply, to tell Evers’s story in song. Curiously, the audience saw in Dylan something larger, unceremoniously anointing him the “voice of a generation” as well as a “protest singer” and “activist.” See HOWARD SOUNES, *DOWN THE HIGHWAY: THE LIFE OF BOB DYLAN* 261 (2001) (noting “the misconception during the folk era that Bob [Dylan] was at heart a radical, a protest singer, a ‘spokesman of his generation.’”).

136. See, e.g., JOHN POWELL, *HOW MUSIC WORKS: THE SCIENCE AND PSYCHOLOGY OF BEAUTIFUL SOUNDS, FROM BEETHOVEN TO THE BEATLES AND BEYOND* 1 (2010) (“[S]ometimes we can be familiar with something we really enjoy, but have no idea what it actually is. This is the relationship most of us have with music—pleasure without understanding.”).

137. See *infra* note 140 and accompanying text.

138. See, e.g., Bob Boilen, *What Does Music Mean to You?*, ALL SONGS CONSIDERED: THE BLOG (July 22, 2010, 2:02 PM), <http://www.npr.org/blogs/allsongs/2010/07/22/128690717/what-does-music-mean-to-you> (soliciting opinions about what music means to his readers as both listeners and players of music).

ignore in their enjoyment and, for our purposes, legal analysis of a work. Much of what endears music to a person's ears and foments a personal connection—i.e., *Arnstein's* “what is pleasing” language—is entirely inarticulable to and impossible for a layperson to neutrally evaluate.¹³⁹ By contrast, professionals who understand the larger body of recorded popular music are better able to view songs objectively and may therefore be less susceptible to the subjective whims of humanity's emotional connection or lack thereof.¹⁴⁰ In none of the other enumerated works is the disparity between professional and lay critical capabilities so vast. Music is different because of its status as an art form with which nearly all laypersons have difficulty remaining objective. Subjectivity plagues the infringement analysis and leads to undesirable verdicts, diminished creativity, and unnecessary licensing brought on by fear of litigation. Music is worthy of different treatment in the judiciary to bring the courts in line with humanity's already altered approach to music.

E. Refuting the Objections

Some would argue that just because particular expertise is required to evaluate music does not make it any different than the numerous other juror-expertise problems in other areas of the law.¹⁴¹ But the juror-expertise issue discussed herein refers to instances when juries think they understand the law.¹⁴² Copyright infringement derives from the simple notion of plagiarism with which every American is surely familiar. Across all of the creative arts, there is some potential for juror-expertise problems in this respect, but with music, where nearly everyone considers themselves an expert, this problem is amplified greatly. The mathematically precise nature of music means that there are fundamental facts about two works which a layperson will not adequately grasp.¹⁴³ Thus, the juror-expertise problem is enlarged even further because persons think they know what they hear when a professional analysis would elucidate their actual lack of context.

139. See, e.g., *Consonance and Dissonance*, GA. ST. U., <http://hyperphysics.phy-astr.gsu.edu/hbase/music/mussca.html> (“You can get into such intense debate about what is ‘pleasing’ that some have come to define music as ‘sounds organized by human beings’ to accede the endless variety.”).

140. See ANTHONY STORR, *MUSIC AND THE MIND* xii (1992) (“It is true that those who have studied the techniques of musical composition can more thoroughly appreciate the structure of a musical work than those who have not.”).

141. See, e.g., *Parke-Davis & Co. v. H.K. Mulford Co.*, 189 F. 95, 115 (S.D.N.Y. 1911) (noting, in an intellectual property case, that more expert guidance is necessary for the fair administration of justice and musing, “How long we shall continue to blunder along without the aid of unpartisan and authoritative scientific assistance in the administration of justice, no one knows; but all fair persons not conventionalized by provincial legal habits of mind ought, I should think, unite to effect some such advance”), *aff'd in part, rev'd in part*, 196 F. 496 (2d Cir. 1912).

142. See, e.g., *Kidd & Coch, Patents and Jury Trials*, 2 J. PROPRIETARY RTS. 17 (1990) (“Jurors do not evaluate cases on the facts alone, but assess facts based upon their own attitudes, values, prejudices, and emotions.”).

143. See *supra* subpart III(A).

Others might suggest, as the author has acknowledged throughout this part, that arguments for differentiation of music as an art form are null in that they would apply to all works of "art" which fall under the heading of copyrightable subject matter. Fundamentally, however, it is an extremely rare occasion when the issues of infringement in the other works of art hinge on such technical, precise, and mathematically certain qualities.¹⁴⁴ Music is far more pervasive and hence ingrained in the American subconscious than the other works.¹⁴⁵ Psychologically, music is different than the other works covered by existing law.¹⁴⁶ Social philosopher Theodore Adorno articulated the distinction between music and the other arts as mystifying: "It is at once completely enigmatic and totally evident. It cannot be solved, only its form can be deciphered."¹⁴⁷ It is not surprising then that confusion is so rampant in musical-copyright law. Doctrinal confusion leads to litigation, of which there has been much.¹⁴⁸ And yet, as Debra Presti Brent observed, "[A] musical infringement case is judged by the same principles that are applied in suits involving other copyrightable works including plays, novels, sculptures, maps, television programs, directories, photographs, and paintings. Although courts have recognized this problem, they have been reluctant to change their approach."¹⁴⁹

F. Conclusion

As the preceding subparts make clear, music is already treated differently than the other enumerated works by the consumer public, by its creators, and in the law. Additionally, music's fundamental numerology renders musical copyrights capable of complex, precise analysis which many laypersons simply cannot or will not grasp. It is time for the courts to fall in line with the differing treatment extended to music by academics, legislators, and laypersons alike. As the failures of the present system of musical-copyright infringement demonstrate, the time for reform is now.

144. See *supra* notes 74-81 and accompanying text.

145. See STORR, *supra* note 140, at 45 ("Music is so freely available today that we take it for granted and may underestimate its power . . ."); Vladimir J. Konečni, *Social Interaction and Musical Preference*, in *THE PSYCHOLOGY OF MUSIC* 497, 499 (Diana Deutsch ed., 1982) (commenting on the "penetration of music into every corner of people's lives, literally and metaphorically"); see also *id.* at 498 ("[M]usic appreciation has been radically altered by the technological and social changes in the twentieth century.").

146. Cf. T.W. ADORNO, *AESTHETIC THEORY* 178 (Gretel Adorno & Rolf Tiedemann eds., C. Lenhardt trans., 1970) (holding music apart from other arts as the "prototypical example" of what makes art cognitively difficult to understand).

147. *Id.*

148. See *infra* note 196 and accompanying text.

149. Debra Presti Brent, *The Successful Musical Copyright Infringement Suit: The Impossible Dream*, 7 U. MIAMI ENT. & SPORTS L. REV. 229, 229-30 (1990); see also Franklin Mint Corp. v. Nat'l Wildlife Art Exch., Inc., 575 F.2d 62, 65 (3d Cir. 1978) (recognizing this problem as "troublesome").

IV. Failures of the Present System

In the academic literature, there is no shortage of commentators who dispute the merits of the present system of musical-copyright infringement. The *Arnstein* formulation has been criticized as “failing to consider the effect of changes the defendant may have made in the plaintiff’s material” and hence leaving open the “question of how much copying is enough to incur liability.”¹⁵⁰ Additionally, one commentator said of the *Krofft/Baxter* framework, “In music there is no ‘idea’ or ‘expression’ to be distinguished. The conceptual framework simply has no relationship to the type of artwork involved.”¹⁵¹ The problems are extensive, but above all else, the biggest issues arise from the overriding confusion caused by the present system’s numerous indeterminacies.

One major problem is the bias shown against popular music and similar nonclassical musical works. A California judge in 1942 differentiated “[a] phrase from Beethoven, or from any other great composer” from “the trite phrasing of an ordinary popular song, with its limitations.”¹⁵² Certainly, this negative predisposition is less prevalent in 2010, when many of our nation’s most esteemed judges have grown up exposed to and enjoying rock ‘n’ roll.¹⁵³ Nevertheless, it is a virtual certainty that each successive generation will develop its own new form of popular music¹⁵⁴ to which the existing judicial establishment will be unaccustomed and therefore unfit to make judgments. This is evidenced by the judicial mistreatment of infringement cases involving more recent genres of music.¹⁵⁵ Presumably, a jury panel, by

150. Keyt, *supra* note 50, at 441.

151. *Id.* at 443.

152. *Carew v. R.K.O. Radio Pictures, Inc.*, 43 F. Supp. 199, 200 (S.D. Cal. 1942).

153. *See, e.g., Arista Records, LLC v. Launch Media, Inc.*, 578 F.3d 148, 160 (2d Cir. 2009) (repeatedly using The Beatles as an illustrative example in the opinion); *Wachovia Bank, N.A. v. United States*, 455 F.3d 1261, 1262 (11th Cir. 2006) (quoting The Beatles’ “Taxman”); *Gorbach v. Reno*, 219 F.3d 1087, 1095 (9th Cir. 2000) (quoting The Beatles’ “All You Need is Love”).

154. *See* Robin Hilton, *The Sound of a Generation*, ALL SONGS CONSIDERED: THE BLOG (June 6, 1998, 9:46 AM), http://www.npr.org/blogs/allsongs/2008/06/the_sound_of_a_generation.html (“Every generation has its own soundtrack. The Silent Generation (people born in the ’20s and ’30s) had big band and swing. Baby Boomers (born in the ’40s and ’50s) had rock and soul. Generation X (born in the ’60s and ’70s) had grunge and hip-hop.”).

155. *See* *Bridgeport Music, Inc. v. Dimension Films LLC*, 230 F. Supp. 2d 830, 841–42 (M.D. Tenn. 2002) (comparing two compositions and arguing that while one “is a celebratory song . . . essentially about dancing,” the other involves a man’s “anger, anxiety and fatalism” while evading the police, and concluding that “a balance must be struck between protecting an artist’s interests, and depriving other artists of the building blocks of future works”), *aff’d in part and rev’d in part*, 383 F.3d 390 (6th Cir. 2004). Clearly, there are few who would disagree that the Master P song and the George Clinton song at issue in *Bridgeport* had different subjects, but to use these subjects dispositively belies artistic ignorance. It is unlikely that any true fan or consumer of hip-hop/rap music would dispute the fact that people will dance to rap songs with all sorts of non-celebratory lyrical content and that few rap songs are not in some way “danceable.” Judge Higgins, it can be assumed, is not a fan of the music at issue (as one might imagine from a man who graduated law school in 1957), nor does he understand its appeal in the broadest sense of the term. *See* Judge Thomas Aquinas Higgins, U.S. DISTRICT CT. MIDDLE DISTRICT TENN., <http://>

virtue of its process of constitution, will, on average, be more “hip” than judges,¹⁵⁶ but still, the problem exists and persists to some extent in the jury room as well.

Another major problem with the present copyright system is the strategic element of venue selection. Michael Landau and Donald E. Biederman have pointed out that “[t]here is, or has been, disagreement regarding almost every possible doctrine covered under federal copyright law from threshold issues of originality and ownership to infringement and damages.”¹⁵⁷ Further, “[m]any of the differences—often polar disagreements—still remain unresolved and will inevitably lead to forum shopping. The outcome of a case should not depend upon the jurisdiction in which the suit is initiated; it should be a function of the law.”¹⁵⁸ Additionally, the increasingly national nature of the recording industry renders virtually every venue subject to a potential suit against a major arm of the music industry.¹⁵⁹ Thus, the ability for strategic selection is seemingly limitless. This Note is ill-suited to address every instance of circuit court disagreement, but certainly *Bridgeport*’s controversial reluctance to find a *de minimis* exception to the copyright infringement of sound recordings is one glaring example.¹⁶⁰ Similarly, *La Cienega Music Co. v. ZZ Top*,¹⁶¹ a Ninth Circuit decision that declined to follow the only then-available appellate precedent, asserting that a phonograph recording did in fact constitute a “copy” within the meaning of the Copyright Act,¹⁶² was equally problematic. The court veered from Second Circuit precedent and created “an obvious conflict between the decisions of the Second Circuit and the Ninth Circuit,” each equally important for the largely Los Angeles and New York-based music business.¹⁶³ Fortunately, legislation was soon enacted that effectively

www.tnmd.uscourts.gov/higgins_bio (noting that Judge Higgins graduated from law school in 1957). Furthermore, it is easy to cringe when reading a judge’s analysis of music clearly absent from his or her personal knowledge base. See, e.g., *Armour v. Knowles*, 512 F.3d 147, 151 n.1 (5th Cir. 2007) (“The lyrics differ in word and substance: Armour’s evince somewhat less love for her ‘ba-by’—‘a l’il bit’—and somewhat more control over such love, which she can ‘stop’ or ‘make grow;’ Beyoncé’s suggest a deeper and more stubborn love, which pervades her thoughts, dreams, and ‘fan-ta-sies.’ Yet, though the nature and depth of their loves may differ, Armour claims they find expression by way of the same musical melody.”).

156. This assertion is based on the fact that juries are composed of twelve individuals over the age of eighteen. It is rare for a judge to ascend to a federal bench before the age of thirty-five, and usually not for a decade or so thereafter. Theoretically then, the average median juror will be younger, and hence “hipper,” than the average judge.

157. Michael Landau & Donald E. Biederman, *The Case for a Specialized Copyright Court: Eliminating the Jurisdictional Advantage*, 21 HASTINGS COMM. & ENT. L.J. 717, 738 (1999).

158. *Id.*

159. See *id.* at 723 (“[T]he potential explosion of jurisdictional possibilities makes more poignant the need for a more uniform, nationwide interpretation of the laws on copyright.”).

160. See *supra* notes 65–68 and accompanying text (articulating the *Bridgeport* controversy).

161. 53 F.3d 950 (9th Cir. 1995).

162. *Id.* at 953.

163. Landau & Biederman, *supra* note 157, at 750.

overruled the decision.¹⁶⁴ In the interim, however, as one commenter put it, “a typical response . . . of what to do subsequent to [*La Cienega*] was: ‘Sue in New York.’ Success or failure would have to depend on the outcome of a race to the courthouse.”¹⁶⁵ The doctrine has developed inconsistently in part because of its improper treatment. Additionally, in the instance of circuit splits, perhaps even more troubling is the “inevitable . . . ever more confusing body of law” which “arise[s] due to the disparate treatment.”¹⁶⁶

Another problem is the idea espoused by the term *battle of the experts*,¹⁶⁷ a frequent subject in legal academia. The concern is that the triable issues become less about the actual findings of fact required of the jury and more about an evaluation of which side’s experts are most believable.¹⁶⁸ Presently, most jurisdictions do not allow experts on intrinsic issues like the perspective of the lay audience but do permit expert testimony on extrinsic aspects of the plaintiff’s case.¹⁶⁹ Nevertheless, some jurisdictions allow experts on certain aspects of the lay-audience element of a successful copyright-infringement case.¹⁷⁰ For the elements in which experts are allowed to attempt to persuade the audience, the issue of the battle of the experts becomes less clear. Since this is the case in various facets of law, it

164. Pub. L. No. 105-80, 111 Stat. 1529 (1997). The legislation added § 303(b) to the Copyright Act. The remedial provision reads: “The distribution before January 1, 1978, of a phonorecord shall not for any purpose constitute a publication of the musical work embodied therein.” *Id.*

165. Landau & Biederman, *supra* note 157, at 750.

166. Alan Korn, *Issues Facing Legal Practitioners in Measuring Substantiality of Contemporary Musical Expression*, 6 J. MARSHALL REV. INTELL. PROP. L. 489, 497 (2007).

167. See Keyes, *supra* note 115, at 430–31 (“[T]he overabundant (and almost exclusive) reliance on the opinions of music experts has turned music copyright litigation into a ‘battle of the experts’ forum.”).

168. See *id.* at 436 (questioning whether it is appropriate to rely on experts to testify about how lay listeners perceive music).

169. Leon Friedman, *Substantive and Procedural Aspects of Copyright Litigation*, in *ADVANCED SEMINAR ON COPYRIGHT LAW 2004*, at 349, 380 (2004); see also *Segrets, Inc. v. Gillman Knitwear Co.*, 207 F.3d 56, 66 n.11 (1st Cir. 2000) (“While expert testimony may be used to establish that there was copying (from evidence of substantial similarity), where factual copying has been established, as here, there should be no expert testimony to establish whether or not there was substantial similarity.”).

170. In particular, there is one very important precedent for allowing expert testimony on the improper-appropriation prong. See *Dawson v. Hinshaw Music Inc.*, 905 F.2d 731, 736–38 (4th Cir. 1990) (“[W]e remand with instructions that the district court determine whether definition of a distinct audience is appropriate in this case. Assuming such a definition is appropriate, the district court should then take additional evidence . . .”). This has occurred most frequently in the context of computer software and other technical areas. See, e.g., *Kohus v. Mariol*, 328 F.3d 848, 858 (6th Cir. 2003) (holding that substantial similarity of a drawing of a latch should be considered from the viewpoint of the intended audience (not the lay public)); *Sturdza v. United Arab Emirates*, 281 F.3d 1287, 1300–01 (D.C. Cir. 2002) (noting that only the Fourth Circuit has allowed expert evidence to show substantial similarity outside of the context of computer programs); *Computer Assocs. Int’l, Inc. v. Altai, Inc.*, 982 F.2d 693, 713 (2d Cir. 1992) (instructing the district court to decide whether expert opinion is warranted regarding the substantial similarity of computer programs). This “technical” rationale is potentially applicable to future musical-copyright cases due to the increasingly electronic constitution of some newer genres of music.

is not always easily apparent why it is in any way particularly problematic in copyright law.¹⁷¹ The principal reason is that an American jury innately influenced by our nation's celebrity-obsessed culture is far more likely to trust a well-known musician's take on an allegedly infringing song than they would an unknown performer or songwriter.¹⁷² Furthermore, as Dolly Parton demonstrated, "you can make things sound like other things" rather easily, as she did in the courtroom by stylistically converting the Ray Charles song "I Can't Stop Loving You" to that of her hit "9 to 5."¹⁷³ Additionally, as one commentator said, "It is the extreme persuasiveness of expert testimony that most implicates public policy concerns. Testimony presented by an expert in the area of music might sway jurors unnecessarily."¹⁷⁴ In cases where the experts are allowed to hold sway over the lay-audience prong of infringement, this problem is even larger. There is something notably amiss when a group of lay individuals hears testimony bearing on what lay individuals—presumably not unlike the jurors themselves—should, and do, think.

The central underlying problem, which encompasses the other aforementioned concerns, is the notion of predictability. One commentator argued that "[copyright] must offer a degree of certainty and predictability in regulating and protecting [compositional] activity."¹⁷⁵ In order for copyright

171. See, e.g., Mark S. Brodin, *Behavioral Science Evidence in the Age of Daubert: Reflections of a Skeptic*, 73 U. CIN. L. REV. 867, 870 (2005) (assessing the costs and benefits of expert testimony on behavioral science); Thomas M. Crowley, *Help Me Mr. Wizard! Can We Really Have "Neutral" Rule 706 Experts?*, 1998 MICH. ST. L. REV. 927, 928 (noting that the "battle of the experts" is a cliché in legal culture); Gregory Dolin, *A Healer or an Executioner? The Proper Role of a Psychiatrist in a Criminal Justice System*, 17 J.L. & HEALTH 169, 171 (2002–2003) (noting the "battle of the experts" that occurs when psychiatrists participate in the criminal justice system); Michelle M. Mello, *Of Swords and Shields: The Role of Clinical Practice Guidelines in Medical Malpractice Litigation*, 149 U. PA. L. REV. 645, 647 (2001) (discussing the use of clinical practice guidelines, which derive from the expert consensus, as evidence in medical malpractice suits); Neil Vidmar & Shari Seidman Diamond, *Juries and Expert Evidence*, 66 BROOK. L. REV. 1121, 1127 (2001) (discussing social scientists' views on how juries handle expert testimony).

172. See, e.g., Alan M. Dershowitz, Op-Ed., *With Lawyers Like These . . .*, WALL ST. J., Mar. 8, 2004, at A16 ("The virtue of putting a celebrity witness on the stand is that her testimony would likely overshadow the other evidence. The jury would base its verdict largely on whether they believed or disbelieved her (or liked or disliked her)."). Indeed, there is reason to believe that the opinions of celebrities hold greater weight with the general public than those of noncelebrities. See Harry C. "Neil" Strine IV, *Stars on Capitol Hill: Explaining Celebrity Appearances in Congressional Committee Hearings 4* (Jan. 7–10, 2004) (unpublished manuscript), http://citation.allacademic.com/meta/p_mla_apa_research_citation/0/6/7/8/2/pages67823/p67823-1.php ("In total, more than 400 celebrities have appeared as witnesses in 288 congressional committee hearings [from] 1969 [to 2004]."). Much has been written on the problem in the context of criminal trials for celebrities. See, e.g., Jonathan D. Glater, *Weighing Celebrity Justice: Blind or Biased?*, N.Y. TIMES, June 15, 2005, at A14 ("[J]urors may be more willing to sympathize with celebrities than they have been in the past, as the lives of the famous have been subject to more scrutiny in recent years.").

173. Keyt, *supra* note 50, at 435 n.64.

174. Miah Rosenberg, *Do You Hear What I Hear? Expert Testimony in Music Infringement Cases in the Ninth Circuit*, 39 U.C. DAVIS L. REV. 1669, 1684 (2006).

175. Keyt, *supra* note 50, at 464.

law to exist as the industry and artist-supportive mechanism for creative incentivization it is frequently posited to be, there must be some semblance of formalistic order. Realistically, the law is only functional if its bounds are known by those expected and intended to rely upon it.¹⁷⁶ The present system provides no means for musicians to know what is and is not permissible. Who better to determine and enforce such uniformity than those with the best understanding of music? Despite the seemingly limitless possibilities presented in both mathematics and music, there are still, to some extent, a finite number of ideas and expressive possibilities within music.¹⁷⁷ Although new instruments¹⁷⁸ and genres¹⁷⁹ frequently arise, it is simply not the case that there is all too much real originality in popular music.¹⁸⁰ And even in cases of virtually undisputed innovation, there are often acknowledgements of impropriety.¹⁸¹ The inability to predict “makes it difficult to determine the portion that is available for the public to use and where that permission stops.”¹⁸² This shameful lack of clarity is in desperate need of reform.

176. See Carol M. Rose, *Possession as the Origin of Property*, 52 U. CHI. L. REV. 73, 81 (1985) (“Similar ideas of the importance of communication, or as it is more commonly called, ‘notice,’ are implicit in our recording statutes and in a variety of other devices . . .”).

177. See *Darrell v. Joe Morris Music Co.*, 113 F.2d 80, 80 (2d Cir. 1940) (“[W]hile there are an enormous number of possible permutations of the musical notes of the scale, only a few are pleasing; and much fewer still suit the infantile demands of the popular ear.”).

178. See, e.g., Jefferson Graham, *What IS That Thing? Futureman Explains the Drumitar*, USA TODAY (Aug. 6, 2004), http://www.usatoday.com/tech/news/techinnovations/2004-08-05-drumitar_x.htm (describing the “‘Drumitar,’ an electronic drum set-up in the shape of a beat-up guitar”); Jon Niccum, *That 1 Guy Concocts His Own Pipe Dream*, LJWORLD.COM (June 9, 2007), http://www.ljworld.com/news/2007/jun/09/1_guy_concocts_his_own_pipe_dream/ (discussing the “magic pipe,” a two-stringed percussive instrument that also produces smoke).

179. Wikipedia’s list of music genres is categorized alphabetically and divided into four different pages. Wikipedia, *List of Music Styles*, http://en.wikipedia.org/wiki/List_of_music_styles (last modified Feb. 24, 2012).

180. See *Everyone’s a Rock Critic: The Lost Lester Bangs Interview*, ROCKCRITICS.COM, <http://rockcriticsarchives.com/interviews/lesterbangs/lesterbangs.html> (transcribing a 1980 interview in which famed rock critic Lester Bangs disagreed with the notion that rock music was undergoing any major changes). Lester Bangs bemoans such lack of originality in rock music:

I mean, everyone’s acting like there’s this big renaissance going on [in rock music today], and it’s all the emperor’s new clothes. I mean, there’s a few groups that are doing really exciting things, and then there’s like all these phony power-pop groups on one side and all these phony synthesizer groups on the other, and I think it’s a big hype. I think it’s a lot of garbage It’s not like in 1977 when you had, you know . . . all these groups . . . and they all stood for something, they were *about* things. . . . [T]hey all had a real point of view about the world and they really, you know, really *meant* something, and these groups now, they’re all just interchangeable, they’re just singing piddly little love songs that don’t even matter.

Id.; see also *Emerson v. Davies*, 8 F. Cas. 615, 619 (C.C.D. Mass. 1845) (No. 4,436) (“In truth, in literature, in science and in art, there are, and can be, few, if any, things, which, in an abstract sense, are strictly new and original throughout.”).

181. See *Playboy Interview: Paul and Linda McCartney*, PLAYBOY, Dec. 1984, at 75, 107 (“We were the biggest nickers in town. Plagiarists *extraordinaires*.”).

182. Liebesman, *supra* note 79, at 334.

For example, one specific area in which the predictability concerns are relevant is self-reference. As popular music has matured, its genres and forms of art have grown increasingly self-referential.¹⁸³ Just as popular films and popular television shows often derive dialogue from characters aware of the influence of past media, various forms of music acknowledge this inherent interconnectivity.¹⁸⁴ With the audiovisual works, however, there can be no doubt where the line between idea and expression exists in analyzing, for example, a Kevin Smith or Quentin Tarantino film.¹⁸⁵ By contrast, with music, this line is often blurred.

Musicians and songwriters frequently err on the side of licensing, seeking to avoid any complications, particularly in cases where the referential aspects of the newer work are most readily apparent.¹⁸⁶ Conversely, perceiving conformity to industry specifications as impossible, artists will completely ignore—particularly in the rap or electronic context—the possibility of legitimacy and instead record without concern for copyright infringement and distribute illicitly.¹⁸⁷ Along with the circuit splits

183. One of the earliest uses of popular culture's self-reference canon of creativity came from John Lennon. *THE BEATLES, Glass Onion, on THE BEATLES* (Apple Records 1968) ("I told you about Strawberry Fields. You know the place where nothing is real . . . The walrus was Paul."). Much of contemporary rap and hip-hop music now routinely makes use of the canon in varying contexts and manners. See *Britney Spears, Eminem, Lady Gaga Play the Self-Reference Game*, MTV NEWSROOM BLOG (Dec. 20, 2010, 5:30 PM), <http://newsroom.mtv.com/2010/12/20/britney-spears-eminem-lady-gaga-sing-names> (suggesting that 2010 was "the Year That Everybody Sang Their Own Names").

184. See, e.g., *BUILT TO SPILL, You Were Right, on KEEP IT LIKE A SECRET* (Warner Brothers Records 1999) (quoting classic rock lyrics by artists like Pink Floyd, Bob Marley, Bob Dylan, The Rolling Stones, and The Doors by prefacing each quotation with either "you were right when you said . . ." or "you were wrong when you said . . ."). Given its penchant for self-reference, it is not surprising that lyrical references to previous music are prevalent in the genre of hip-hop. See, e.g., *JAY-Z, Hola' Hovito, on THE BLUEPRINT* (Roc-a-Fella 2001) (referencing Frank Sinatra, "I'm the compadre, the Sinatra of my day. Ol' Blue Eyes my nigga, I did it my way" and the Notorious B.I.G., "if I ain't better than [B.I.G.], I'm the closest one"); *J. COLE ET AL., Looking For Trouble, on FRIDAY NIGHT LIGHTS* (Dreamville Records 2010) (referencing Jay-Z's line in "Hola' Hovito" but switching Notorious B.I.G. with Jay-Z, the man who gave J. Cole his first recording contract: "Never said I'm better than [Jay-Z], but I'm the closest one").

185. *But see JAY AND SILENT BOB STRIKE BACK* (View Askew Productions 2001) (jokingly offering a character's dialogue up as prophetic: "I think George Lucas is going to sue somebody").

186. See *KID ROCK, All Summer Long, on ROCK N ROLL JESUS* (Atlantic Records 2007) (using multiple samples of Lynyrd Skynyrd's "Sweet Home Alabama" and Warren Zevon's "Werewolves of London" to address the Skynyrd song's influence on Kid Rock's musical and social upbringing). The increasingly popular trend of musical "mashups" is also relevant to this point. E.g., *Glee* (Ryan Murphy Productions 2009).

187. This is perhaps best exemplified by the popularity and credibility of the rap mixtape market. See Shaheem Reid, *Mixtapes: The Other Music Industry*, MTV.COM, http://www.mtv.com/bands/m/mixtape/news_feature_021003/ ("The other music industry, the one where labels don't exist and there are no highly paid Lizzie Grubmans to publicize your new release, where the CDs are sold by vendors hawking them off dirty blankets on city streets, and bootlegging is encouraged."). *But see, e.g., GIRL TALK, NIGHT RIPPER* (Illegal Art 2006) (utilizing innumerable popular hit songs in newly created mashups and releasing the songs on albums via traditional methods of distribution without paying any royalties to, or even attempting to license from, the actual copyright owners).

exemplified by *Bridgeport's* incongruence, these creative inhibitions demonstrate that uncertainty is rampant. A more easily discernible, clarified, and universally applicable system of infringement analysis would, over time, establish a more consistent framework within which artists would know where they stand. In some instances, artists would likely have more leeway to avoid the transactional costs of licensing, and in others, they might be able to alter a few aspects of their recordings for the purpose of creating music legally and obtaining all the benefits of promotion and distribution that come with it. A predictable musical-copyright scheme is therefore essential to the long-term legitimacy of the legal doctrine as a whole and its ostensible goal of promoting creativity, particularly with respect to the proliferation of trends like self-reference and sampling.

V. Solution

As Part IV demonstrated, there are several problems with the present system. Some argue that the ideal solution is also the least novel: a return to “the *Arnstein* lay listener inquiry in its purest form.”¹⁸⁸ Others have proposed slight alterations in application or formulation that fail to solve problems of uniformity.¹⁸⁹ Distinctively, this Note argues that real reform will require legislative action. The fundamental practical and academic duality of musical copyrights prevents the lay listener from properly performing its task. What is necessary is nothing short of a separate and distinct musical-copyright-infringement tribunal staffed by academic musicologists and armed with jurisdiction to handle the vital determination in all qualifying cases and, more broadly, to unify the doctrine’s inconsistencies.

A. *What Is at Stake*

As this Note has outlined, the ultimate crux of the argument boils down to whether we as legal scholars, laypersons, and influential legislators believe that there is in fact a correct answer as to whether one song infringes on the copyright of an earlier song. Ontologically, one would assume that those most knowledgeable about music from a scientific perspective are those best equipped to make this determination. Epistemologically, given that we as a society are intensely opinionated about matters of originality (and taste), there must be a reasonably discernible, objectively accurate answer. If not, this would seem to suggest that the fundamental system of musical copyright,

188. Austin Padgett, *The Rhetoric of Predictability: Reclaiming the Lay Ear in Music Copyright Infringement Litigation*, 7 PIERCE L. REV. 125, 146 (2008).

189. See, e.g., Brent, *supra* note 149, at 253 (stressing the necessity for a “slightly restructured *Arnstein* . . . analysis specifically tailored to music’s special qualities and needs”); Keyt, *supra* note 50, at 421 (“[A] better way to balance competing creative interests would be to focus on rights more specific than the vague ‘right not to be infringed’ which underlies the existing formulations. One of these more specific rights is the right to exploit fully the market for one’s creative product, free from competition by works copied from it.”).

if not the doctrine as a whole, is flawed. Even most legal realists acknowledge the allure of formalistic predictability in the law.¹⁹⁰ If a verdict of infringement or non-infringement based on similarity hinges on factors totally unrelated to what is actually at stake (the composition or sound recording itself), predictability, at least to the *de minimis* degree of desirability, is nonexistent. This opens the field of music law and the music industry as a whole, which obtains much of its legitimacy (or what remains of it)¹⁹¹ from the legal doctrine of musical-copyright law, to a whole slew of criticisms. These criticisms are increasingly viable and relevant in an age when copyright entitlements are routinely and exponentially violated via digital technology.¹⁹² One commentator noted that “[b]ecause economic considerations suppl[ied] the justification for copyright law in the United States, changed economic conditions should lead policymakers to re-examine the scope and content of copyright law as applied to music.”¹⁹³ As the music industry struggles to redefine its core business model, the legal field can do its part by offering suggestions to stem the tide of diminishing legitimacy and increasing cynicism toward the law of musical copyrights. Nationwide reform of the infringement doctrine could serve to encourage positive reform in other areas.

Unfortunately, for several reasons it is very difficult to really know the extent of the problem in the infringement doctrine that this Note proposes to reform. Firstly, because the music industry has repeatedly shown a propensity to avoid litigation—particularly where the allegedly infringing song is successful as is usually the case if litigation is brought—it is very difficult to determine the frequency of allegations.¹⁹⁴ Litigation is both costly and

190. See, e.g., James E. Moliterno, *Why Formalism?*, 49 U. KAN. L. REV. 135, 136–38 (2000) (noting that lawyers and clients alike are more comfortable with formalism, as opposed to realism, due to its level of predictability).

191. See *What Will Music Fans Pay For?*, ZED EQUALS ZEE (Oct. 9, 2009), <http://zedequalszee.com/2009/10/09/what-will-music-fans-pay-for/> (discussing the questionable viability of a business model based on charging consumers for a product they can obtain for free).

192. See Marc F. Bellemare & Andrew M. Holmberg, *The Determinants of Music Piracy in a Sample of College Students* 21 (Working Paper, May 20, 2010), available at <http://ssrn.com/abstract=1481272> (analyzing the illicit downloading habits of college students and finding, among other alarming statistics, that more than 30% of respondents “reported that their last song had been obtained illegally”).

193. Carroll, *supra* note 72, at 1411.

194. *Music Copyright Infringement Resource: Case List*, UCLA SCH. L. [hereinafter *Case List*], available at <http://cip.law.ucla.edu/cases/Pages/default.aspx> (“Stories about claims of music copyright infringement appear fairly regularly in *Variety* and other mainstream publications. Few of these disputes go to trial, and fewer still generate published judicial opinions. Since the 1850s federal courts have published over 100 opinions dealing with this issue, but the frequency with which these cases arise has increased markedly over the past twenty years.”); see also Mark Caro, *Music Lyric Infringement Cases Are a Murky Area*, CHI. TRIB. (July 11, 2010), http://articles.chicagotribune.com/2010-07-11/entertainment/ct-ae-0711-ripoff-20100711_1_infringement-copyright-lyric (“The Beatles’ ‘Come Together’ shared a similar cadence and melody with Chuck Berry’s ‘You Can’t Catch Me,’ but the clincher may have been John Lennon’s singing ‘Here come old flat-top’ as opposed to Berry’s ‘Here come a flat-top.’ (Lennon settled.). Given the musical and

potentially deleterious for the longevity of the song and artist because of the negative press it tends to generate.¹⁹⁵ Secondly, as Michael Keyes noted, “[f]rom 1950 through 2000, there were forty-three reported cases dealing with music copyright infringement—nearly twice as many as compared to the period between 1900–1950—and many more disputes that never ripened into litigation as a result of out-of-court settlements.”¹⁹⁶ Thus, while forty-three cases over a fifty-year period might not seem indicative of a problem worthy of reform, the number is indicative of a larger problem. If the purpose of copyright law, and intellectual property law as a whole, is to encourage creativity,¹⁹⁷ it is easy to see how the doctrine’s indeterminacy can damage artistic freedom. Firstly, it may make musicians shy away from those creative issues about which they are confused, particularly with respect to increasingly relevant self-referential musical statements. Secondly, given that today’s breakout artists are not “discovered” in the traditional sense anymore,¹⁹⁸ the doctrine’s ambiguities may dissuade newer artists from

lyrical similarities of their ‘Forever Young’ songs, Rod Stewart wound up giving Bob Dylan a songwriter credit.”); *Musical Copyright Infringement Cases*, HARTFORD COURANT, <http://www.courant.com/features/hc-copyrightcases-pg,0,3310392.photogallery> (providing a click-through slide show of various infringement cases and threatened lawsuits as well as their outcomes).

195. See, e.g., *Case List*, *supra* note 194 (chronicling all instances of musical-copyright-infringement litigation and threatened litigation in the U.S. and abroad); see also Rohan Ramakrishnan, *The 5 Most Famous Musicians Who Are Thieving Bastards*, CRACKED (May 6, 2010), http://www.cracked.com/article_18500_the-5-most-famous-musicians-who-are-thieving-bastards.html (“[S]ome of the most successful musical acts in history based huge chunks of their careers entirely on plagiarism.”).

196. Keyes, *supra* note 115, at 418.

197. See, e.g., Rufus Pollock, *Copyright and the Digital Age*, in PROMOTING INNOVATION AND REWARDING CREATIVITY: A BALANCED INTELLECTUAL PROPERTY FRAMEWORK FOR THE DIGITAL AGE 18, 18 (2006), available at http://www.thersa.org/_data/assets/pdf_file/0006/54834/adephi-promoting-innovation-and-rewarding-creativity-january-2007.pdf (“[C]hanges in entry costs imply alterations in the structure of the industries affected: as entry costs fall there should be a large, and rapid, increase in the number and diversity of those making and distributing copyrightable work (a democratisation of creativity).”).

198. See, e.g., Bob Lefsetz, *I Want My MTV*, LEFSETZ LETTER (Oct. 17, 2011), <http://lefsetz.com/wordpress/index.php/archives/2011/10/17/i-want-my-mtv/> (describing how new technologies have changed the music industry). Lefsetz explains,

Lady Gaga is the first Web star. She wasn’t broken by radio, but by videos online, she established that paradigm. But even more important Gaga was the first social network superstar. She realized it was about the relationship with the fan first and foremost. You could create that kind of bond online, the customer wasn’t the label or the radio station or the TV outlet, it was the end user. As for the music itself, that came last.

In the MTV era the music came first. If it wasn’t an approved genre, you were SOL. Sure, you had to look good and have money behind you, but if you didn’t sound like what MTV was playing, you had no chance. Now just the opposite is true. The way you connect with your fans comes first. The bond is the initial attraction. The music comes last. The point is the music can sound like anything, there are no limits, no rules, no genres you must fall into. You’ve just got to have a relationship with your fans and consistently honor it, put them ahead of all other interests.

Id.

spreading their work through modern distribution channels¹⁹⁹ for fear of being ripped off without recourse. As the music industry's struggles continue, the trend of increasing litigation is likely to increase over the next half century unless steps are taken to reform a doctrine that values confusion and uncertainty over creativity.

It would be disingenuous to assert in any serious discussion of public policy that the musical copyrights are as important, from either a financial or a philosophical perspective, as the field of patents. Nevertheless, this Note recommends sweeping reform similar to that accomplished by the creation of the Federal Circuit for the purpose of bringing uniformity to American patent doctrine. There are several reasons why the proposal outlined below would be financially and logistically worthwhile to the federal government. Firstly, as rampant piracy and declining streams of revenue have become reality, and as the survival of the major record labels has been placed squarely in doubt,²⁰⁰ no effort has yet been made to restore legitimacy to the musical copyrights for younger generations raised in an age of Napster and Kazaa. Imprecise, impractical doctrines like Creative Commons threaten the stability of the musical copyrights.²⁰¹ Reform is necessary to reestablish the dominance of the copyright regime in the musical arts. Secondly, in the age of the Internet, domestic intellectual property law is routinely threatened by the behavior of various international actors. By virtue of how interconnected the world has become, piracy in one part of the world is no longer confined just to that region.²⁰² The "creation" of a substantially similar infringing work is no less a form of piracy than file sharing. Music industries around the world often look to the United States—given the size and scope of its industry—as a progressive innovator.²⁰³ Thus, the United States is in a

199. See, e.g., Steve Knopper, *The New Economics of the Music Industry: How Artists Really Make Money in the Cloud—or Don't*, ROLLING STONE (Oct. 25, 2011), <http://www.rollingstone.com/music/news/the-new-economics-of-the-music-industry-20111025> ("Today, music fans play free music videos on YouTube, stream songs for free on Spotify, MOG or Rdio, customize Internet radio stations on Pandora or Slacker and consume music a zillion different ways. The fractions of pennies artists make for each of these services are nearly impossible to track . . .").

200. See Jason Feinberg, *What Will Record Labels Look Like in the Future?*, PBS.ORG (Aug. 18, 2009), <http://www.pbs.org/mediashift/2009/08/what-will-record-labels-look-like-in-the-future230.html> ("The majority of label income has traditionally come from selling physical product, which is the area now seeing the sharpest decline. As a result, record companies have experienced a rocky decade, seeing wave after wave of layoffs and massive decline in revenues.").

201. See John C. Dvorak, *Creative Commons Humbug*, PCMAG (July 18, 2005), <http://www.pcmag.com/article2/0,2817,1838249,00.asp> (postulating that the Creative Commons regime may actually weaken traditional copyright protection by overcrowding copyright doctrine with redundant rules and provisos).

202. See Lamar Smith, *Why We Need a Law Against Online Piracy*, CNN (Jan. 20, 2012), <http://edition.cnn.com/2012/01/20/opinion/smith-sopa-support/> (noting that the growth of online piracy promulgated by overseas websites costs the United States as much as \$100 billion each year).

203. See generally Michael F. Mertens, *Thieves in Cyberspace: Examining Music Piracy and Copyright Law Deficiencies in Russia as It Enters the Digital Age* 15 (Berkeley Elec. Press, Working Paper No. 663, 2005) (comparing U.S. policies to combat piracy with approach taken by the Russian music industry); *id.* at 14 (describing the influence of international organizations on the

unique position to make a statement to the world that it will not tolerate piracy of any kind and that it is serious about creating a coherent, fiscally sound musical-copyright policy. Finally, the biggest benefit of bringing patent-like uniformity to musical copyrights is the benefit of predictability to the musicians and artists. As music becomes increasingly referential both musically and lyrically, the line between acceptable and illicit “copying” has been blurred beyond recognition. Money does not just flow from consumer to owner in the music industry. Rather, money flows to content owners from content licensors in various ways across several types of businesses.²⁰⁴ Given how pervasive music is in all of our lives, music and the business of music is incredibly important. When increasingly vast sums of money are at stake, it is easy to justify this Note’s proposal on a policy basis for the greater good of preserving the legitimacy of the various small and large business models that are affected every day by a musical-copyright system that does not work correctly.

B. The Proposal

This Note proposes a set of musicology tribunals under one umbrella entity headed by a chief justice, with three musicologists comprising each tribunal and making legal determinations on similarity for each case. The tribunals would not replace the traditional judicial process and would determine a small portion of only those musical infringement cases held to satisfy other necessary elements. Cases would originate in venues compliant with appropriate procedural rules where jurisdiction is proper. The issue of a valid copyright and the copying prong would still be heard in the ordinary venue under existing rules. If the plaintiff satisfies these elements, the case would proceed to one of the musicologist-staffed tribunals for the all-important determination of substantial similarity under the improper-appropriation prong. Different tribunals would exist for specific genres and styles; for example, there would be a rap-hip-hop tribunal comprised of musicologists well-versed in the unique process of writing, creating, and recording music in these genres and sub-genres.²⁰⁵ This would be the case for everything from country to bluegrass and classical to vocal, including special exceptions for newer styles like trance and drum ’n bass. In the instance that a case arises involving these highly particularized genres, the tribunal fluent in the genre would hear the case. Otherwise, tribunal

development of Russian copyright law, namely the U.S.-driven International Intellectual Property Alliance); *id.* at 40–42 (illustrating how U.S. copyright policies have forced China and Thailand to adopt more stringent antipiracy regulations).

204. See Gary Myers & George Howard, *The Future of Music: Reconfiguring Public Performance Rights*, 17 J. INTELL. PROP. L. 207, 211–24 (2010) (describing the development of performance royalties, mechanical licenses, and record-label contracts).

205. There would not be a tribunal for every genre. The author acknowledges that this would be impossible and impractical. Rather, each tribunal would be well-versed in several genres and sub-genres, and there would be much overlap.

assignments would be relatively random, based primarily on caseload and nonconflict.

Generally speaking, the tribunal's mode of analysis will not significantly differ from the initial *Arnstein* formulation. Crucially, however, it will specifically tailor the framework to take advantage of the many ways in which music differs from other copyrightable works. The tribunal's approach, uniquely constituted to impose predictability on the present system of chaos, will restore credibility to musical copyrights. The tribunal will consider similarity from a dual perspective. First, drawing inspiration from Pandora's objectification of music, the tribunal will ask to what degree the allegedly infringing work shares certain traits with the plaintiff's work. For both musical composition and sound recording cases, there would be five traits: lyrics (if there are any, evaluating both content and theme), melody, harmony, rhythm, and miscellany.²⁰⁶ Notably, this process, presided over by musicologists with a specialized understanding of music, will be far less susceptible to inherent human subjectification and inconsistency. With respect to a determination of objective similarity, the traits would be considered on a sliding scale of significance, dependent on musical genre or other circumstances and as determined on a case-by-case basis by the tribunal's members.²⁰⁷

If the tribunal, proceeding on an analysis of all five traits, deems objective similarity to exist, the tribunal will then determine if the allegedly infringing work has subjectively taken "what is pleasing" from the plaintiff's work. It is this second portion of the analysis that will enable the musicologists to gauge infringement from the perspective of the lay listener. The tribunal's members are qualified to analyze from this perspective because their academic backgrounds permit them to understand both what is commercially viable about a work and also what is personally enjoyable about that work. In the sound-recording context, the distinctiveness of the defendant's usage of the plaintiff's work will be taken into account. This is particularly relevant in the rap-hip-hop genre, in which the process of sampling has been significantly complicated by *Bridgeport*. The usage will be only part of the evaluation process of subjective similarity of the copyrightable portions of the plaintiff's and defendant's works as a whole. Primarily, the tribunal will analyze whether, from the perspective of the

206. See ROOKSBY, *supra* note 81, at 10 ("A song has four basic elements: words (the lyric[s]) are sung to a tune (melody) that is supported by chords (harmony) and played to a certain combination of beat and tempo (rhythm). . . . In different styles of music the balance of importance between these elements will change."). Miscellany is included to allot for subtleties not taken into account by the four principal traits, as articulated by Stephanie Jones. See *supra* notes 123-24 and accompanying text.

207. The methods of analysis will be based on a loose rubric structure and drawn from commonly accepted professional standards. The standards will be determined by a committee comprised of academic experts in musicology, copyright, and procedural law prior to the first action of any of the tribunals.

average listener of the allegedly infringing work, the elements of the work that draw the listener to it are sufficiently similar. The tribunal may also view independently collected data from surveys of laypersons as funded and provided by the plaintiffs. Thus, this secondary lay determination fixes many of the problems associated with the present system without rewriting the law. Vitality, the tribunal's dual analysis keeps the subjective and objective determinations of similarity separate—as they ought to be—rather than conflated and confused by an uninformed lay jury.

The appeals process would be one of statutorily prescribed uniformity, further ensuring consistency. A termed appointment of a chief justice of the musicology tribunals would be made by the Librarian of Congress.²⁰⁸ Ideally, the chief justice would be both a music scholar and a juris doctorate. Any dissatisfaction with a ruling on substantial similarity by one of the musicology tribunals could be appealed to the chief justice. The chief justice's word, whether in the form of a complete denial of certiorari (requiring no explanation) or a full review and reversal (necessitating a written opinion), would be *ipso facto* final. After the tribunal assigned to a particular case makes its ruling and the appeals process is exhausted or not sought, the case would be remanded to the court from whence it came for a judgment consistent with the tribunal's opinion. This would include, in cases where the improper-appropriation prong is held to be satisfied, a judge or jury determination of proper damages.²⁰⁹ The ordinary appellate process

208. This is the general process by which appointments are made to what is presently called the Copyright Royalty Board. See MERGES ET AL., *supra* note 13, at 518 (“A 2004 law replaced the [Copyright Arbitration Royalty Panels] process with three Copyright Royalty Judges (CRJ), full-time employees of the Library of Congress appointed for six-year terms with an opportunity for reappointment.”). But see *SoundExchange, Inc. v. Librarian of Cong.*, 571 F.3d 1220, 1226 (D.C. Cir. 2009) (Kavanaugh, J., concurring) (questioning without provocation the constitutionality of the statutory structure of the Copyright Royalty Board); Complaint for Declaratory and Injunctive Relief at 2, *Live365, Inc. v. Copyright Royalty Bd.*, No. 1:09-cv-01662 (D.D.C. Aug. 31, 2009) (“This action seeking declaratory and injunctive relief is a facial constitutional challenge under the ‘Appointments Clause’ of the U.S. Constitution, U.S. Const. art. II, § 2, cl. 2, to the formation of the Copyright Royalty Board (‘CRB’).”). It is unclear what the result of this proceeding will be. The author suggests that if a court finds the CRB appointment process to be unconstitutional, that the musicology tribunal members should alternatively be appointed consistent with that ruling.

209. It is important to note that this Note's proposal does not inhibit the Seventh Amendment right to a trial by jury because only one part of a much larger consideration is actually taken from a jury, if one is constituted, and given to the musicologist-staffed tribunal. Crucially, the determination of damages remains unaffected. Should a challenge to the proposed tribunals be raised on Seventh Amendment grounds, there is some authority to suggest the possibility of a “complexity exemption.” See Jason Weeden, *Historically Immune Defendants and the Seventh Amendment*, 74 TEXAS L. REV. 655, 656–57 & n.9 (1996) (surveying the history and scholarship around the “complexity exemption”). The exemption was first raised by Justice White in a 1970 opinion, in a footnote from which much academic speculation and authorship ensued. See *Ross v. Bernhard*, 396 U.S. 531, 538 n.10 (1970) (positing a third factor to the existing two-part test, “the practical abilities and limitations of juries”); see also Douglas King, Comment, *Complex Civil Litigation and the Seventh Amendment Right to a Jury Trial*, 51 U. CHI. L. REV. 581, 613 (1984) (“The seventh amendment calls for the preservation of the right to a jury trial in ‘Suits at common law.’ In implementing this constitutional provision, the Supreme Court has developed a ‘flexible’ historical test that focuses primarily on the substantive aspects of the litigation, namely the rights

with respect to all elements of the case not determined by the musicology tribunals would not be altered. However, where a previous circuit split is presented, the chief justice would always be required to file a special amicus brief, taking into account all sides of an issue and offering an unbiased determination that, by virtue of the position's reverence, would likely have great influence on the court's final determination. Additionally, parties displeased by a judge's or jury's holding on access, given its inclusion under the same improper-appropriation prong, could also seek recourse with the chief justice, albeit of a different kind. In these cases, the chief justice would lack formal jurisdiction to issue a controlling order. But over time, the chief justice's persuasive rebukes and recommendations would help to establish uniformity in the rest of the doctrine.

Audio recordings would always be necessary in sound-recording-infringement cases. For compositions, recordings would only be relevant where the songs were constructed in a less traditional arrangement process.²¹⁰ Also, some written form of the works in question would always be required for compositions. The mode of tangible expression must be variable given the diversity of genres. This would be permissible, so long as the written version corresponded, at least vaguely, with some preexisting format of musical scholarship. Amicus briefs would be permissible in order to allow the tribunal to understand the larger context, whether it be legal, economic, or in some rare instances, political.²¹¹

asserted and remedies sought. This comment has argued for a reevaluation of the jury right under the seventh amendment. Statistical analysis of litigation in the common law and equity courts of England during the period of 1789 to 1791 demonstrates that, by comparison to the equity courts, common law cases were very narrow in scope and, by virtue of the existing procedural system, complex cases were not possible at common law.”)

210. It is unquestionably the case that composition no longer always takes place via traditional tangible songwriting methodologies. See, e.g., Michele Traverso, *Desktop Composition Software Changes the Face of Music*, WIRED (Aug. 23, 2011, 7:35 PM), <http://www.wired.com/gadgetlab/2011/08/nodal-music-composition/> (reviewing Nodal, a new music composing software program “designed to create and visualize music,” which aims to revolutionize how music is composed). Songs are often written by individual band members and only fully fleshed out in the full-band setting afterwards. See POWELL, *supra* note 136, at 225 (describing an improvisational composition process that begins with a “tune as it was originally written . . . and then making up your own variations”). Rarely do popular-music bands compose in the formal sense of traditional notation as was the case in the early days of popular music typified by the Brill Building. See Ian Inglis, “*Some Kind of Wonderful*”: *The Creative Legacy of the Brill Building*, 21 AM. MUSIC 214, 218 (2003) (“[T]he working practices of the Brill Building presented an archetypal example of ‘vertical integration.’ Containing the offices of record companies, publishers, managers, composers, and promoters, it functioned in a way that has been rather glibly described as a ‘production line’ or ‘songwriting factory.’”); see also Carroll, *supra* note 72, at 1492 (“[D]igital technology also makes possible new ways to create music. . . . [C]omputers can ‘listen’ to recorded or live music and analyze and then translate what they ‘hear’ into visual representations in musical notation. . . . Other forms of composition software extend the abilities of amateurs to express themselves in musical notation.”).

211. It is certainly possible that future musical works and their alleged infringement might bear on major political or foreign policy issues. The obvious example would be with regard to the harmonization of American and international copyright law. Another possibility would involve the ramifications of particular findings with regard to larger American foreign policy, however unlikely

In *McCulloch v. Maryland*,²¹² Chief Justice John Marshall was insistent in advocating against any semblance of judicial discretion.²¹³ And yet, although this problem is certainly not unique to the copyright context, it prevails to a fault, particularly with musical copyrights. The fact remains that (most) judges are not musicians, songwriters, or even, to say the least, creative-minded persons,²¹⁴ nor are they intended to be. As Justice Oliver Wendell Holmes Jr., a relatively artistic adjudicator, admitted, "It would be a dangerous undertaking for persons trained only to the law to constitute themselves final judges of the worth of [certain creative works], outside of the narrowest and most obvious limits."²¹⁵ Furthermore, given the procedural complications regarding when judges should question a jury's decision,²¹⁶ and given the confusing application of the varied case law,²¹⁷ something must be done to limit judicial discretion in the musical-copyright context. Some form of standardized judicial analysis is necessary to render uniform decisions and clearly and quickly elucidate conflicts toward the goal of precision and efficacy. Musicology is "[t]he historical and scientific study of music."²¹⁸ Thus, the musicology tribunals would be staffed by academics fully cognizant of the varying ways in which music is written and recorded and the complex processes through which humans interact with music.

this eventuality might seem. The viability of music with regard to the political world has not gone unnoticed by legal scholars. See generally Carol Weisbrod, *Fusion Folk: A Comment on Law and Music*, 20 CARDOZO L. REV. 1439 (1999) (using music as a metaphor for politics and law). Additionally, there is no shortage of musicians and songs credited with fomenting or at least planting the seeds for political change. See, e.g., BOB DYLAN, *Blowin' in the Wind*, on THE FREEWHEELIN' BOB DYLAN (Columbia Records 1963) (stirring support for an increasingly influential civil rights movement by asking rhetorically, "How many years can some people exist before they're allowed to be free?"); USA FOR AFRICA, *We Are the World*, on WE ARE THE WORLD (Columbia Records 1985) (calling on the people of the world to unite to help people dying in Africa); U2, *Sunday Bloody Sunday*, on WAR (Island Records 1983) (lamenting the death of the Irish civil rights protesters who were infamously shot by British paratroopers in Northern Ireland on "Bloody Sunday").

212. 17 U.S. (4 Wheat.) 316 (1819).

213. *Id.* at 423 ("But where the law is not prohibited, and is really calculated to effect any of the objects entrusted to the government, to undertake here to inquire into the degree of its necessity, would be to pass the line which circumscribes the judicial department, and to tread on legislative ground.").

214. *But see* Noble v. Bradford Marine, Inc., 789 F. Supp. 395, 396–97 (S.D. Fla. 1992) (structuring the largely procedural legal opinion around quotations and references from the film WAYNE'S WORLD (Paramount Pictures 1992)).

215. *Bleistein v. Donaldson Lithographing Co.*, 188 U.S. 239, 251 (1903).

216. See *Selle v. Gibb*, 741 F.2d 896, 902–03 (7th Cir. 1984) (sustaining judgment notwithstanding the verdict over a jury finding in favor of the plaintiff on grounds that "although proof of striking similarity may permit an inference of access, the plaintiff must still meet some minimum threshold of proof which demonstrates that the inference of access is reasonable").

217. See Brent, *supra* note 149, at 252 ("Most courts have not adhered to the *Arnstein* test in its literal terms; rather, they have combined and altered the *Arnstein* test into [a] more viable, workable approach.").

218. THE AMERICAN HERITAGE DICTIONARY OF THE ENGLISH LANGUAGE 1191 (3d ed. 1992).

C. Precedents for Reform

The solution proposed within this Note is not as novel as it may first seem. Landau and Biederman have “recommend[ed] the creation of a national copyright court that follows the model of the Court of Appeals for the Federal Circuit.”²¹⁹ The Federal Circuit was established legislatively in 1982 with the express goal of hearing all patent appeals.²²⁰ Additionally, it was instituted, albeit perhaps less directly, with the goal of bringing the existing patent law doctrine into uniformity in order to serve the original interests of the system and the doctrine as a whole.²²¹ One commenter noted that “[i]t is a model that can be studied and perhaps adapted to other fields to ward off . . . [the] growing threat to doctrinal coherence posed by ever-increasing appellate caseloads and judgeships.”²²² Musical copyrights occupy the field most in need of such coherence. Logistically, the Federal Circuit has been largely successful in “bring[ing] cohesiveness to the patent law.”²²³ Donald R. Dunner enthused that “the court not only cleared up the confusion that thirty years of practice under the 1952 Patent Act had not, but resolved more than a dozen conflicts which had existed in the patent law.”²²⁴ It has dramatically reduced forum shopping and ensured that venue selection is “no longer driven by judicial interpretation or application of the law.”²²⁵ The Federal Circuit is a perfect example of an Article III court that, although a seemingly radical reform at its time of inception, has ushered in an unprecedented era of patent-doctrine coherence.²²⁶ As such, it illustrates the viability of the notion that the proposed musicology tribunals will bring consistency to musical copyrights.

Another parallel to the proposal lies in medical malpractice tribunals. Massachusetts has found success in its “special panel[s] consisting of a judge, a lawyer, and an expert or specialist familiar with the subject matter of the particular lawsuit.”²²⁷ More importantly, it has altered its “system of fee

219. Landau & Biederman, *supra* note 157, at 719.

220. MERGES ET AL., *supra* note 13, at 123.

221. *Id.*

222. *Special Session of the United States Court of Appeals for the Federal Circuit Commemorating its First Ten Years*, 2 FED. CIR. B.J. 267, 273 (1992) (remarks of Daniel J. Meador).

223. Landau & Biederman, *supra* note 157, at 777. *But see id.* at 777 n.301 (providing examples where the Supreme Court intervened to resolve Federal Circuit disputes).

224. *Federal Circuit Bar Association and American Bar Association Tenth Anniversary Celebration of the United States Court of Appeals for the Federal Circuit*, 2 FED. CIR. B.J. 297, 301 (1992) (remarks of Donald R. Dunner).

225. *Id.*

226. *See supra* note 224 and accompanying text. The Court of International Trade is another good example. It has exclusive jurisdiction over a number of different civil actions centered on international business transactions and, in fact, has its decisions reviewed for appellate purposes by the Federal Circuit. 17 CHARLES ALAN WRIGHT, ARTHUR R. MILLER, EDWARD H. COOPER & VIKRAM DAVID AMAR, *FEDERAL PRACTICE AND PROCEDURE* § 4103 (3d ed. 2007).

227. John M. Greaney, *Reflections on Solomon and Other Forms of Progress*, 30 NEW ENG. L. REV. 919, 924 (1996).

distribution, and has managed to eliminate a significant number of non-meritorious cases.”²²⁸ The state even offers parties the option of presenting their case to “blue ribbon” juries, comprised of specialists familiar with the context, facts, and technical considerations of particular cases.²²⁹ The proposed musicology tribunal would initiate similar changes in the overall constitution of the remedy, providing protection from frivolity to established songwriters and musicians just as Massachusetts provides protection to respected doctors and medical professionals. Of paramount significance, the tribunals would be most focused on getting the correct answer to the question of liability.

D. *Refuting the Objections*

Perhaps the biggest objection to the proposed solution comes from those who claim that the “lay ear” ought to be the exclusive arbiter of musical-copyright law.²³⁰ These arguments primarily proceed from the notion that academics are predominantly not the audience enjoying the benefits of the music, so their opinions are no more relevant to the legal issue of infringement than those of the layperson.²³¹ But we would not entrust the interpretation of our Constitution to everyday persons just because they would be more frequently affected by the results than our most esteemed legal scholars and judges. The fact remains that people creating and analyzing music are nearly always more knowledgeable about its structure than the average lay listener.²³² Furthermore, music is capable of precise mathematical analysis, which sets it apart from the other enumerated works and demands exceptional consideration.²³³

Another objection likely to arise from the proposed solution is the cynical notion that there is no guarantee of any resulting positive reform. To these detractors, the author must concede that nothing is ever certain,

228. *Id.* at 925.

229. *Id.*

230. *See, e.g.,* Padgett, *supra* note 188, at 126 (“[T]he simplest and best approach to music copyright infringement litigation is to accept the jury’s determination of substantial similarity in its most classic form.”).

231. *See id.* at 149 (“[A]rtists, publishers, and record companies (and their lawyers) should accept the risk [posed by the ‘lay ear’] that their industry operates within a realm where art and commerce may lead to ‘indescribable’ results.”); *cf.* *Markman v. Westview Instruments, Inc.*, 517 U.S. 370, 388–90 (1996) (granting responsibility for construction of terms of art to the judge because “[t]he construction of written instruments is one of those things that judges often do and are likely to do better than jurors unburdened by training in exegesis”).

232. *See supra* note 140 and accompanying text; *see also* STORR, *supra* note 140, at xii (“It is also true that people who can play an instrument, or who can sing, can actively participate in music in ways which enrich their understanding of it.”). *But cf.* CLAUDE LÉVI-STRAUSS, *THE RAW AND THE COOKED* 18 (John Weightman & Doreen Weightman trans., Harper & Row 1969) (1964) (“[S]ince [music] is the only language with the contradictory attributes of being at once intelligible and untranslatable, the musical creator is a being comparable to the gods, and music itself the supreme mystery of the science of man . . .”).

233. *See supra* subpart III(A).

particularly in the complicated socioeconomic and technological balancing act of domestic intellectual property policy. But this Note's proposal and its predicted results seem not only logical but also to some extent inevitable. Nevertheless, even in the absence of the expected improvements of gaining predictability, a correct result, and the ability to deal properly with newly and recently created genres, the musical-copyright-infringement regime would be greatly benefitted by the uniformity unquestionably offered by the musicology tribunals. This institutional mechanism would serve as a larger counterweight to all of the confusion and chaos that has prevailed until now. Furthermore, it is certainly possible that over time, should the necessity arise, changes could easily be initiated—either statutorily via amendment or administratively via the Librarian of Congress—to remedy any concerns.

VI. Conclusion

This Note proposes a set of musicology tribunals headed by a chief justice equally versed in both music and law. There is no way to look into the future and know for certain how successful the proposed tribunals will be in providing solutions to the numerous aforementioned problems with musical-copyright-infringement law. Nevertheless, there are several reasons why reform is preferable and necessary. The most essential of these rationales is the comfort and stability of uniformity, exemplified by the Federal Circuit's successful reform of the patent doctrine. There are countless calls for reform across virtually every aspect of the music industry, as wide-ranging in scope and justification as new experimental sources of revenue and pending legislation. The industry suffers increasingly from a growing perception of illegitimacy and impropriety. Legislative reform would demonstrate substantial commitment to positive change. It could also serve to inspire future improvement and creative progress on a similar scale.

Nobody would dispute that music is integral to the American experience. Furthermore, few would doubt the objective existence of those who understand, analyze, and process music in a manner vastly superior to that of the larger general populace.²³⁴ Surely, those with bona fide qualified and quantifiable knowledge should be in charge of steering the course of music law over the next century and beyond. Speculation, flat-out guessing, celebrity-induced bias, and rampant indeterminacy have no place in the mathematically driven field of musical copyrights. Appropriately, neither does William Hung.

—Eric M. Leventhal

234. The author does not mean to question the importance or viability of the uniquely American and deeply ingrained process of jury trials, nor to somehow diminish the authority and prescience of French historian and political thinker Alexis de Tocqueville. Rather, in this one minor context, it is believed that the existing formulation is wrought with inaccuracy because it has been improperly constituted from its inception.



JAMAIL CENTER FOR LEGAL RESEARCH
TARLTON LAW LIBRARY
THE UNIVERSITY OF TEXAS SCHOOL OF LAW

The Tarlton Law Library Oral History Series features interviews with outstanding alumni and faculty of The University of Texas School of Law.

Oral History Series

- | | |
|---|--|
| No. 1 - <i>Joseph D. Jamail, Jr.</i> 2005. \$20 | No. 6 - <i>James DeAnda</i> 2006. \$20 |
| No. 2 - <i>Harry M. Reasoner</i> 2005. \$20 | No. 7 - <i>Russell J. Weintraub</i> 2007. \$20 |
| No. 3 - <i>Robert O. Dawson</i> 2006. \$20 | No. 8 - <i>Oscar H. Mauzy</i> 2007. \$20 |
| No. 4 - <i>J. Leon Lebowitz</i> 2006. \$20 | No. 9 - <i>Roy M. Mersky</i> 2008. \$25 |
| No. 5 - <i>Hans W. Baade</i> 2006. \$20 | |

Forthcoming:

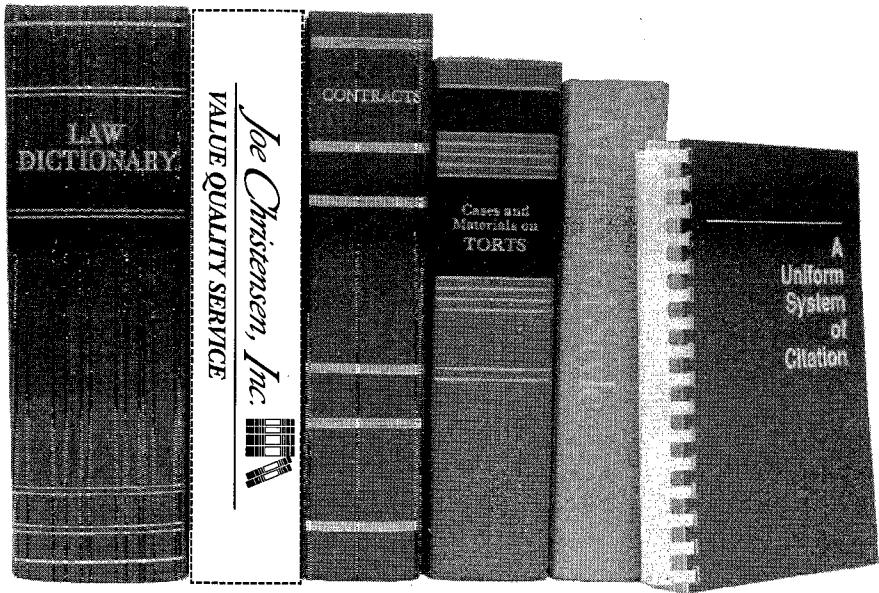
Gloria Bradford, Patrick Hazel, James W. McCartney,
Michael Sharlot, Ernest E. Smith, John F. Sutton, Jr.

*Other Oral Histories Published by the
Jamail Center for Legal Research*

- Robert W. Calvert* (Texas Supreme Court Trilogy, Vol. 1). 1998. \$20
Joe R. Greenhill, Sr. (Texas Supreme Court Trilogy, Vol. 2). 1998. \$20
Gus M. Hodges (Tarlton Law Library Legal History Series, No. 3). 2002. \$20
Corwin Johnson (Tarlton Law Library Legal History Series, No. 4). 2003. \$20
W. Page Keeton (Tarlton Legal Bibliography Series, No. 36). 1992. \$25
Jack Pope (Texas Supreme Court Trilogy, Vol. 3). 1998. \$20

Order online at <http://tarlton.law.utexas.edu/> click on Publications
or contact Publications Coordinator,
Tarlton Law Library, UT School of Law,
727 E. Dean Keeton St., Austin, TX 78705


phone (512) 471-6228; fax (512) 471-0243;
email tarltonbooks@law.utexas.edu



We Complete the Picture.

In 1932, Joe Christensen founded a company based on Value, Quality and Service. Joe Christensen, Inc. remains the most experienced Law Review printer in the country.

Our printing services bridge the gap between your editorial skills and the production of a high-quality publication. We ease the demands of your assignment by offering you the basis of our business—customer service.

Joe Christensen, Inc. 

1540 Adams Street
Lincoln, Nebraska 68521-1819
Phone: 1-800-228-5030
FAX: 402-476-3094
email: sales@christensen.com

Value

Quality

Service

Your Service Specialists

THE UNIVERSITY OF TEXAS SCHOOL OF LAW PUBLICATIONS

What the students print here changes the world

Journal	domestic/foreign
Texas Law Review http://www.TexasLRev.com	\$47.00 / \$55.00
Texas International Law Journal http://www.tilj.org	\$45.00 / \$50.00
Texas Environmental Law Journal http://www.texenrls.org/publications_journal.cfm	\$40.00 / \$50.00
American Journal of Criminal Law http://www.ajcl.org	\$30.00 / \$35.00
The Review of Litigation http://www.thereviewoflitigation.org	\$30.00 / \$35.00
Texas Journal of Women and the Law http://www.tjwl.org	\$40.00 / \$45.00
Texas Intellectual Property Law Journal http://www.tiplj.org	\$25.00 / \$30.00
Texas Hispanic Journal of Law & Policy http://www.thjlp.org	\$30.00 / \$40.00
Texas Journal On Civil Liberties & Civil Rights http://www.txjclcr.org	\$40.00 / \$50.00
Texas Review of Law & Politics http://www.trolp.org	\$30.00 / \$35.00
Texas Review of Entertainment & Sports Law http://www.tresl.net	\$40.00 / \$45.00
Texas Journal of Oil, Gas & Energy Law http://www.tjogel.org	\$30.00 / \$40.00
Manuals:	
<i>The Greenbook: Texas Rules of Form</i> 12th ed. ISBN 1-878674-08-0	
<i>Manual on Usage & Style</i> 11th ed. ISBN 1-878674-55-2	

To order, please contact:
 The University of Texas School of Law Publications
 727 E. Dean Keeton St.
 Austin, TX 78705 U.S.A.
 Publications@law.utexas.edu

ORDER ONLINE AT:
<http://www.texaslawpublications.com>

* * *

* * *

Texas Law Review

The Greenbook: Texas Rules of Form *Twelfth Edition*

A comprehensive guide for Texas citation, newly revised in 2010.

Texas Law Review Manual on Usage & Style *Twelfth Edition*

A pocket reference guide on style for all legal writing.

Newly revised and released in Fall 2011

School of Law Publications
University of Texas at Austin
727 East Dean Keeton Street
Austin, Texas USA 78705
Fax: (512) 471-6988 Tel: (512) 232-1149
Order online: <http://www.utexas.edu/law/publications>

