Texas Law Review

April 2010 Vol. 88 No. 5 Pages 917 to 1150

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 *Texas Rules of Form* (11th ed. 2006) and *The Bluebook: A Uniform System of Citation* (18th ed. 2005). Except when content suggests otherwise, the *Texas Law Review* follows the guidelines set forth in the *Texas Law Review Manual on Usage & Style* (11th ed. 2008), *The Chicago Manual of Style* (15th ed. 2003), and Bryan A. Garner, *A Dictionary of Modern Legal Usage* (2d ed. 1995).

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.texaslrev.com

THE UNIVERSITY OF TEXAS SCHOOL OF LAW

ADMINISTRATIVE OFFICERS

LAWRENCE G. SAGER, B.A., LL.B.; *Dean, John Jeffers Research Chair in Law, Alice Jane Drysdale Sheffield Regents Chair.*
MECHELE DICKERSON, B.A., J.D.; *Associate Dean for Academic Affairs, Arthur L. Moller Chair in Bankruptcy Law and Practice.*
ALEXANDRA W. ALBRIGHT, B.A., J.D.; *Associate Dean for Administrative Services, Senior Lecturer.*
EDEN E. HARRINGTON, B.A., J.D.; *Assistant Dean for Clinical Education, Director of William Wayne Justice Center for Public Interest Law.*
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.*
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.*
LESLIE OSTER, B.A., J.D.; *Assistant Dean for Strategic Planning.*
REYMUNDO 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.*
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 and 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 and 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.A., LL.M., J.S.D.; *Thomas Shelton Maxey Professor in Law.*
LYNN A. BAKER, B.A., Honours 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.*
BERNARD S. BLACK, A.B., M.A., J.D.; *Hayden W. Head Regents Chair for Faculty Excellence, Director of Center for Law, Business, and Economics.*
LYNN E. BLAIS, A.B., J.D.; *Leroy G. Denman, Jr. Regents Professor in Real Property Law.*
ROBERT G. BONE, B.A., J.D.; *Professor.*
NORMA V. CANTU, B.A., J.D.; *Professor of Law and Education.*
LOFTUS C. CARSON, II, B.S., M. Pub. Aff., M.B.A., J.D.; *Ronald D. Krist Professor.*
ROBERT M. CHESNEY, B.S., J.D.; *Charles I. Francis Professor in Law.*
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.*
JOHN DEIGH, B.A., M.A., Ph.D.; *Professor of Law and Philosophy.*
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 Professor in Law, Director of Bernard and Audre Rapoport Center for Human Rights and Justice.*
WILLIAM E. FORBATH, A.B., B.A., Ph.D., J.D.; *Lloyd M. Bentsen Chair in Law.*
JULIUS G. GETMAN, B.A., LL.B., LL.M.; *Earl E. Sheffield Regents Chair.*
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.*
PATRICIA I. HANSEN, A.B., M.P.A., J.D.; *J. Waddy Bullion Professor.*
HENRY T. HU, B.S., M.A., J.D.; *Allan Shivers Chair in the Law of Banking and Finance.*
DEREK P. JINKS, B.A., M.A., J.D.; *The Marrs McLean Professor in Law.*
STANLEY M. JOHANSON, B.S., LL.B., LL.M.; *Fannie Coplin Regents Chair, University Distinguished Teaching Professor.*
CALVIN H. JOHNSON, B.A., J.D.; *Andrews & Kurth Centennial Professor.*
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 and W. St. John Garwood, Jr. Centennial Chair in Law, Professor of Government.*
STEFANIE A. LINDQUIST, B.A., J.D., Ph.D.; *The Thomas W. Gregory Professorship in Law.*
BASIL S. MARKESINIS, 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. and Teresa Lozano Long Endowed Chair in Administrative Law.*
LINDA S. MULLENIX, B.A., M. Phil., J.D., Ph.D.; *Morris & Rita Atlas Chair in Advocacy.*

ROBERT J. PERONI, B.S.C., J.D., LL.M.; *James A. Elkins Centennial Chair in Law.*
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, The University of Texas at Austin, Hines H. Baker and Thelma Kelley Baker Chair, University Distinguished Teaching Professor.*
DAVID M. RABBAN, B.A., J.D.; *Dahr Jamail, Randall Hage Jamail, and Robert Lee Jamail Regents Chair, University Distinguished Teaching Professor.*
ALAN S. RAU, B.A., LL.B.; *Burg Family Professorship.*
R. A. REESE, B.A., J.D.; *Arnold, White & Durkee Centennial Professor.*
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.*
DANIEL B. RODRIGUEZ, B.A., J.D.; *Minerva House Drysdale Regents Chair in Law.*
WILLIAM M. SAGE, A.B., M.D., J.D.; *James R. Dougherty Chair for Faculty Excellence.*
JOHN J. SAMPSON, B.B.A., LL.B.; *William Benjamin Wynne Professor.*
THOMAS K. SEUNG, Ph.D.; *Jesse H. Jones Regents Professor in Liberal Arts.*
MICHAEL M. SHARLOT, B.A., LL.B.; *Wright C. Morrow Professor.*
CHARLES M. SILVER, B.A., M.A., J.D.; *Roy W. and Eugenia C. MacDonald Endowed Chair in Civil Procedure, Co-Director of Center on Lawyers, Civil Justice, and the Media.*
ERNEST E. SMITH, B.A., LL.B.; *Rex G. Baker Centennial Chair in Natural Resources Law.*
JANE STAPLETON, B.S., Ph.D., LL.B., 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.; *Stanley D. and Sandra J. Rosenberg Centennial Professor in Property Law.*
SHIRLEY E. THOMPSON, A.B., A.M., Ph.D.; *Professor.*
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 in Law.*
PATRICK WOOLLEY, A.B., J.D.; *Beck, Redden & Secrest Professor.*

ASSISTANT PROFESSORS

OREN BRACHA, LL.B., S.J.D.
DANIEL M. BRINKS, A.B., J.D., Ph.D.
JENS C. DAMMANN, J.D., LL.M., Dr. Jur., J.S.D.
JUSTIN DRIVER, B.A., M.A., M.A., J.D.
ZACHARY S. ELKINS, B.A., M.A., Ph.D.
MIRA GANOR, B.A., M.B.A., LL.B., LL.M., J.S.D.
JOHN M. GOLDEN, A.B., J.D., Ph.D.
EMILY E. KADENS, B.A., M.A., Dipl., M.A., Ph.D., J.D.
JENNIFER E. LAURIN, B.A., J.D.
ANGELA K. LITTWIN, B.A., J.D.
MARY ROSE, A.B., M.A., Ph.D.
SEAN H. WILLIAMS, B.A., J.D.
HANNAH J. WISEMAN, A.B., J.D.

SENIOR LECTURERS, WRITING LECTURERS, AND CLINICAL PROFESSORS

WILLIAM P. ALLISON, B.A., J.D.; *Clinical Professor, Director of Criminal Defense Clinic.*
MARJORIE I. BACHMAN, B.S., J.D.; *Clinical Instructor.*
PHILIP C. BOBBITT, A.B., J.D., Ph.D.; *Distinguished Senior Lecturer.*
KAMELA S. BRIDGES, B.A., B.J., J.D.; *Lecturer.*
CYNTHIA L. BRYANT, B.A., J.D.; *Clinical Professor, Director of Mediation Clinic.*
SARAH M. BUEL, B.A., J.D.; *Clinical Professor.*
MARY R. CROUTER, A.B., J.D.; *Assistant Director of William Wayne Justice Center for Public Interest Law.*
TIFFANY J. DOWLING, B.A., J.D.; *Clinical Instructor.*
LORI K. DUKE, B.A., J.D.; *Clinical Professor.*
ARIEL E. DULITZKY, J.D., LL.M.; *Clinical Professor.*
DENISE L. GILMAN, B.A., J.D.; *Clinical Professor.*
BARBARA HINES, B.A., J.D.; *Director of Immigration Clinic.*
KRISTINE A. HUSKEY, B.A., J.D.; *Director of National Security & Human Rights Clinic.*
JEANA A. LUNGWITZ, B.A., J.D.; *Clinical Professor, Director of Domestic Violence Clinic.*
ROBIN B. MEYER, B.A., M.A., J.D.; *Lecturer.*
TRACY W. MCCORMACK, B.A., J.D.; *Director of Trial Advocacy Program.*
RANJANA NATARAJAN, B.A., J.D.; *Clinical Professor.*
ROBERT C. OWEN, A.B., M.A., J.D.; *Clinical Professor, Co-Director of Capital Punishment Clinic.*
SEAN J. PETRIE, B.A., J.D.; *Lecturer.*
WAYNE SCHIESS, B.A., J.D.; *Senior Lecturer, Director of Legal Writing.*
PAMELA J. SIGMAN, B.A., J.D.; *Lecturer, Director of Juvenile Justice Clinic.*
LESLIE L. STRAUCH, B.A., J.D.; *Clinical Professor.*
DAVID S. SOKOLOW, B.A., M.A., J.D., M.B.A.; *Distinguished Senior Lecturer, Director of Student Life.*
JAN SUMMER, B.A., M.A., J.D.; *Executive Director of Center for Public Policy Dispute Resolution.*
MELINDA E. TAYLOR, B.A., J.D.; *Senior Lecturer, Executive Director of Center for Global Energy, International Arbitration, and Environmental Law.*
HEATHER K. WAY, B.A., B.J., J.D.; *Lecturer, Director of Community Development Clinic.*
ELIZABETH M. YOUNGDALE, B.A., M.L.I.S., J.D.; *Lecturer.*

ADJUNCT PROFESSORS AND OTHER LECTURERS

WILLIAM R. ALLENSWORTH, B.A., J.D.
JAMAL K. ALSAFFAR, B.A., J.D.
MARILYN ARMOUR, B.A., M.S.W., Ph.D.
WILLIAM G. BARBER, B.S.Ch.E., J.D.
WILLIAM G. BARBER, III, B.A., LL.M.
NICOLAS G. BARZOUKAS, B.S., M.B.A., J.D.
SHARON C. BAXTER, B.S., J.D.
KATHERINE E. BEAUMONT, B.A., J.D.
KARA BELEW, B.A., B.B.A., 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.
MURFF F. BLEDSOE, B.A., J.D.
WILLIAM P. BOWERS, B.B.A., J.D., LL.M.
ANTHONY W. BROWN, B.A., J.D.
JAMES E. BROWN, LL.B.
TOMMY L. BROYLES, B.A., J.D.
PAUL J. BURKA, B.A., LL.B.
W. A. BURTON, JR., B.A., M.A., LL.B.
AGNES E. CASAS, B.A., J.D.
RUBEN V. CASTANEDA, B.A., J.D.
EDWARD A. CAVAZOS, B.A., J.D.
CHARLES G. CHILDRESS, B.A., J.D.
DAN J. CHRISTENSEN, B.B.A., J.D.
JOSEPH A. CIALONE, II, B.A., J.D.
JEFF CIVINS, A.B., M.S., J.D.
JUDGE LEIF M. CLARK, B.A., M.Div., J.D.
DANA L. COBB, B.A., J.D.
GARY COBB, B.A., J.D.
MARK W. COCHRAN, B.J., LL.M., J.D.
ELIZABETH COHEN, B.A., M.S.W., J.D.
JAMES W. COLLINS, B.S., J.D.
TED CRUZ, A.B., J.D.
PATRICIA J. CUMMINGS, B.A., J.D.
WILLIAM H. CUNNINGHAM, B.A., M.B.A., Ph.D.
HECTOR DE LEON, B.S., J.D.
DICK DEGUERIN, B.A., LL.B.
MICHELE Y. DEITCH, B.A., M.S., J.D.
ADAM R. DELL, B.A., J.D.
CASEY D. DUNCAN, B.A., M.L.I.S., J.D.
PHILIP DURST, B.A., M.A., J.D., Ph.D.
ELANA S. EINHORN, B.A., J.D.
JAY D. ELLWANGER, B.A., J.D.
LOWELL P. FELDMAN, B.A., J.D.
KENNETH FLAMM, Ph.D.
JOHN C. FLEMING, B.A., J.D.
MARIA FRANKLIN, B.A., M.A., Ph.D.
DAVID C. FREDERICK, B.A., Ph.D., J.D.
GREGORY D. FREED, B.A., J.D.
ELIZABETH FRUMKIN, B.A., M.A., J.D.
FRED J. FUCHS, B.A., J.D.
MICHAEL GAGARIN, Ph.D.
GERALD J. GALOW, B.A., J.D.
JAMES B. GAMBRELL, B.S.M.E., M.A., LL.B.
FRANCIS J. GAVIN, B.A., M.S.T., Ph.D.
CHARLES E. GHOLZ, B.S., Ph.D.
MICHAEL J. GOLDEN, A.B., J.D.
JULIE E. GRANTHAM, B.A., J.D.
SHERRI R. GREENBERG, B.A., M.Sc.
ROBERT L. GROVE, B.S., J.D.
DAVID HALPERN, B.A., J.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.
AMY J. SCHUMACHER, B.A., J.D.
AARON R. SCHWARTZ, B.A., LL.B.
MARCUS F. SCHWARTZ, B.B.A., J.D.
SUZANNE SCHWARTZ, B.J., J.D.
ROBERT L. HARGETT, B.B.A., J.D.
BARBARA J. HARLOW, B.A., M.A., Ph.D.
JAMES C. HARRINGTON, B.A., M.A., J.D.
CHRISTOPHER S. HARRISON, Ph.D., J.D.
AMBER L. HATFIELD, B.S.E.E., J.D.
JOHN R. HAYS, JR., B.A., J.D.
PAUL M. HEBERT, A.B., J.D.
STEVEN L. HIGHLANDER, B.A., Ph.D., J.D.
SUSAN J. HIGHTOWER, B.A., M.A., J.D.
WILLIAM M. HINES, III, B.B.A., J.D.
JAMES C. HO, B.A., J.D.
KENNETH E. HOUP, JR., J.D.
RANDY R. HOWRY, B.J., J.D.
MONTY G. HUMBLE, B.A., J.D.
BOBBY R. INMAN, B.A.
PATRICK O. KEEL, B.A., J.D.
CHARI L. KELLY, B.A., J.D.
ROBERT N. KEPPLE, B.A., J.D.
MARK L. KINCAID, B.B.A., J.D.
KEVIN S. KUDLAC, B.S., 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.
MAURIE A. LEVIN, B.A., J.D.
ANDREW F. MACRAE, B.J., J.D.
VIJAY MAHAJAN, M.S.Ch.E., Ph.D.
JIM MARCUS, B.A., J.D.
PETER D. MARKETOS, B.A., J.D.
FRANCES L. MARTINEZ, B.A., J.D.
RAY MARTINEZ, III, B.A., J.D.
ALOYSIUS P. MARTINICH, B.A., M.A., Ph.D.
LISA M. MCCLAIN, B.A., J.D., LL.M.
STEPHANIE G. MCFARLAND, B.A., J.D.
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.
DARYL L. MOORE, B.A., M.L.A., J.D.
STEVEN A. MOORE, B.A., Ph.D.
EDWIN G. MORRIS, B.S., J.D.
MARCO M. MUNOZ, LL.B., M.C.J.
MANUEL H. NEWBURGER, B.A., J.D.
STEVEN P. NICHOLS, B.S.M.E., M.S.M.E., J.D., Ph.D.
JANE A. O'CONNELL, B.A., J.D.
PATRICK L. O'DANIEL, B.B.A., J.D.
GUILLERMO A. PADILLA, J.D., Ph.D.
M. A. PAYAN, B.A., J.D.
MARK L. PERLMUTTER, B.S., J.D.
LOUIS T. PIRKEY, B.S.Ch.E., J.D.
JUDGE ROBERT L. PITMAN, B.S., 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.
GRETCHEN RITTER, B.S., Ph.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.
WILLIAM F. STUTTS, B.A., J.D.
MATTHEW J. SULLIVAN, B.S., J.D.
GRETCHEN S. SWEEN, B.A., M.A., Ph.D., J.D.
BRADLEY P. TEMPLE, B.A., J.D.

RICHARD J. SEGURA, JR., 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.
LOUISE SINGLE, B.S., M.T.A., Ph.D.
BEA A. SMITH, B.A., M.A., J.D.
CORY W. SMITH, B.A., J.D.
DWAYNE W. SMITH, B.A., J.D.
TARA A. SMITH, Ph.D.
LYDIA N. SOLIZ, B.B.A., J.D.
JUSTICE ROSE B. SPECTOR, B.A., J.D.
LEWIS J. SPELLMAN, B.B.A., M.B.A., M.A., Ph.D.
DAVID B. SPENCE, B.A., J.D., M.A., Ph.D.
WILLIAM J. SPENCER, B.A., M.S., Ph.D.
LAURA L. STEIN, B.A., M.A., Ph.D.
JAMES B. STEINBERG, B.A., J.D.
PAUL J. STEKLER, Ph.D.
CHARLES H. STILL, B.B.A., J.D.
SHERINE E. THOMAS, B.A., J.D.
TERRY O. TOTTENHAM, B.S., LL.M., 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.
VALERIE L. TYLER, B.J., J.D.
SUSAN S. VANCE, B.B.A., J.D.
LANA K. VARNEY, B.J., J.D.
CLARK C. WATTS, B.A., M.D., M.A., M.S., J.D.
JANE M. WEBRE, B.A., J.D.
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.
JUDGE PAUL L. WOMACK, B.S., J.D.
NOLAN L. WRIGHT, B.A., M.A., M.L.I.S., J.D.
LARRY F. YORK, B.B.A., LL.B.
DANIEL J. YOUNG, B.A., J.D.

VISITING PROFESSORS

ANTONIO H. BENJAMIN, LL.B., LL.M.
PETER F. CANE, B.A., LL.B., D.C.L.
DAVID ENOCH, LL.B., Ph.D.
IAN P. FARRELL, B.A., LL.B., M.A., LL.M.
VICTOR FERRERES, J.D., LL.M., J.S.D.
FRANCESCO FRANCIONI, J.D., LL.M.

* * *

Texas Law Review

Volume 88 Number 5 April 2010

* * *

Texas Law Review

Volume 88, Number 5, April 2010

* * *

Texas Law Review

Volume 88, Number 5, April 2010

Articles

Torts as Wrongs

John C.P. Goldberg* & Benjamin C. Zipursky**

I. Introduction

All of the standard substantive first-year law courses seem to address a basic legal category. All, that is, save one. Property is about the relationship of persons to things that can be owned and alienated—land, chattels, and patents, for example. Criminal Law, at its core, concerns rules so important that their violation elicits from the state its harshest action: punishment. Contract Law introduces students to the ways in which law can empower individuals to enter into mutually advantageous transactions. Civil Procedure provides students with an overview of the litigation process. Constitutional Law is about guarding the guardians. Each of these subjects stands out for being ancestral, essential, or both.

The odd man out, it seems, is Torts. As it tends to be taught today, Torts is "accident-law-plus." Its most noted chestnuts involve claims for negligence or strict liability.[1] Accidents—in the sense of unintended outcomes—are even at the center of the most commonly taught intentional tort cases.[2] The "plus" comes from decisions that serve as a platform for

---

* Professor, Harvard Law School.

** Professor and James H. Quinn '49 Chair in Legal Ethics, Fordham University School of Law. For helpful comments on prior drafts, we are grateful to Glenn Cohen, Ronald Dworkin, Thomas Nagel, Arthur Ripstein, Tony Sebok, Henry Smith, and Jeremy Waldron, as well as to participants in the New York University School of Law's Colloquium on Social, Legal, and Political Philosophy; the Queens University Law Faculty's Law, Economics and Public Policy Workshop; and the University of Pennsylvania Law School's Institute for Law and Philosophy Workshop. Remaining errors are our own.

1. *E.g.*, United States v. Carroll Towing Co., 159 F.2d 169 (2d Cir. 1947); Sindell v. Abbott Labs, 607 P.2d 924 (Cal. 1980); Dillon v. Legg, 441 P.2d 912 (Cal. 1968); Summers v. Tice, 199 P.2d 1 (Cal. 1948); Escola v. Coca Cola Bottling Co., 150 P.2d 436 (Cal. 1944); Brown v. Kendall, 60 Mass. (6 Cush.) 292 (1850); Palsgraf v. Long Island R.R., 162 N.E. 99 (N.Y. 1928); Rylands v. Fletcher, (1868) 3 L.R.E. & I. App. 330 (appeal taken from Eng.) (U.K.).

2. *E.g.*, Garrat v. Dailey, 279 P.2d 1091 (Wash. 1955); Vosburg v. Putney, 50 N.W. 403 (Wis. 1891).

discussions of economic or moral theory.[3] In sum, Torts seems often to be conceived as a course that teaches students how common law allocates the costs of accidents, while also providing some instruction on law and economics, or law and philosophy. So defined, the course seems ad hoc and esoteric, not basic. Somehow law professors have lost their grip on its subject matter.

The goal of this Article is to put us back on track, not just pedagogically but theoretically. Tort is indeed a basic category of law. To see this, however, one must abandon the notion, now deeply entrenched, that tort law is law for allocating the costs of accidents. As its name indicates, tort law is about *wrongs*.[4] The law of torts is a law of wrongs and recourse—what Blackstone called "private wrongs."[5]

Of course tort law is in many ways public. It sets generally applicable standards of conduct.[6] It is developed and applied by officials who may have in mind various policy concerns as they render judgments in particular cases. And its operation can advance or interfere with the operation of other public institutions. But tort is private in two basic senses. It defines duties to refrain from injuring (or to protect from injury) that are owed by certain persons to others: duties that, when breached, constitute *wrongs to those others*, as opposed to wrongs to the world.[7] Second, precisely because torts are private wrongs, they provide the basis for a private response.[8] For a wrong to be a tort it must in principle generate for its victim *a private right of action*: a right to seek recourse through official channels against the wrongdoer.

As the law of private and privately redressable wrongs, tort law is rightly treated as a cornerstone of legal education along with criminal law (the law of public and publicly redressable wrongs) and contract law (the law

---

3. *See, e.g.*, Boomer v. Atl. Cement Co., 257 N.E.2d 870 (N.Y. 1970) (inviting discussion of when entitlements should be protected by injunction rather than liability); *see also* Vincent v. Lake Erie Transp. Co., 124 N.W. 221, 221–22 (Minn. 1910) (raising the issues of whether and why it might be just or efficient to impose liability for injurious acts taken out of "necessity").

4. The noun "tort" means "wrong." *See* BLACK'S LAW DICTIONARY 1626 (9th ed. 2009) (defining "tort" as a civil wrong, other than breach of contract, for which a remedy may be obtained).

5. *See* 3 WILLIAM BLACKSTONE, COMMENTARIES *2 (describing "private wrongs" as "an infringement or privation of the private or civil rights belonging to individuals"); *id.* at *115–19 (treating causes of action for infringing the rights of persons or property as articulating private wrongs for which the law provides a remedy to victims).

6. *See* MacPherson v. Buick Motor Co., 111 N.E. 1050, 1053 (N.Y. 1915) (observing that the duties recognized by negligence laws are duties of conduct grounded in law rather than defined by agreement).

7. *See, e.g.*, Palsgraf v. Long Island R.R., 162 N.E. 99, 101 (N.Y. 1928) ("Negligence, like risk, is thus a term of relation. Negligence in the abstract, apart from things related, is surely not a tort, if indeed it is understandable at all. Negligence is not a tort unless it results in the commission of a wrong, and the commission of a wrong imports the violation of a right . . . .").

8. *See* Benjamin C. Zipursky, *Rights, Wrongs, and Recourse in the Law of Torts*, 51 VAND. L. REV. 1, 85 (1998) (explaining the centrality to tort of the private right of action).

of consensually defined duties). Looked at through the lens of litigation, Torts is about the wrongs that a private litigant must establish to entitle her to a court's assistance in obtaining a remedy and the remedies that will be made available to her. Looked at through the lens of daily life, Torts is about which duties of noninjury owed to others are counted as legal duties and what sorts of remedial obligations one will incur for failing to conduct oneself in accordance with those duties.[9] In turn, the places to look for contemporary extensions of tort law are *not* the compensation systems with which tort law is frequently coupled.[10] Rather, they are found in the rules governing 10b-5 suits, civil RICO actions, Title VII claims for workplace discrimination, constitutional tort claims, and intellectual-property-infringement actions. To study torts is to learn what sort of conduct our legal system defines as wrongfully injurious toward another such that, when committed, the victim is entitled to exact something from the wrongdoer. This is the domain of law that was born centuries ago with the recognition of the writ of *trespass vi et armis* and that today is defined by state and federal common law, as well as state and federal statutory and constitutional law.

How is it that academics have lost their feel for this basic legal category? The dominant tendency among modern scholars has been to dismiss the language of wrongs as dated, squishy, and inapt, and to dismiss as bankrupt any distinction between "public" and "private."[11] Although we see no reason to suppose that talk of "wrongs" is inherently less meaningful than talk of "cognitive biases" or "marginal utility," we grant that it is essential for torts to be understood as *legal wrongs* rather than moral wrongs. In addition, while we conceive of torts as private wrongs, we also concede that government is central to the tort system's operation in a manner that many scholars have overlooked and that a challenge for tort theory is to explain what is distinctively "private" about tort, given the state's role. In short, there is a need for a cogent and doctrinally grounded account of two distinct concepts and the connection between them: *tortious wrongdoing* ("wrongs") and *civil recourse* ("recourse"). This Article begins to fill that need.

Part II describes the move in torts scholarship away from the idea of Torts as a law of private wrongs in favor of a conception of Torts as law for the allocation of accidentally caused losses. In doing so, it demonstrates the pervasiveness of this mistake, which cuts across standard divides in tort

---

9. Needless to say, the wrongs of tort include negligence and other wrongs that can be committed unintentionally. RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL & EMOTIONAL HARM §§ 3, 20–23 (Proposed Final Draft No. 1, 2005). In this sense, accidents are central to tort law—not because tort is a law of accidents, but because many instances of accidental injury constitute wrongs for which victims are entitled to redress.

10. The section of the American Association of Law Schools devoted to tort law, established in 1972, is titled the "Section on Torts and Compensation Systems." *See generally* The Association of American Law Schools, Section on Torts and Compensation Systems, https://memberaccess.aals.org/eWeb/dynamicpage.aspx?webcode=ChpDetail.

11. *See infra* Part III.

theory. Part III posits that tort scholars have placed losses at the center of tort theory out of jurisprudential, moral, political, and conceptual concerns that together create a challenge for those who wish to retain a tort theory centered on wrongs, and it sets forth that challenge. This challenge is met in Part IV, in which we set forth a framework for understanding torts as instances of a distinctive kind of wrong.

Having established the availability of a notion of Torts as wrongs, Part V argues for the descriptive superiority of a wrongs-based view: basic features of tort law are inexplicable from within a view of Torts as law for allocating losses, yet are perfectly intelligible if one understands the subject of Torts as covering a special kind of wrong. Part VI explains that there is value to having law that defines private wrongs and provides recourse to victims of those wrongs, and that this value does not reduce down to other values such as enhancing safety, compensating persons in need, or achieving justice. Part VII briefly identifies ways in which a wrongs-and-recourse approach to tort law can illuminate contemporary and enduring debates within and about tort law while also providing an agenda for further research.

II. Torts, Losses, and Accidents

Most tort scholars would accede to the hornbook definition of torts as "civil wrong[s], other than breach[es] of contract, for which the court will provide a remedy."[12] At the least, they probably would not deny that the word "tort" means "wrong," that tort law is on the civil side of the civil–criminal divide, and that torts—while different from breaches of contract—involve individuals bringing lawsuits seeking damages or other forms of relief.

Yet it is equally commonplace for scholars to insist that concepts such as "wrong," "legal wrong," "civil wrong," and "private wrong" do no real work in explaining how the field hangs together, or what its point is.[13] These scholars are also wont to suggest that we would do better to focus instead on certain allegedly observable features of tort law—i.e., when it tends to be invoked and what tends to happen when it is invoked.[14] Doing so, they suppose, allows us to see that tort law hangs together as law for the allocation of costs, especially the costs of accidents.

---

12. W. PAGE KEETON ET AL., PROSSER AND KEETON ON THE LAW OF TORTS § 1, at 2 (5th ed., Lawyer's ed. 1984); *see also, e.g.*, WILLIAM L. PROSSER, HANDBOOK OF THE LAW OF TORTS § 1, at 1 (1941) ("'Tort' is a term applied to a miscellaneous . . . group of civil wrongs . . . for which a court of law will afford a remedy in the form of an action for damages.").

13. *See, e.g.*, PROSSER, *supra* note 12, § 1, at 3–4 (recognizing the great difficulties scholars have faced in trying to give a coherent definition of the field of tort law).

14. *See, e.g.*, *id.* § 1, at 8 ("Enough has been said to indicate that definition or description of a tort in terms of generalities distinguishing it from other branches of the law is difficult, or impossible. It is somewhat easier to consider the function and purpose of the law of torts.").

Holmes's writings on torts are commonly taken to mark the birth of modern tort theory.[15] That they are so regarded is itself a significant fact. Holmes's writings are accorded this status *because* they provide perhaps the first effort to offer a comprehensive account of the field in terms of losses and accidents.[16] This effort was itself part of Holmes's larger thesis that the transition from ancient and medieval law to modern law was marked by the law's de-moralization.[17]

Determined to get behind or underneath the "moral phraseology . . . of wrongs" that abounds in tort law,[18] Holmes quite consciously chose in *The Common Law* to commence his treatment with accidents.[19] He did so because it permitted him to confront his readers with a view that he believed most of them held, namely, the view that liability under the old trespass writ was *strict*—without regard to fault.[20] How can the moral phraseology be taken at face value, he argued, if trespass liability has always been strict?[21] Of course, Holmes was at the same time determined to rebut a particular version of the claim that trespass liability was strict and to establish instead the historical pedigree and normative superiority of fault-based liability, properly understood.[22] In effect, he credited proponents of the strict-liability claim with having overstated an otherwise sound point. Liability in tort law for accidentally caused harms had always been fault based, not strict. But the

---

15. *See, e.g.*, Thomas C. Grey, *Accidental Torts*, 54 VAND. L. REV. 1225, 1256 (2001) (describing Holmes's early writings as the "first serious attempt in the common law world to give torts both a coherent structure and a distinctive substantive domain").

16. *See, e.g., id.* at 1282 ("[Holmes's] approach centered tort doctrine around . . . accidental personal injuries.").

17. *See* OLIVER WENDELL HOLMES, JR., THE COMMON LAW 33 (Mark DeWolfe Howe ed., Harv. Univ. Press 1963) (1881) (arguing that the law "is continually transmuting . . . moral standards into external or objective ones").

18. *Id.* at 65.

19. *See id.* (noting that trespass "affords a fair field for a discussion of the general principles of liability for unintentional wrongs").

20. *See id.* (stating that the historically more popular view of liability was that "a man is answerable for all the consequences of his acts").

21. *Id.* at 65–66.

22. *See id.* at 76 (arguing that the general principle of the common law is that losses from genuine accidents lie where they fall). Here we reject David Rosenberg's provocative argument that Holmes sought only to resist "absolute" forms of strict liability but endorsed strict liability for acts causing losses that could have been foreseen and thus avoided. *See* DAVID ROSENBERG, THE HIDDEN HOLMES: HIS THEORY OF TORTS IN HISTORY 5–6 (1995) (concluding that Holmes's "theory held that rules of strict liability qualified by a foresight condition . . . were just and rational"). Holmes did believe that the propriety of imposing liability in tort turns on the foreseeability and the avoidability of the plaintiff's loss. *See* HOLMES, *supra* note 17, at 45–46 ("[T]here must be actual present knowledge of the present facts which make an act dangerous . . . [and the act] must be made with a chance of contemplating the consequence complained of . . . ."). Rosenberg is mistaken, however, to suppose that these aspects of Holmes's theory stood apart from his commitment to the idea that tort liability turns on a failure of the defendant to meet the law's standard for prudent conduct. *See, e.g.*, HOLMES, *supra* note 17, at 97 (endorsing a nonsuit granted in a slip-and-fall case on the ground that the defendant "had done all that it was bound to do in maintaining [its] staircase").

standard for fault was *legal* rather than *moral* fault—i.e., objective rather than subjective.[23] In Holmes's view, tort law was thus fault based in the sense of rejecting strict liability, but was not fault based in the sense of conditioning liability on the commission of a morally wrongful act.[24]

In this manner, Holmes decoupled tort law from notions of wrong and wrongdoing. Yet he did not draw from this effort a skeptical conclusion about the integrity of the field. Quite the opposite, he famously was an early convert on the question of whether Torts is a "proper subject."[25] What, then, did he suppose the subject matter of tort law to be? The answer: law for the allocation of losses. "The business of the law of torts is to fix the dividing lines between those cases in which a man is liable for harm which he has done, and those in which he is not."[26] Tort law is law that specifies when a liberal state will depart from its default rule of nonintervention and force *A* to indemnify *B*. Generally speaking, it is prepared to do so for any instance in which (1) *A* causes *B* a loss and (2) such a loss was a foreseeable consequence of *A*'s actions, thus rendering it avoidable in principle through the exercise of what the law defines as ordinary prudence.[27]

Now fast-forward to the mid-twentieth century and the hugely influential Prosser treatise.[28] It offers a view of Torts not very different from Holmes's. Prosser begins with the standard definition of a tort as a civil wrong (other than breach of contract), only to reject it as unhelpful.[29] In its place, he hesitantly offers that "[t]he common thread woven into all torts is the idea of unreasonable interference with the interests of others."[30] It is

---

23. *See* HOLMES, *supra* note 17, at 88 ("What the law really forbids, and the only thing it forbids, is the act on the wrong side of the line, be that act blameworthy or otherwise.").

24. Holmes next addressed torts such as deceit, which seem not to be about accidents, and to be all about "actual wickedness." *Id.* at 104. Retracing Lecture I's argument for a de-moralized understanding of criminal law, he again insisted that appearances are misleading and that the default rule of tort liability is action causing harm under circumstances that would alert an ordinary person to the risk of harm, in turn permitting him to avoid causing harm through the exercise of prudence. *Id.* at 104–05. For example, liability for deceit rests on the defendant's having uttered a statement in circumstances in which there was a foreseeable likelihood that it might prove to be false and might induce someone to rely on it to his detriment. *Id.* at 109.

25. Grey, *supra* note 15, at 1232.

26. HOLMES, *supra* note 17, at 64; *see also* Commonwealth v. Pierce, 138 Mass. 165, 176 (1884) (Holmes, J.) (characterizing civil liability for negligence as "the redistribution of losses" and contrasting it with criminal law, which sets limits on conduct "in the interest of the safety of all").

27. *See* HOLMES, *supra* note 17, at 115 ("[T]he general purpose of the law of torts is to secure a man indemnity against certain forms of harm to person, reputation, or estate, at the hands of his neighbors, not because they are wrong, but because they are harms . . . . [Fault-based liability] is intended to reconcile the policy of letting accidents lie where they fall, and the reasonable freedom of others with the protection of the individual from injury."). Of course in Holmes's view, the fact that a particular defendant is incapable of exercising ordinary prudence, and hence incapable of taking advantage of the opportunity provided by the state to avoid liability, is not a ground for excusing him from the duty to indemnify. *Id.*

28. PROSSER, *supra* note 12.

29. *Id.* § 1, at 3.

30. *Id.* at 8.

notable that Prosser here superimposes onto the general concept of a tort a distinctive feature of the particular tort of nuisance, in which the adjective "unreasonable" is *not* used to describe the alleged tortfeasor's conduct but instead the type or degree of loss caused by that conduct.[31] As in nuisance, so in tort law as a whole, Prosser seems to say, the question is not whether an actor has failed to comply with a norm of conduct, but whether his actions have produced a sufficiently serious adverse impact on another. Prosser is also anxious to note that almost anything can count as unreasonable interference—the test being whether it is unreasonable "from the point of view of the community as a whole, rather than the sole matter of individually questionable conduct."[32]

Having established that the "wrongs" of tort consist of the (nearly limitless) set of acts that might be deemed to generate unreasonable interferences with others, Prosser suggests that we will do better to define tort law by reference to its function. Distinguishing a tort from a crime, he observes that "the civil action for a tort . . . is commenced and maintained by the injured person himself, and its purpose is to compensate him for the damage he has suffered at the expense of the wrongdoer."[33] Prosser adds that, "[i]n recent years, there has been a growing appreciation that the law of torts is concerned chiefly with the distribution of the losses inevitable in a civilized community, in accordance with the court's conception of social justice."[34] The thought seems to be this: one can expect that, in a crowded, industrialized world, injuries will happen and that they will generate losses. Tort law determines—on the basis of judges' and jurors' sense of what counts as an unreasonable interference—whether to let those losses lie where they fall or to shift them to others.

By the mid-1960s, with scholarly attention squarely on automobile accidents, the linkage of the idea of Torts as law for loss allocation to the idea of tort law as accident law was transformed from an already strong tendency to an axiom. This much is clear in three important works from this period. First is John Fleming's *An Introduction to the Law of Torts*,[35] which opens with this passage:

> [T]he economic costs of accidents represents a constant and mounting drain on the community's human and material resources. The task of the law of torts is to play an important regulatory role in the adjustment of these losses and the eventual allocation of their cost.[36]

---

31. *Id.*
32. *Id.* at 10.
33. *Id.* § 2, at 10.
34. *Id.* at 11.
35. JOHN G. FLEMING, AN INTRODUCTION TO THE LAW OF TORTS (1967).
36. *Id.* at 1.

No less clear on this point is Patrick Atiyah's *Accidents, Compensation and the Law*.[37] Tort law, he says, is best viewed as but one method by which the state deals with "the problem of compensation for misfortunes."[38] He adds that the subject matter of tort law is primarily the rules by which compensation is paid (or not paid) for "road accidents and industrial accidents."[39]

Finally, although pathbreaking in its reliance on economic analysis in the service of a deterrence-based argument for certain forms of liability, Judge Calabresi's *The Costs of Accidents* fits quite comfortably within the tradition we are describing.[40] His premise, like Atiyah's, is that tort law is one of several means by which government might address the problem of accident costs.[41] Also like Atiyah, Calabresi concludes that tort law—particularly negligence law—is not well-designed to achieve accident-cost reduction.[42] For civil litigation to reduce the cost of accidents, it must allocate losses in a way that properly incentivizes those who are in the best position to take cost-efficient precautions against such losses.[43] For this task, a regime of strict liability—and really a form of liability disconnected from backward-looking inquiries into causation and responsibility—would be vastly preferable. In challenging the propriety of the law's setting fault as the trigger of liability, Calabresi, like Fleming James Jr. before him,[44] only further weakened the idea that the right to prevail in a tort suit has any principled connection to the notion of a wrong having been done. Tort law is about shifting losses to achieve policy objectives, not wrongs and recourse.

The work of each of the aforementioned scholars accepts that, when it comes to understanding tort law, the path to enlightenment begins by dissociating the concept of a tort from notions of wrong and wrongdoing. To say that torts are civil wrongs other than those arising from contract is to utter an unhelpful platitude. It tells us nothing about why individuals in the various instances called "torts" are entitled to damages. In reality, tort liability always begins with a loss, the cost of which the plaintiff aims to shift, and

37. P.S. ATIYAH, ACCIDENTS, COMPENSATION AND THE LAW (3d ed. 1980).

38. *Id.* at 6.

39. *Id.* at 239; *see also id.* at 59 (describing negligence law as a scheme for "decid[ing] if compensation should be paid to an innocent accident victim").

40. GUIDO CALABRESI, THE COSTS OF ACCIDENTS: A LEGAL AND ECONOMIC ANALYSIS (1970).

41. *See id.* at 239 (asserting that the "the fault system" is properly assessed in terms of how well it performs as a scheme for minimizing the costs of accidents).

42. For one thing, it asks lay jurors to focus on highly particularized facts rather than conditions that generally obtain and to decide where liability should fall within an artificially constrained universe of potential loss-bearers. *Id.* at 246–49. Worse, it shifts back and forth between the goals of deterrence and loss spreading in a way that undermines its ability to achieve either, while also generating large administrative costs. *Id.* at 274–77.

43. *See id.* at 312 (suggesting that accident costs can be reduced by assigning costs to those that can avoid accidents most cheaply).

44. *See* John C.P. Goldberg, Comment, *Misconduct, Misfortune, and Justice Compensation: Weinstein on Torts*, 97 COLUM. L. REV. 2034, 2045–47 (1997) (summarizing James's view that tort law has come primarily to serve an insurance function).

hence tort law must be understood as law devoted to identifying the instances in which losses should be shifted.

Of course the enterprise of tort theory did not grind to a halt with the publication of *The Cost of Accidents*. Indeed, it was precisely at this time that the academy witnessed a resurgence of efforts to analyze law by reference to notions of rights and justice. Did not this resurgence correspond with a revitalization of the idea of torts as wrongs? For the most part, no. Fairness- and justice-oriented scholars tended to be concerned with "public law" issues of civil rights and distributive justice. This sort of thinking dovetailed quite naturally with loss-based conceptions of Torts. Illustrative is George Fletcher's influential 1972 effort to craft a fairness-based conception of Torts.[45] Fletcher argued that tort law calls on judges to shift losses in accordance with a principle of reciprocity.[46] A critical piece of his argument is that the defendant's having committed a legal wrong against the plaintiff is irrelevant to the fairness of shifting it. Hence, one of Fletcher's main goals was to defend certain forms of strict liability.[47]

What about corrective-justice theory? Many would suppose it to be precisely the school of thought that has reestablished the centrality of wrongs to Torts. Among the leaders of this school is Jules Coleman, whose most important book to date in tort theory is titled *Risks and Wrongs*.[48] And a central idea in his work and that of other corrective-justice theorists is that tort law holds defendants responsible for injuries they have caused others through wrongful conduct.[49] Yet even in the hands of Coleman, wrongs turn out to be less central than they first seem. Much the same is true, we believe, for the work of other corrective-justice theorists, including the early work of Richard Epstein, as well as that of Stephen Perry and perhaps Arthur Ripstein.[50] In fact, among corrective-justice theorists, Ernest Weinrib has distinguished himself for having unequivocally embraced the idea of torts as wrongs.[51]

---

45. George P. Fletcher, *Fairness and Utility in Tort Theory*, 85 HARV. L. REV. 537 (1972).

46. *See id.* at 550 ("All individuals in society have the right to roughly the same degree of risk.").

47. *See id.* at 544–49 (arguing that fairness requires the imposition of strict liability for activities that impose nonreciprocal risks of loss).

48. JULES L. COLEMAN, RISKS AND WRONGS (1992).

49. *See id.* at 329 ("There are two essential components in the concept of corrective justice . . . wrongfulness and responsibility.").

50. *See, e.g.*, ARTHUR RIPSTEIN, EQUALITY, RESPONSIBILITY, AND THE LAW 48–58 (1999) (explaining the concept of fault by examining the "fair" allocation of risk); Richard A. Epstein, *A Theory of Strict Liability*, 2 J. LEGAL STUD. 151, 151 (1973) (attempting to reconcile tort theory with "common sense notions of individual responsibility" by exploring concepts of causation); Stephen R. Perry, *The Moral Foundations of Tort Law*, 77 IOWA L. REV. 449, 450 (1992) (identifying outcome-responsibility and fault as the conditions set by tort law for the recognition of a duty of repair).

51. *See* ERNEST J. WEINRIB, THE IDEA OF PRIVATE LAW 134–35 (1995) (claiming that a tortious act is "an act of wrongdoing").

Coleman himself goes out of his way to insist that tort law is fundamentally about *losses*, not wrongs.[52] In his mind, tort law distinguishes itself from criminal law precisely on this score. If law is going to respond to wrongs *qua* wrongs, he says, it should be in the business of punishment.[53] Tort law, by contrast, shifts losses.[54] It is true that, for Coleman, as for Perry, Epstein, and Ripstein, the determination of when a loss is to be shifted hinges on the identification of grounds for holding the defendant morally responsible. Still, the type of responsibility that generates tort liability is the *moral responsibility for a loss* one has caused, rather than responsibility for having committed a wrong, a point that Stephen Perry makes clear by invoking Honoré's notion of "outcome-responsibility" and distinguishing it from responsibility for actions.[55] Instantiating the principle of corrective justice, tort law specifies that an actor's having caused a loss to another, and having done so by means of conduct that falls short of an applicable moral standard, is a sufficient reason for deeming the loss to be the defendant's moral responsibility and not the plaintiff's. The fundamental question to which the principle of corrective justice provides an answer is thus: Whose mess is it?[56]

Among broadly influential approaches to modern tort theory, the one that may sit least comfortably within the loss-allocation framework is Judge Posner's efficient-deterrence account.[57] Yet even this view is fairly depicted as an allocative one, in which judges, at the behest of private litigants, shift the costs of certain losses or harms in order to promote the efficient expenditure of resources on accident prevention. That Posner's view does not present itself as a loss-allocation theory stems partly from the fact that it attributes no inherent value to providing compensation to an injured plaintiff; that a loss has fallen on one person rather than anyone else is a distributional issue according to Posner, and, as such, is irrelevant to the question of whether resources are being used efficiently.[58] And yet the related question of when an accident cost should be assigned to a defendant is fundamental.

---

52. *See* COLEMAN, *supra* note 48, at 330–32 (arguing that tort law is concerned to correct wrongful losses, not wrongs per se).

53. *See id.* at 222–24 (distinguishing risk allocation by tort law and punishment by criminal law).

54. *See id.* at 314–18 (criticizing "relational" views that treat tort law as responding to wrongs, as opposed to losses); Perry, *supra* note 50, at 486–87 (criticizing Weinrib's characterization of torts as wrongs).

55. Perry, *supra* note 50, at 489–96. *See generally* Tony Honoré, *Responsibility and Luck*, 104 L.Q. REV. 530 (1988) (creating the concept).

56. *See* Jules L. Coleman, *Second Thoughts and Other First Impressions*, *in* ANALYZING LAW 257, 302 (Brian Bix ed., 1998) ("Tort law is about messes. A mess has been made, and the only question before the court is, who is to clean it up?").

57. *See generally* WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW (1987) (arguing that tort law is best understood as a means of inducing cost-effective precaution taking).

58. *See id.* at 16–17 (arguing that tort law operates to achieve Kaldor–Hicks efficiency).

The challenge for courts is to figure out when, given the goal of efficient deterrence, an accident victim's costs should be imposed on someone else.[59]

It is also true that, in Posner's view, the trigger for loss allocation is the defendant's having engaged in conduct that is "wrongful."[60] Yet Posner insists that torts are wrongs only in the sense that they involve a failure to use scarce resources efficiently.[61] Given this account of what makes conduct tortious, as well as Posner's efforts to disconnect the "wrong" of waste from ordinary conceptions of wrongdoing,[62] one may question whether Posner's account, in the end, really is about wrongs. Regardless, it is quite clearly not about *private wrongs*. As both Weinrib and Coleman have emphasized,[63] Posner is concerned with conduct that is (on the most charitable reading) wrongful to the world—a misuse of the resources in principle available to us all—which is why, on his view, tort suits must be understood as enforcement actions brought by plaintiffs acting in the capacity of private attorneys general.

***

Loss-allocation views of tort come in many shapes and sizes. Some assert that a defendant is morally responsible for a harm he has caused and therefore can be held liable for the costs associated with it. Some say that it is fairer for a defendant who has caused a loss to bear it, and that is when and why tort law reallocates a loss. Some focus on the better capacity of certain risk creators to pay for and to insure against such costs. Some say that loss shifting will permit efficient deterrence. Perhaps the most notable difference is not about *which values* are realized through tort law's shifting of losses but *how* values are realized through its operation. Progressives and Posnerians alike view tort law as carrying out one or more of several public goals, be they egalitarian, libertarian, insurance providing, or efficiency enhancing.

Not all of these allocative views depict or need to depict tort law as coextensive with accident law. For several reasons, however, contemporary tort theories that see Torts as loss-allocation law tend also to see it as law for allocating losses arising out of accidents. First, the most notable expansions of tort liability for losses in the modern era have occurred in part because of

---

59. *See id.* at 6–8 (tracing the roots of positive economic analysis of tort law to the work of Coase and Calabresi, and arguing that judicial decisions in fact impose liability when doing so furthers the goal of efficient deterrence).

60. *See* Richard A. Posner, *A Theory of Negligence*, 1 J. LEGAL STUD. 29, 32 (1972) (acknowledging that harm caused through negligence arouses indignation).

61. *See id.* at 31–32 (suggesting torts are "wrongs" in the particular sense of being failures to take cost-justified precautions).

62. *Id.*

63. *See* JULES L. COLEMAN, THE PRACTICE OF PRINCIPLE: IN DEFENCE OF A PRAGMATIST APPROACH TO LEGAL THEORY 13–24 (2001) (criticizing economic analysis of tort law for its inability to capture adequately the structure of tort law); WEINRIB, *supra* note 51, at 48 (criticizing economic analysis for presuming the plaintiff's suit is "not to secure redress for wrongful injury but to claim a bounty for prosecuting inefficient economic activity").

the increase in accidental injuries.[64] Second, one's perception of the importance of an area of law, and the source of that importance, is related to what one thinks it does. If tort law is fundamentally about permitting persons to shift the costs of their injuries to others, then those areas in which there is a pronounced need to have the costs of injuries shifted will figure especially prominently in thinking about what the law is.[65] Third, when the fundamental question in Torts is posed as the question of when to shift losses, instances of intentional infliction of physical harm seem trivially easy and thus really require no attention: virtually everything interesting in the subject area involves cases in which a loss is caused accidentally. Finally, when one focuses upon losses caused by accidents, one tends to focus upon a certain feature of conduct, namely, the creation of *risks* of physical harm. In a contemporary legal culture understandably eager to avoid moralism in describing legal wrongs, the idea that we should focus on the degree to which individuals can impose risks on others has been highly appealing to all of these theorists. But once risk creation is the analytical lens through which one assesses conduct, one is thinking very much in terms of Torts as the realm of accident law.

III. Torts as Wrongs: Four Challenges

Torts scholars did not always think of tort law as fundamentally concerned with accidents or with loss allocation. Blackstone plainly conceived of Torts as the law of private wrongs,[66] as did influential nineteenth-century American jurists.[67] A handful of important modern theorists—most notably Professor Weinrib—offer wrongs-based views of Torts. Robert Stevens holds a view that is closer to being wrongs based than loss based; perhaps the same can be said of John Gardner and Arthur Ripstein's recent work.[68] Overwhelmingly, however, allocative models

64. *See* John Fabian Witt, *Toward a New History of American Accident Law: Classical Tort Law and the Cooperative First-Party Insurance Movement*, 114 HARV. L. REV. 690, 694 (2001) (asserting that the common law of torts expanded in the late nineteenth century as an alternative institutional mechanism for dealing with "an accident crisis like none the world had ever seen," stemming largely from industrial accidents).

65. *See, e.g.*, LANDES & POSNER, *supra* note 57, at 2–3 (suggesting that tort law was an unimportant field until the rise of litigation over railroad accidents).

66. *See supra* note 5.

67. *See* John C.P. Goldberg, *Ten Half-Truths About Tort Law*, 42 VAL. U. L. REV. 1221, 1237 & n.40 (2008) [hereinafter Goldberg, *Ten Half-Truths*] (noting that early U.S. treatises by Hilliard and Addison defined torts, respectively, as "Private Wrongs" and "Wrongs"); John C.P. Goldberg, *Two Conceptions of Tort Damages: Fair v. Full Compensation*, 55 DEPAUL L. REV. 435, 444–45, 456–59 (2006) [hereinafter Goldberg, *Two Conceptions*] (discussing the adoption of a Blackstonian view of torts by Nathan Dane, Zephaniah Swift, and Simon Greenleaf).

68. *See* ROBERT STEVENS, TORTS AND RIGHTS 2–3 (2007) (arguing that the "rights model" of tort law, which analyzes torts as a "species of wrong," is clearly preferable over the "loss model" for torts actionable per se). Gardner clearly takes wrongs to be central to tort law. *See* John Gardner, *What Is Tort Law for? Part 1: The Place of Corrective Justice* 19 (Oxford Legal Stud. Research Paper Series, Paper No. 1, 2010), *available at* http://ssrn.com/abstract=1538342 (describing tort law

dominate tort theory, and the idea of torts as wrongs is mentioned only as a matter of form or etymology.

The fading of an idea is sometimes warranted: it is good that scientists no longer talk of phlogiston. Yet it is not as if the general idea of a wrong has proved itself to be conceptually bankrupt, nor has our law generally dissociated itself from notions of wrongdoing. If anything, the last half-century has seen in the area of criminal law a rebirth of theories of punishment dependent on moral concepts and a concomitant rejection of therapeutic understandings of criminals as not-to-be-blamed.[69] Likewise, the last great wave of enthusiasm in tort law and tort theory for broad forms of "enterprise liability" is now almost twenty years behind us,[70] and there is no question that noninstrumentalist tort theory has gained significant ground.[71]

So we are left with a mystery. With hindsight, one can see how an obsession with accidents prompted mid-twentieth-century jurists to emphasize tort law's potential as a source of compensation while deemphasizing its foundation in a notion of wrongs. But great resistance to wrongs-based theories of tort law remains. Why?

Our diagnosis is that a cluster of powerful jurisprudential, moral, and political ideas have exerted—and continue to exert—tremendous force on American legal thought, and have done so in a way that has made the notion of Torts as private wrongs appear unavailable. These ideas do not attach to any one tort theorist or philosopher and do not fit easily into any pre-set pattern. Thus, it will be useful to instead examine in the abstract four kinds of problems that seem to have led theorists to abandon the notion of torts as wrongs.

The first problem concerns what might be called the *normative* aspect of the notion of a wrong, and it presents a range of different problems that come

---

as requiring those who commit legal wrongs to pay reparative damages "in respect of" those losses occasioned by their wrongs). Yet he is also inclined to describe tort law as implementing corrective justice and to maintain that corrective justice is a matter of allocation. *Id.* at 59–60. Ripstein's book *Equality, Responsibility, and the Law* is fairly read as being at least equivocal about how central a role losses should play in tort theory, for the concept of risk-ownership carries substantial weight in that work, and it surely resonates deeply with the loss-orientation of Coleman's work from the late 1980s and early to mid-1990s. *Compare* RIPSTEIN, *supra* note 50, at 53–58 (discussing how "the person who exposes another to a risk 'owns' the risk, and if the risk ripens into an injury, that person owns the injury"), *with* COLEMAN, *supra* note 48, at 253 ("[I]f a victim can show that her loss is wrongful in the appropriate sense, the burden of making good her loss falls to the individual responsible for it."). However, Ripstein's work over the past several years has become increasingly wrongs based. *See* Arthur Ripstein, *As If It Had Never Happened*, 48 WM. & MARY L. REV. 1957, 1960 (2007) (linking legal remedies to wrongs).

69. *See, e.g.*, James Q. Whitman, *A Plea Against Retributivism*, 7 BUFF. CRIM. L. REV. 85, 87 (2003) (noting, critically, the renewed emphasis in criminal law theory on retributive rather than therapeutic conceptions of criminal law).

70. *See generally* 1 REPORTERS' STUDY, ENTERPRISE LIABILITY FOR PERSONAL INJURY (1991). The study "never received the imprimatur of the ALI, and the project was subsequently abandoned by the ALI Council." Jerry J. Phillips, *Comments on the Reporters' Study of Enterprise Responsibility for Personal Injury*, 30 SAN DIEGO L. REV. 241, 241 (1993).

71. *See supra* notes 45–56 and accompanying text.

to a head in a dilemma: Should we understand torts as *moral* wrongs or as *legal* wrongs? For sound doctrinal reasons, tort theorists have been disinclined to cast torts as moral wrongs. For a different set of jurisprudential reasons, they have instead treated torts as legal wrongs. Yet in doing so, they have felt compelled to concede that this choice necessarily drains the normative aspect of the idea of a wrong from torts, leaving only an empty conceptual shell. Let us call this "The Moral–Legal Dilemma."

The second problem arises from the fact that torts generate a specific kind of response from the legal system: the imposition of liability when a private plaintiff successfully asserts a claim against the defendant. At a theoretical level, many tort scholars have been troubled by the idea of a tort as a wrong because they would have expected the legal system to respond very differently to torts if they were really a kind of wrong. Specifically, they would have expected that it would respond to this sort of wrong by punishing the wrongdoer. Instead, tort law "responds" by shifting the cost of the loss imposed upon the plaintiff back to the defendant. Both the absence of a punishment and the presence of a financial liability for a loss render the notion of a tort as a legal wrong unhelpful at best and incoherent at worst, the argument goes. We refer to this as "The Inaptness-of-Liability Problem."

The third problem that has pushed theorists away from wrongs-based conceptions of Torts derives from the fact that torts are defined in an *injury-inclusive* manner. A defendant is not considered to have committed a tort unless his potentially injurious conduct has actually ripened into an injury. From the point of view of dominant ways of thinking about morality, however, the idea of an *injury-inclusive wrong* makes no sense—*acts* are the proper subject of moral evaluation, not *consequences*. So if torts cannot be deemed wrongs without being treated as injury-inclusive wrongs, then we must abandon the notion of torts as wrongs. We label this "The Realization Problem."

Finally, scholars have supposed that, even if coherent, the idea of private wrongs is unhelpful. The notion that Torts can be organized as a field around the concept of a wrong is no more useful—and hence no less ripe for abandonment—than the medieval practice of organizing Torts around the writs of trespass and case. What we end up with is an essentially random catalogue of the types of acts that, if they cause certain types of harms, subject an actor to liability. We call this "The Hodgepodge Problem."

A. The Moral–Legal Dilemma

Killing an innocent person is both a moral wrong and a legal wrong, while chewing gum is neither. But the two categories are by no means coextensive. Failing to save a drowning child, though it can easily be done, is a moral wrong but usually not a legal wrong. Using the land of another person, while reasonably believing it to be one's own land, is a legal wrong but not a moral wrong. For very different reasons, Holmes, Prosser, Calabresi, and Posner have emphasized this point. Paradoxically, it turns out

that an appreciation of the distinctively *legal* aspect of tortious wrongdoing is both the beginning of all wisdom and a source of profound confusion—the basis for grasping the particular sense in which tort law is a law of wrongs and the basis for failing to do so such that one is led down a false path toward loss-allocation theories.

The capacity for the phrase "legal wrong" to cause mischief derives in part from the implied suggestion that perhaps there is no moral (or immoral) aspect at all to the term "wrong" within that phrase. Simply put, many find it odd to suppose that a moral wrong is simply one kind of wrong among several, and yet the replacement of "moral" by "legal" in the phrase "legal wrong" tends to suggest just that. To say of a certain pair of earrings that they are gold, but not yellow gold, allows for the possibility that they are made of white gold, which is really a kind of gold. By contrast, to say of them that they are gold, but contain no precious metal, is to say that they are not really gold earrings. To the modern lawyerly ear, a description of torts as wrongs, but not moral wrongs, is akin to describing earrings as gold, but without precious metal. The supposition is that the phrase "legal wrongs" functions like the phrase "fool's gold." The adjective "legal" purports to identify the kind of wrong at issue but ends up signifying that torts are not genuine wrongs at all and instead bear only a superficial resemblance to them.

How did this come to be a dominant instinct among tort theorists? We conjecture that the answer rests partly in the influence of a positivistic conception of law traceable back to Holmes and ultimately Austin. According to this conception, legal wrongs are those acts that violate a command or dictate issued by a political superior or sovereign.[72] It is only when the sovereign issues the sort of command that specifies for its subjects what is wrong to do or not do that the category of legal wrongs is created.

Significantly, this model of law much more comfortably accommodates legislation issued within a parliamentary system than common law.[73] By the same token, it renders awkward the idea of common law torts as wrongs. When the careless acts of an imprudent person[74] or a car manufacturer[75] are deemed "wrongful," they are deemed so notwithstanding the absence of a readily located *ex ante* command that has been violated. Instead,

---

72. *See generally* JOHN AUSTIN, THE PROVINCE OF JURISPRUDENCE DETERMINED 11–13 (Hackett Publ'g Co. 1998) (1832) (defining laws to include only rules passed "by persons exercising supreme and subordinate *government*" and placing norms set only by public opinion in a separate category of "positive morality").

73. Of course, Austin himself thought that his jurisprudence could account for the common law. *See* 1 JOHN AUSTIN, LECTURES ON JURISPRUDENCE, OR, THE PHILOSOPHY OF POSITIVE LAW 102 (Robert Campbell ed., 5th ed. 1885) ("Now when judges transmute a custom into a legal rule (or make a legal rule not suggested by custom), the legal rule which they establish is established by the sovereign legislature. A subordinate or subject judge is merely a minister. The portion of the sovereign power which lies at his disposition is merely delegated.").

74. Vaughan v. Menlove, (1837) 132 Eng. Rep. 490 (C.P.) (U.K.).

75. MacPherson v. Buick Motor Co., 111 N.E. 1050 (N.Y. 1915).

"wrongdoing" is determined by the application of a judicially recognized norm that incorporates a standard of ordinary prudence that itself is not traceable to a clear sovereign directive. And if careless conduct does not violate a government-issued prohibition, the only meaningful sense in which it can be a wrong is if it amounts to the violation of a *moral* (rather than a legal) prohibition against it. In this way, the Holmes–Austin notion that law's normativity—its directive force—resides in the issuance of a command by a political authority tends to eliminate the possibility that legal wrongs are distinctive, at least for a common law area such as Torts. In turn, it pushes toward the idea that there are really only two options: either Torts is a law of wrongs only in the sense that it happens to attach official sanctions to the commission of conduct that is wrongful in the sense of morally wrongful, or Torts is not really about wrongs at all.

Some who accept this dilemma have considered it so important to depict torts as wrongs that they have tried to equate torts with moral wrongs. But too many well-settled doctrines stand in the way of this move. There are several torts, such as trespass to land, that deem an actor to be a tortfeasor notwithstanding that he has not acted immorally—indeed, has acted reasonably and blamelessly.[76] Even in negligence, the breach standard is judged from an objective point of view, and this drives a wedge between those actors one would deem to have acted immorally and those who can be held liable for negligently injuring someone. Thus, the effort to see torts as wrongs seems to present only two options, both of which seem unacceptable: Either equate torts with moral wrongs, which runs afoul of settled doctrine, or categorize torts as legal wrongs, which is doubly unacceptable because there are no commands of which tortious conduct can be violative and because there is no independent content to the concept of wrongfulness—whatever government prohibits is by definition wrongful. A proper understanding of the subject of Torts must therefore eschew reference to wrongfulness and instead define torts as not-necessarily-culpable failures to take opportunities to avoid harming others (Holmes); or as conduct that can be deemed to interfere unreasonably with the interests of others (Prosser); or as conduct that wastes resources (Posner).

B. The Inaptness-of-Liability Problem

Our effort to reclaim the notion of wrongs for tort law has been motivated in part by the work of leading corrective-justice theorist Ernest Weinrib. Weinrib, however, has been criticized by Jules Coleman and Stephen Perry, who have expressly argued that losses, not wrongs, are more basic to corrective justice and to tort law.[77] Their fundamental objection is

76. *See, e.g.*, RESTATEMENT (SECOND) OF TORTS § 164 (1965) (stating that an intentional physical invasion of another's land is a trespass even if made in the reasonably mistaken belief that the land was not owned by the other).

77. *See supra* notes 48–52 and accompanying text.

that the legal system has a particular response to torts, and it is quite different than it would be if torts were really a kind of wrong.[78] We shall look at Coleman's account of what the state's response *should be* if torts are wrongs, turning then to Perry's account of what the response to a tort *is* and why a conception of torts as wrongs is ill-suited to explaining this.

Coleman is, at least in principle, sympathetic with the idea that government might think it a matter of justice to respond to a wrong that one person has done to another. Invoking language from some of his earliest attempts to articulate a notion of corrective justice, he argues that it might want to respond by "annulling" the wrong.[79] However, he insists that the business of annulling wrongs is a matter of retributive justice, not corrective justice: "There is a legal institution that, in some accounts anyway, is designed to do retributive justice, namely, punishment."[80] On this basis, he rejects the possibility that tort law identifies and responds to wrongs *qua* wrongs.

Coleman's argument can be depicted as relying upon three premises: (a) if torts are wrongs then tort liability will have to be understood as—in some sense—the legal system's effort to do justice in responding to wrongs; (b) the imposition of tort liability for wrongs is not the imposition of retributive punishment for conduct deemed worthy of criminal punishment; (c) the imposition of retributive punishment for conduct deemed worthy of criminal punishment is the way the legal system tries to do justice in responding to wrongs.[81] We agree with the first two premises. The third has plausibility but is ambiguous. It is quite plausible that punishment is in some sense about responding to wrongs. The argument is not complete, however, unless the third premise states that retributive punishment is the legal system's *only* (systematic) way of doing justice in responding to wrongs. Coleman does not explicitly make this claim; his point is more that punishment is the only well-developed legal institution that has been understood as responding to wrongs (as opposed to losses)—and that tort liability is plainly not punishment. Thus interpreted, this part of Coleman's argument poses a challenge to a wrongs-based theory of tort law that has not been met: How is tort liability a response to wrongs, given that it is not punishment? Because he sees no good answer to this challenge, yet he is able to provide an answer when tort law is understood as responding to losses, Coleman opts for the latter approach.

Perry, like Coleman, also offers a more particular critique of Weinrib's effort to depict tort liability as a response to torts as legal wrongs. Tort liability, on Perry's view, involves first and foremost the imposition of a duty

78. *See supra* notes 53–54.

79. COLEMAN, *supra* note 48, at 325.

80. *Id.*

81. *See supra* text accompanying notes 49–53.

of repair upon a tortfeasor, a duty that runs to the injured plaintiff.[82] Weinrib argues that the duty of repair flows from the defendant's having committed a wrong.[83] The wrong is depicted as the breach of a primary duty of conduct owed to the victim; the duty of repair is depicted as a secondary duty that the defendant incurs upon breaching the primary duty.[84] The problem, according to Perry, is that there is a nonsequitur here: A duty of *repair* rectifies a loss, not a wrong.[85] Wrongs are defined by Weinrib in a manner that abstracts from interests and are quite detached from the notion of a loss, assuming a loss is understood as a setback to interests.[86] Hence, there is simply no basis for understanding the assignment of responsibility for a loss (which is what the duty of repair involves) as flowing from the commission of a wrong.

In short, for both Coleman and Perry, there is a fundamental mismatch between a body of substantive law that purports to regulate conduct on the ground that it is wrongful and a body of remedial law that typically requires the payment of compensatory damages. If the law of torts were truly a law of wrongs, they reason, its characteristic remedy would not be an award of damages keyed to the *losses* caused by a wrong. Instead, it would be a penalty keyed to the gravity of the wrong itself. A genuine law of wrongs must go hand in hand with a law of punishment, not a law of damages.

C. The Realization Problem

Careless driving that harms no one is in a sense negligent, but it does not amount to the commission of the tort of negligence. Likewise, the writing of a defamatory statement that remains unread is not a libel. In each case there is conduct that might ripen into an injury, but the conduct is not tortious because the ripening never occurs.

Generally speaking, there are two ways of thinking about how conduct and injury come together in the definition of each tort. The first is to treat the conduct component of a tort as the component in which the wrongfulness of the tortfeasor's conduct resides. Conversely, the injury component, though essential to liability, is not part of what makes a tort wrongful. On this view, a tort consists of an act that, taken on its own, meets the legal test for wrongfulness, and that also happens to cause a certain kind of result. To the extent the commission of a tort deserves condemnation, it deserves the same condemnation as does identical conduct that does not harm anyone. Two identically careless drivers, only one of whom hits someone, have committed

82. *See* Stephen R. Perry, *On the Relationship Between Corrective and Distributive Justice*, *in* OXFORD ESSAYS IN JURISPRUDENCE 237, 237–38 (Jeremy Horder ed., 4th ed. 2000) (characterizing corrective justice, which is concerned with the moral duty of repair, as the foundation of tort law).

83. *Id.* at 479–80.

84. *See id.* at 479 ("Weinrib's version [of the volitionist argument for fault liability] claims that accompanying the *primary* duty not to act wrongfully . . . is a *secondary* duty to compensate for harm that results from one's wrongful conduct.").

85. *Id.* at 480, 483.

86. *Id.* at 484.

the same wrong—the wrong of careless driving. That the legal system authorizes different responses to misconduct causing injury and misconduct not causing injury does not reflect a notion that one merits a different kind of response. Instead, it stems from the fact that, in the vast run of cases, there is simply nothing for tort law to do until an injury and a loss has occurred. We call this the *pure-conduct conception of wrongs*.

The alternative view holds that the wrongfulness of torts resides in conduct and result together. Torts, in other words, are defined by the law as wrongs that are only wrongs when completed or realized. Until the injury pregnant in an actor's misconduct occurs, there is no wrong in the tort sense of "wrong," though there might be grounds for condemning the actor or for subjecting him to sanction via criminal or regulatory law. This is because the duties imposed by tort law are duties of *noninjury*.[87] For example, the duty of care in negligence, on this view, is not correctly described as a duty to act with ordinary care toward others. It is instead a *duty not to injure others by conduct that is careless as to them*. By definition, this duty cannot be breached until an actor causes injury to another by failing to act with ordinary prudence toward that other. We call this *the injury-inclusive conception of wrongs*.

Overwhelmingly, modern tort theorists have assumed or insisted that the idea of an injury-inclusive wrong is incoherent.[88] This inclination derives from many sources. The influence of a command-based view of law can once again be detected. In thinking of legal wrongs as sovereign commands, it is quite natural to think of commands that enjoin *conduct* rather than consequences—"Drive carefully!"; "Follow standard medical protocol!"—and that carry with them a sanction for disobedience. The idea is that law identifies for subjects when they have rendered themselves eligible for sanction: They may be lucky enough to avoid a penalty, but the point at which they commit an act meeting the legal definition of a tort is the point at which they have ceded their ability to control this aspect of their fates.[89]

Those inclined to treat torts as moral wrongs to which the law has attached penalties may also have other grounds for being attracted to conduct-based (or will-based) conceptions of wrongdoing. Utilitarianism and Kantianism are the two leading contemporary frameworks for thinking

---

87. *See* Arthur Ripstein & Benjamin C. Zipursky, *Corrective Justice in an Age of Mass Torts*, *in* PHILOSOPHY AND THE LAW OF TORTS 214, 222–23 (Gerald J. Postema ed., 2001) (explaining that a tort plaintiff seeks recourse for the breach of a duty of noninjury).

88. *See, e.g.*, *id.* at 221–22 (outlining the views of scholars who argue that holding defendants responsible only when they breach a duty of noninjury is morally arbitrary).

89. Such was Holmes's view:

> All the rules that the law can lay down beforehand are rules for determining the conduct which will be followed by liability if it is followed by harm,—that is, the conduct which a man pursues at his peril . . . . [I]f he escapes liability, it is simply because by good fortune no harm comes of his conduct in the particular event.

HOLMES, *supra* note 17, at 64.

about moral wrongs. Rightly or wrongly, they have generally been interpreted by legal scholars in such a way as to render opaque the idea of a moral wrong that is not a wrong absent a certain kind of result. For Utilitarians, it is said, the wrongfulness of conduct hinges on the probability that the conduct will produce net disutility (more pain than pleasure). If the ascription of "negligence" to someone's driving means that it is the sort of driving that, on balance, will probably cause more pain than pleasure, it makes sense to condemn the driving as wrongful in and of itself, irrespective of its actual results. For Kantians, the moral quality of someone's conduct turns on the quality of the will of the person who performs it. If a person disregards others' deservingness of respect and acts in a way that is inconsistent with such respect—by, for example, not taking seriously the risks to others' health stemming from one's conduct—then the person has acted immorally. Whether the conduct is wrong is again independent of what happens as a result of the conduct.

On any of the foregoing views, the wrongfulness of tortious conduct must reside in the qualities of the actor's will or her acts, not in the results for the person suing. It makes no sense, on these views, to talk about torts as injury-inclusive wrongs. Every tort involves two qualitatively different components: wrongful conduct and certain results, morally neutral in themselves, that flow from that conduct.

D. The Hodgepodge Problem

Even if the various problems already identified can be overcome, there is a remaining concern associated with the idea of conceptualizing torts as wrongs. It is best captured in a deficiency commonly attributed to the pre-nineteenth-century system under which the various tort causes of action were organized under the writs of trespass and trespass on the case. This practice was deemed unsatisfactory because it seemed to be committed to the incoherent project of organizing a set of substantively defined wrongs by reference to the characteristic remedy available to those who successfully complained about their having been committed.[90] This concern seems to have been the source of Holmes's early disparagement of tort law as a non-subject: "The worst objection to the title Torts, perhaps, is that it puts the cart before the horse, that legal liabilities are arranged with reference to the forms of action allowed by the common-law for infringing them,—the substantive under the adjective law."[91]

Conceived of as wrongs, the various torts seem not to have any common characteristics. Some involve intentional injurings, others concern accidents.

---

90. *See* G. Edward White, *The Intellectual Origins of Torts in America*, 86 YALE L.J. 671, 678–83 (1977) (explaining the factors contributing to the demise of the writ system, including dissatisfaction with its emphasis on remedies and procedure over substantive wrongs).

91. Oliver W. Holmes, Jr., *The Theory of Torts*, 7 AM. L. REV. 652, 659 (1873), *reprinted in* 1 THE COLLECTED WORKS OF JUSTICE HOLMES 326, 331 (Sheldon M. Novick ed., 1995).

Some arise out of bodily harms, others out of property damage, and still others out of interferences with intangibles such as the interest in maintaining one's good name or using and enjoying one's property. Given the array of conduct and consequences recognized by the law as "tortious," it would seem to be the mere fact of actionability that unites the field: The only thing that one can say of all torts is that their commission generates the potential for liability. Because there is no conceptual integrity to the idea of tortious wrongdoing, it is unhelpful to think of torts as wrongs.

IV. Meeting the Challenges: Torts as Wrongs

Having identified and elaborated four significant challenges to the intelligibility or usefulness of the idea of torts as wrongs, we will rebut each of them. In doing so, it will help to address each in the reverse order of its initial presentation.

A. Tort Law's Wrongs

The wrongs of tort law are diverse. Some involve atrocious misconduct, others do not. In some instances, immoral actions that harm give rise to no liability at all. The sorts of injuries that give rise to tort claims range from fatal physical injuries to interferences with intangible property rights. If one were to rank conduct from the most to the least wrongful, and to rank injuries from the most to the least serious, tort law would not map onto these indices in any straightforward way, such that torts correspond to the most serious forms of misconduct and injury.

And yet, as we have explained in a recent book, the wrongs recognized by tort law hardly make for an eccentric or random collection.[92] Tortious wrongdoing always involves an interference with one of a set of individual interests that are significant enough aspects of a person's well-being to warrant the imposition of a duty on others not to interfere with the interest in certain ways, notwithstanding the liberty restriction inherent in such a duty imposition.[93] In part out of a sense of the limitations as to what sorts of interferences and injuries are justiciable, and in part for policy considerations that have changed over time with changes in social norms and economic and political circumstances, courts and legislatures have never sought to render interferences with all such interests actionable. (There is no tort for interference with one's ability to obtain a good education or a decently well-paying job.) Instead, it is to say that tortious wrongdoing is always about interference with some such interests and that tort law can hardly be accused of idiosyncrasy in focusing on the set of interests on which it does focus. It is also to say that tort law does not vindicate public or communal interests,

92. JOHN C.P. GOLDBERG & BENJAMIN C. ZIPURSKY, THE OXFORD INTRODUCTIONS TO U.S. LAW: TORTS ch. 3 (forthcoming 2010).

93. *See id.* (discussing tort law's "gallery of wrongs").

though of course those might be served indirectly by its operation. A brothel might well be a blight on a neighborhood, and an obstruction of a public way might create a big headache for those to whom it would otherwise provide a convenient route.[94] But both are classic examples of public nuisances: They are not wrongs to anyone in particular and are appropriately addressed through executive-branch action.

Unsurprisingly, many torts are concerned to protect and vindicate an individual's interest in her bodily integrity. Thus, it is prima facie actionable for one person to intentionally touch another's body in a manner that is harmful or offensive (battery),[95] to act carelessly toward them so as to cause them physical injury (negligence),[96] or to place a dangerously defective product in the stream of commerce so as to injure someone (products liability).[97] For centuries, it has been actionable to invade physically another's land (trespass),[98] to steal or intentionally destroy property (conversion),[99] to temporarily deprive or carelessly damage or destroy property (trespass to chattel, negligence),[100] and to create an unreasonable interference with another's use and enjoyment of her real property (nuisance).[101] Though statutory in origin, acts that infringe another's copyright, patent, or trademark today form an increasingly important class of legal wrongs that generate private rights of action.[102] As such, they might well be denominated "property torts."[103]

Although tort law often is concerned to address conduct that causes physical harms or property damage, it is a mistake to suppose that these forms of injury have a special claim to being central to the subject of torts. (Indeed, it is a serious defect of courses that teach tort law as "accident-law-plus" that they tend to convey the misimpression that physical harm and property damage are privileged in this way.) Even battery is not precisely concerned with physical harm. Rather, it is the wrong of invading another person's personal space by an unwelcome or inappropriate touching,

94. *See* KEETON ET AL., *supra* note 12, § 90, at 643–46 (discussing examples of public nuisance and the available remedies).

95. RESTATEMENT (SECOND) OF TORTS § 13 (1965).

96. *Id.* § 281.

97. RESTATEMENT (THIRD) OF TORTS: PRODS. LIAB. § 1 (1998).

98. RESTATEMENT (SECOND) OF TORTS § 158; *see also id.* cmts. j, *l* (citing examples of trespass to land from the 1800s).

99. *Id.* § 222A; *see also id.* cmt. d, illus. 1 (citing case examples of the tort of conversion from the 1800s).

100. *Id.* §§ 217–218; *see also id.* § 218, cmts. d, e (citing case examples from the 1800s).

101. RESTATEMENT (SECOND) OF TORTS § 821D (1979); *see also id.* cmt. f (citing case examples of nuisance from the 1800s).

102. *See* GOLDBERG & ZIPURSKY, *supra* note 92 (discussing intellectual property torts).

103. *See* RESTATEMENT (SECOND) OF TORTS § 871 (defining a property tort and providing examples).

irrespective of whether the touching causes harm.[104] One's interest in maintaining one's "personal space" as against intrusions by others is equally central to other torts, including battery's longtime-partner assault, as well as false imprisonment.[105] Claims for intentional and negligent infliction of emotional distress are not vindications of a right to happiness. Rather, they are rooted in the notion that one takes a significant hit in one's ability to live well when placed in the sort of oppressively difficult situation that is sufficient to cause an ordinarily constituted person to fall apart.[106] Malicious prosecution and abuse of process involve a special form of harassment through the spiteful invocation of the legal system.[107]

It takes but a short step to move from the vindication of individuals' dignitary interests via battery, assault, and false imprisonment claims (among others) to the recognition as torts of workplace sexual harassment, civil rights violations, and human rights violations. The latter quite clearly count as torts, unless one supposes that a wrong cannot be a tort unless it is grounded in precedent rather than statute or constitution: Each involves an assertion that the defendant has committed a legal wrong against the plaintiff, for which she or he is now entitled to recourse. (In fact, sexual-harassment claims began as common law tort claims,[108] as did false-arrest claims that now, along with other claims against government actors, fall under the entirely apt heading of "constitutional torts."[109])

We turn next from constitutional torts to contemporary commercial litigation. At the heart of the latter are claims for fraud, a basic tort that protects one's interest in being able to transact with others in an environment unpolluted by false information.[110] Thanks to Section 10(b) of the Securities

---

104. *See* RESTATEMENT (SECOND) OF TORTS § 15 cmt. a (1965) ("A contact which causes no bodily harm may be actionable as a violation of the right to freedom from the intentional infliction of offensive bodily contacts."); *id.* § 18 (defining the tort of battery to include "offensive contact with the person of the other").

105. *See* KEETON ET AL., *supra* note 12, § 10, at 43 (observing that assault protects the individual's interest in being free of threats of touchings); *id.* § 11, at 47 (observing that false imprisonment protects the interest in being free of restraints on bodily movements imposed by others).

106. *See* RESTATEMENT (SECOND) OF TORTS § 46 (requiring that the emotional distress inflicted by reckless or intentional outrageous conduct must be "severe"); *id.* § 312 & cmt. d (noting that negligent infliction of emotional distress is actionable only if it is serious enough to result in illness or bodily harm).

107. RESTATEMENT (SECOND) OF TORTS §§ 653, 682 (1977).

108. *See* CATHARINE A. MACKINNON, SEXUAL HARASSMENT OF WORKING WOMEN: A CASE OF SEX DISCRIMINATION 165 (1979) ("Sexual touching that women do not want has historically been considered tortious under a variety of doctrines . . . .").

109. *See* SHELDON H. NAHMOD ET AL., CONSTITUTIONAL TORTS 1 (2d ed. 2004) ("Constitutional torts are actions brought against governments and their officials and employees seeking damages for the violation of federal constitutional rights, particularly those arising under the Fourteenth Amendment and the Bill of Rights.").

110. *See* RESTATEMENT (SECOND) OF TORTS § 525 (explaining that liability for fraud results from misrepresentations of fact that induce others to rely on that misrepresentation); *see also* Michael D. Green, *Apportionment, Victim Reliance, and Fraud: A Comment*, 48 ARIZ. L. REV.

Act of 1934 and Rule 10b-5, as interpreted by a Supreme Court once receptive to implied rights of action,[111] federal courts are now the preferred forum for commercial fraud cases, and sales and purchases of securities are the most frequently addressed transactions.[112] But since there is precious little in the statute or the rule to help flesh out what the wrong of securities fraud really is, courts have relied on the common law of fraud as they have articulated this new area of civil recourse for private wrongs.[113] Other common law torts that protect the interest in freely entering into transactions include tortious interference with contract and with prospective economic advantage,[114] injurious falsehood,[115] slander of title,[116] and various forms of unfair competition.[117]

Dignitary torts such as battery and transactional torts such as fraud connect in interesting ways both to the hoary torts of libel and slander, and to the distinctly modern actions for invasions of privacy. The injury for these torts consists of an interference with the victim's ability to interact with others—one that involves the defendant acting so as to alter in a deleterious way how others view the victim. In defamation, the interference typically consists of inducing others to attribute to the victim some act or quality that renders her contemptible, untrustworthy, pitiable, ridiculous, or the like.[118] Invasions of privacy involve the disclosure (or false attribution) of an

---

1027, 1036 (2006) (arguing that the tort of fraud protects a person's transactional decision-making interests).

111. *See* Superintendent of Ins. of N.Y. v. Bankers Life & Cas. Co., 404 U.S. 6, 13 n.9 (1971) ("It is now established that a private right of action is implied under § 10(b).").

112. *See* 3 THOMAS LEE HAZEN, THE LAW OF SECURITIES REGULATION § 12.4, at 531 (6th ed. 2006) ("Rule 10b-5 has had a tremendous impact on a broad spectrum of securities litigation. This general antifraud rule is the most commonly used basis for private suits charging fraud in connection with the purchase or sale of securities.").

113. *See* Basic, Inc. v. Levinson, 485 U.S. 224, 253 (1988) (White, J., concurring in part and dissenting in part) ("In general, the case law developed in this Court with respect to § 10(b) and Rule 10b-5 has been based on doctrines with which we, as judges, are familiar: common-law doctrines of fraud and deceit.").

114. *See* Lumley v. Gye, (1853) 118 Eng. Rep. 749 (Q.B.) (U.K.) (establishing liability for inducing the breach of a contract); Charles E. Carpenter, *Interference with Contract Relations*, 41 HARV. L. REV. 728, 728 (1928) (arguing that the tort of interfering with contract relations extends beyond mere inducement of breach).

115. *See* Gale v. Ryan, 31 N.Y.S.2d 732, 734 (N.Y. App. Div. 1941) (recognizing the tort of publishing injurious falsehoods, distinct from libel and analogous to slander of title); William L. Prosser, *Injurious Falsehood: The Basis of Liability*, 59 COLUM. L. REV. 425, 425 (1959) ("There is a tort which passes by many names. . . . It consists of the publication, or communication to a third person, of false statements concerning the plaintiff, his property, or his business, which cause him pecuniary loss.").

116. *See* Wilson v. Dubois, 29 Nw. 68, 68 (Minn. 1886) ("False and malicious statements, disparaging an article of property, when followed, as a natural, reasonable, and proximate result, by special damage to the owner, are actionable.").

117. *See* RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 1(a) (1993) (reporting that one may be liable for unfair competition by causing harm to the commercial interests of another by engaging in practices determined to be actionable).

118. RESTATEMENT (SECOND) OF TORTS § 559 & cmts. b, c (1977).

embarrassing personal fact,[119] or the (unwanted) association of a person with a commercial enterprise.[120]

Bodily integrity, possessory interests, personal space, freedom to transact, the maintenance of one's standing in the eyes of others—each of these is an important interest, interference with which may amount to the commission of a tort. It is thus plausible that the interests protected by tort law are plural and irreducible. Any loss-based account faithful to the content of tort law will therefore need to be similarly capacious. Candor and a refusal to engage in reductive thinking should not count against a comprehensive wrongs-based account of tort law any more than it should against a nonreductive loss-based account. While admitting the plurality of wrongs and types of wrongs may dash the hopes that some theorists have for a certain kind of essentialist theory of Torts, that is far from making a wrongs-based theory of Torts incoherent.

B. Realized Wrongs and Duties of Noninjury

In responding to the Hodgepodge Problem, we have emphasized the type of interests that are protected by tort law. No doubt this emphasis will invite some readers to accuse us of having switched from a wrongs-based account of torts to an interest-based or injury-based account. What is central to torts, we seem now to be suggesting, is not how the injurer conducted himself but instead what has happened to the victim. This imagined objection is off the mark, for reasons best explained by turning to the Realization Problem. For it is really just a variation on the claim that gives rise to that problem—namely, that the act and injury components of every tort must be set apart for separate appraisal.

Recall that the Realization Problem centers on the putative philosophical insight that wrongfulness can only be an attribute of conduct, not results. On this view, the tort of negligence is properly depicted as the wrong of acting carelessly coupled with a harmful consequence caused by that wrong, as opposed to the complex or compound wrong of injuring someone through careless conduct. Yet even from within the sort of command-centered framework that is so strongly slanted against comprehending torts as wrongs, the idea of wrongs built on duties of non-injury can be perfectly intuitive. If one were to reduce the tort of battery to a command, it would be: "No unwelcome touching!" Whether one has complied with this directive turns on whether one has in fact touched another, or caused a touching.

Perhaps one could strain to capture a wrong such as battery in purely conduct-related terms. One sees a related effort in Holmes's convoluted attempt to reduce the tort of deceit—built around a "command" to refrain

119. *Id.* § 652D.
120. *Id.* § 652C.

from actually deceiving another—to the idea of making a statement under circumstances in which one can expect it to be false and to deceive another.[121] Even if one is prone to credit such an effort, our point is made by the fact that intellectual artifice of this sort is required. Torts such as battery, trespass, libel, and fraud help us to see that the idea of a completed wrong is neither incoherent nor esoteric. And if the injunction contained in the law of battery is "No unwelcome touching!," why is it so odd to think that the injunction around which negligence is built is: "Don't injure another by acting carelessly toward her"?

What about the Utilitarian and Kantian objection that, insofar as one is in the business of identifying genuine wrongs (rather than using the term "wrongs" as an empty placeholder), one must focus on acts and not results, lest mere happenstance be allowed to infect what should be noncontingent judgments of right and wrong? Here we arrive at the difficult topic of "moral luck," first staked out in modern analytic philosophy by Bernard Williams and Thomas Nagel.[122] We have elsewhere addressed the significance of moral luck for torts, and of torts for philosophical discussions of moral luck.[123] Here, four points will suffice.

First, tort law plainly involves something of a legal analogue to moral luck. Careless driving that results in a running down is a tort; careless driving that injures no one is not a tort. Second, it is common in ordinary moral evaluation to see a moral difference between these two cases, with the first being deemed worse than the second.[124] From this widely embraced perspective, the wrong of negligently injuring someone is not fully captured by the wrong of negligently driving plus the loss it caused. Insofar as ours is a conventionalist or coherentist conception of tort law that aims to characterize the moral principles that our society and legal system have entrenched within the law of torts, the existence of a framework of moral thought that people deploy regularly in their daily lives and that runs along these lines is of great significance.

Third, if it is really true that there is a conflict between certain systematic moral theories and the ordinary mindset just depicted above, it does not follow that the ordinary mindset is the one that needs to be rejected. Instead, with Williams, one might cogently argue that it is so much the worse for a certain kind of moral philosophizing. Or one might suppose, with Nagel, that we should credit each perspective and work back and forth between them. The point is that one cannot simply assume that clear

121. *See* HOLMES, *supra* note 17, at 106–10 (arguing that it is sufficient for the tort of deceit that the defendant made a misrepresentation under circumstances where it could be expected to mislead another).

122. *See* John C.P. Goldberg & Benjamin C. Zipursky, *Tort Law and Moral Luck*, 92 CORNELL L. REV. 1123, 1126 n.8 (2007) (briefly describing Nagel's and William's views).

123. *See generally id.*

124. *Id.* at 1128–29.

thinking requires the elimination of consequences from any standard that purports to be a genuine standard of wrongful conduct.

Finally, an embrace of injury-inclusive wrongs promises to shed light both on tort law and on the problem of moral luck. Appreciating the availability of completed wrongs as one kind of wrong allows us to see that there are many different reasons for using the language of wrongs and many different reasons for using normative concepts in answering practical questions. It is useful to contrast three of these, one prospective and two retrospective. A person deciding prospectively whether to act in a particular manner will want to know if the conduct counts as a wrong, legal or moral, which typically will (and should) count against it. Alternatively, one looking retrospectively at actions that have already been performed might rely on categories of wrongs (or non-wrongs) in assessing the character or quality of another person. A third usage is retrospective, but different from the second. One might want to ascertain whether some untoward event—the burning down of a house, for example—is so connected with human agency that it should be understood not merely as something that just happened, but ought to be understood as another's *doing*: someone's burning down of the house.[125] Within this kind of assessment of injury-inclusive acts, some may be categorized as wrongs and some may not: The deliberate destruction of a house as a public-health measure is different from a loan shark's willful act of arson or a suburban barbecuer's accidental burning down of his neighbor's house. A person considering whether he is morally responsible for burning down his neighbor's house is asking whether acts and results are linked in such a way as to constitute the moral wrong of negligently burning down the house.

Tort litigation typically presents a legal version of the third kind of question: A person who feels aggrieved or injured is attempting to respond to what he or she perceives as a wronging at the hands of another.[126] Indeed, in tort law it is particularly clear that a defendant's vulnerability to an action by the plaintiff should turn on whether the defendant actually injured the plaintiff, for the injury is intrinsic to the wronging of which the plaintiff complains.[127] From the plaintiff's perspective, it is not correct to say that there just happens to have been a conjunction of her loss and wrongful conduct by the defendant: In her eyes the defendant's wrong is mistreating her or interfering with some aspect of her well-being (or failing to protect or assist her in ways that would have prevented her from suffering a certain kind of setback). More importantly, the court's obligation to provide an avenue of civil recourse against the defendant hinges on the defendant having

125. The distinction between these two dimensions of retrospective examination is explored at greater length in Benjamin C. Zipursky, *Two Dimensions of Responsibility in Crime, Tort, and Moral Luck*, 9 THEORETICAL INQUIRIES L. 97 (2008).

126. *See supra* note 33 and accompanying text.

127. *See supra* subpart III(C).

wronged the plaintiff in a manner that renders her a victim *entitled* to respond to the wrongdoer.[128]

Those puzzling over the relevance of harm to degree of responsibility have always recognized the fact that the victim of the harm will be naturally disposed to feel differently towards the wrongdoer when the wrongful conduct ripens into harm than she would if no harm ensues.[129] What they have wondered about is why it should make a difference to how the wrongdoer's acts are categorized and evaluated from a more objective perspective.[130] As to this question, tort theory helps moral theory. One of the things we are asking about when we evaluate someone's conduct is what acts he has done. And there is no ground for insisting that a classification that abstracts from results carve at the normative joints—often, in fact, the opposite may be true.[131] A morally significant aspect of what an actor has done is whether his acts—described in a result-inclusive way—are ones that another person could fairly demand that he be held accountable for.[132] That is, even assuming that the increased resentment felt by the victim is not itself to be converted into an attribution of greater blameworthiness to the author of the injurious act, tort law helps us to see a distinct but related point. The question of whether a defendant's blameworthiness is greater is a question of whether the degree and nature of the resentment (not improperly) felt by a victim of the result-inclusive wronging is greater than that of a victim of a harmless wronging. A heightened degree of blameworthiness does not necessarily entail an increased level of wrongfulness, but it may reflect an increase in the level of blame by others to which a third party (like the state) would regard the wrongdoer as properly vulnerable. To say that an actor could reasonably be resented to a greater degree is not to say that there was some respect in which the actor's conduct ought to be deemed more wrongful; on the other hand, increased blameworthiness in the sense of increased grounds for resentment may indeed be an attribute of the actor's actions, not simply a reification or projection of the spontaneous or natural reactions of others.

Torts are not wrongful acts that happen to cause certain kinds of injuries. They are wrongings. For every tort, there is an inquiry into the nature of the tortfeasor's actions (whether intentional in some sense or careless), the nature of the setback suffered by the victim, and the connection between the two.[133] A driver who accidentally loses control of his car and slides onto the ice-covered pond of his neighbor without further incident has

---

128. *See supra* subpart IV(A).

129. *See* Goldberg & Zipursky, *supra* note 122, at 1154 ("[V]ictims of these norm violations are likely to regard themselves as having been wronged and tend to have concomitant feelings of resentment and blame in response.").

130. *Id.* at 1153–54.

131. *Id.* at 1156–59.

132. *Id.* at 1160.

133. *E.g.*, *supra* notes 27–33 and accompanying text.

not committed a trespass or negligence. There is no trespass because the driver did not set out to make contact with this neighbor's land.[134] There is no negligence because there was no damage to the property.[135] As it is with respect to every other tort, so with respect to these: There must be conduct and a result of a certain description for there to be a wrong in the tort sense of "wrong."

C. *Accountability, Relational Wrongs, and Civil Recourse*

Coleman, Perry, and others have inferred from tort law's distinctive form of accountability that torts cannot be characterized as wrongs.[136] The inference is unwarranted. It is true, of course, that crimes are punished and torts are typically not, at least *qua* torts.[137] But this fact merely forces us to ask whether there is a form of accountability for legal wrongs that is different from punishment imposed at the behest of a public prosecutor. In fact, there is. But before explaining this form of accountability, it is crucial to develop further our account of torts as legal wrongs.

As we have explained elsewhere, torts are a special kind of legal wrong not only because they are injury-inclusive or realized wrongs but also because they are *relational wrongs*.[138] Some crimes or regulatory infractions are wrongs *tout court*—wrongs to the world or to the state. It makes perfect sense for the law to recognize a crime of narcotics possession or of selling alcohol without a license, or to adopt a regulatory law against littering. Each of these is cogently described as a "legal wrong." Notably, the directives that enjoin these kinds of action are of a certain form. Each is what might be called a *simple legal directive*: For all x, x shall not A. *Simple legal wrongs* are violations of simple legal directives.[139]

Torts are violations of legal directives with a different analytic structure. We call these *relational* directives. They are of the form: For all x and for all y, x shall not do A to y.[140] These are directives that enjoin people not to treat others in certain ways or that require them to treat others in certain ways. For example, the legal directive underlying the tort of trespass enjoins

---

134. RESTATEMENT (SECOND) OF TORTS § 166 (1965).

135. *Id.*

136. *See supra* notes 48–50 and accompanying text.

137. *See supra* notes 53–54 and accompanying text.

138. *See* Zipursky, *supra* note 8, at 59–63 (providing an overview of the concept of relational wrongs).

139. *Id.* at 59–60.

140. *Id.* at 59. Some relational legal wrongs are defined in a way that requires the first relatum to be a member of a less-than-universal group (e.g., the class of physicians, not the class of persons) or requires the second relatum to be a member of a less-than-universal group (e.g., the class of patients, not the class of persons), or requires both, or requires that one of the relata be connected in some way (e.g., for all physicians, and all persons who are patients of that physician, the physician shall not injure the patient by failing to take the care she owes him; for all attorneys, and all clients of that attorney, the attorney shall not divulge information provided to him by that client to any other person).

persons from interfering with a *possessor's* right to exclusive possession.[141] Battery involves an intentional harmful or offensive touching *of another*.[142] Medical malpractice (usually) involves *harming a patient by treating her incompetently*.[143]

One of the special features of relational and injury-inclusive wrongs is that they are defined in such a manner that their commission entails that there is a person who counts as having done the wrong and a person who counts as having been the victim of the wrong. If the tort of libel has occurred, then there is a person who libeled someone and there was someone who has been libeled.[144] If the tort of fraud has occurred, then there is a person who committed the fraud and someone who has been defrauded.[145] If a car accident turns out to have involved the tort of negligence, then there is someone who negligently injured someone else.[146]

It is not hard to see that relational and injury-inclusive wrongs, so understood, simultaneously confer both primary duties and primary rights. The tort of libel contains a relational directive that creates a legal duty not to libel others and a legal right not to be libeled.[147] Fraud generates a duty not to defraud others and a right not to be defrauded.[148] Negligence imposes a legal duty not to injure others through conduct that is careless toward them and a legal right against being injured by such conduct.[149]

With this framework in place, it should now be clear how and why the response authorized by tort law to the commission of torts is distinct from the response provided by criminal law to crimes. Tort law permits victims of relational, injury-inclusive wrongs to obtain a court's assistance in redressing the wrongs that have been done to them. It gives each such victim a right of action, the aim of which is to obtain a remedy—usually, but not always, money damages.[150] We have generalized the point and summed it up in a phrase: Tort law provides victims with an avenue of *civil recourse* against those who have committed relational and injurious wrongs against them.

Tort law is thus plainly private law in the sense that it is about empowering *private parties* to initiate proceedings designed to hold

141. RESTATEMENT (SECOND) OF TORTS § 158 (1965).

142. *Id.* § 18.

143. *Id.* § 299A.

144. RESTATEMENT (SECOND) OF TORTS § 558 (1977).

145. *Id.* § 525.

146. RESTATEMENT (SECOND) OF TORTS § 281 (1965).

147. *See* RESTATEMENT (SECOND) OF TORTS § 558 (1977) (establishing liability from one person to another for libel, thereby creating a duty not to libel and a right not to be libeled).

148. *See id.* § 525 (establishing liability from one person to another for fraud, thereby creating a duty not to defraud and a right not to be defrauded).

149. *See* RESTATEMENT (SECOND) OF TORTS § 281 (1965) (establishing liability from one person to another for negligence, thereby creating a duty not to injure negligently and a right not to be injured negligently).

150. *See* RESTATEMENT (SECOND) OF TORTS § 901 (1979) (explaining how damages are awarded in tort law); *id.* § 902 (defining damages as a monetary payment).

tortfeasors accountable. This marks one of a number of fundamental differences between tort law and criminal law, which empowers *the state* to hold wrongdoers accountable.[151] Another is that both the stigma and the human consequences of what is done to the wrongdoer are very different for criminal wrongs than for tortious wrongs; part of what we do to render tort law civil, and amenable to private dispute resolution, is to cash out the remedy in monetary payments.[152] State prosecutions and punishments are quite different, of course, from tort suits and the exaction by tort plaintiffs of damages. The differences between seeking and obtaining punishment and seeking and obtaining civil liability exist both in consequences and in social meaning. Yet both involve courts responding to a demand to hold a wrongdoer accountable for having committed a wrong. Punishment and civil liability to a victim are different forms of accountability that both figure in our legal system's response to wrongdoing.

Finally, and *contra* Coleman and Perry, it is quite straightforward to explain why a body of law that defines private wrongs would frequently offer a remedy linked to the plaintiff's loss. In a large number of cases, particularly with respect to accidents, the wrong alleged is the careless infliction of physical harm.[153] In these cases, the successful victim will have the right to exact a remedy, and our courts will apply principles of remedies to select the appropriate level of money damages.[154] A longstanding principle of remedies for nonwillful wrongs sets make-whole as the default remedy.[155] This does not mean that the legal system has somehow reallocated the loss of the victim; it means that the victim's loss normally sets the outer boundary of the remedy that courts will provide to victims of wrongs when they successfully sue those who wrong them.

D. Torts as Legal Wrongs

Now that we have a better sense of the distinctive features of torts as wrongs, we are in a better position to address the Moral–Legal Dilemma. Recall that, according to it, one cannot characterize torts as moral wrongs without losing the ability to account for large swaths of doctrine, yet one cannot characterize torts as legal wrongs without rendering the concept of

---

151. *See, e.g.*, 18 U.S.C. § 2 (2006) (authorizing punishment by the state against anyone who commits a crime).

152. *See* Zipursky, *supra* note 8, at 82–93 (discussing the link between torts understood as relational wrongs and the asserting of claims by victims against alleged wrongdoers).

153. *See* Kenneth S. Abraham, *The Trouble with Negligence*, 54 VAND. L. REV. 1187, 1188 (2001) (describing the centrality of negligence to modern tort law).

154. *See* RESTATEMENT (SECOND) OF TORTS § 910 (describing tort victims' default entitlement to "make-whole" compensation).

155. *See* JOHN C.P. GOLDBERG, ANTHONY J. SEBOK & BENJAMIN C. ZIPURSKY, TORT LAW: RESPONSIBILITIES AND REDRESS 476, 484 (2d ed. 2008) (comparing "[e]xtra-compensatory" damages that are only available to victims of certain "'aggravated' forms of mistreatment" with "compensatory damage[s]," commonly described as efforts to make "the plaintiff whole," which are available to nearly any tort plaintiff).

"wrong" vacuous (a legal wrong being anything the law defines as a legal wrong). The solution to this dilemma can be found by invoking a set of ideas made famous by H.L.A. Hart, though we should be quick to emphasize that our use of these ideas is not meant to entail, and does not entail, that we embrace all aspects of Hart's jurisprudence (such as the so-called "social facts" thesis).[156]

Hart argued that certain rules of conduct carry directive or injunctive force, whether they are putative moral rules, legal rules, rules of etiquette, or rules associated with some other sort of social association (e.g., a club) or institution (e.g., a school).[157] While their force is injunctive, their grammatical form need not be imperatival.[158] For example, a parent uttering to his child, "The fork goes to the left of the plate and the knife goes to the right," is expressing a rule of etiquette but not issuing an imperative. A swimming teacher who utters, "Bathing caps are required," is similarly issuing (an institutional) rule of conduct. We call such rules "directives." When uttered, they are typically meant to guide conduct, at least in part. Hart's account did not depend upon treating such rules as utterance tokens or utterance types: rules of conduct of this form could exist in social settings by virtue of a social practice of complying with the rule and expecting such compliance with others—more generally, by being entrenched in social practice in a certain way.[159] Moreover, at least in the case of legal rules, a legal rule could exist in some sense by qualifying under a meta-rule that was itself entrenched in social practices.[160]

For certain families of social practices, according to Hart, the kind of pressure putatively imposed by the directive, in combination with other

---

156. We have previously analyzed Hart's thought as it bears on tort law. *See* John C.P. Goldberg & Benjamin C. Zipursky, *Seeing Tort Law from the Internal Point of View: Holmes and Hart on Legal Duties*, 75 FORDHAM L. REV. 1563, 1572–81 (2006) (applying Hart's framework to tort law to develop a "duty-accepting" concept of tort); *see also* Zipursky, *supra* note 8, at 58 (discussing Hart's influence on modern views of legal rules).

157. *See* H.L.A. HART, THE CONCEPT OF LAW 86 (2d ed. 1994) (arguing that rules are conceived of as "imposing obligations when the general demand for conformity is insistent and the social pressure brought to bear upon those who deviate or threaten to deviate is great").

158. *See id.* at 9–10 (explaining that for social rules it is not necessary to use imperatival words such as "must," "should," and "ought to," and that using those words for "mere convergent behaviour" would be confusing).

159. *Id.* at 86.

160. *See id.* at 86–87 (explaining that there may be social pressure to conform both indirectly by appealing to the individual's feelings, as is common for rules of morality, and directly through physical sanction, as is common for legal rules). Despite our reliance on a Hartian framework for thinking about the place of rules within a legal system, we need not and do not accept a social-facts based form of legal positivism. For an articulation of this "mixed" jurisprudential position, see Benjamin C. Zipursky, *The Model of Social Facts*, *in* HART'S POSTSCRIPT 219, 268–70 (Jules Coleman ed., 2001), accepting certain aspects of Dworkin's critique of the social-facts thesis but also salvaging aspects of Hart's analytic framework, and Benjamin C. Zipursky, *Pragmatism, Positivism, and the Conventionalistic Fallacy*, *in* LAW AND SOCIAL JUSTICE 285, 308 (Joseph K. Campbell et al. eds., 2005), rejecting Coleman's inclusive positivism for relying on the social-facts thesis.

features of the entrenchment of the practice in institutions and in language, leads the directives to have a special quality; he called these "duty-imposing" rules.[161] Moral directives and legal directives both have this quality according to Hart.[162] The statement that it is wrong to lie, spoken by a parent to a child or written by an opinion columnist for newspaper readers, contains injunctive force: it condemns lying and conversely urges refraining from lying. Making such a statement is, moreover, identifying a way of treating other people as unacceptable. The same is true when a court holds liable a broker who has misrepresented a company's financial condition to an investor who relied on that misrepresentation to his detriment. The court is articulating a norm of conduct that requires certain actors to refrain from deceiving other to their detriment and condemns doing so as wrongful. The first would be said to be a duty-imposing rule of morality, the second a duty-imposing rule of law.

There is a big difference between uttering a rule of conduct as a moral rule and uttering a rule of conduct as a legal rule, however. The parent and the opinion writer are claiming that the norm of truth telling should be complied with *because it is morally sound.* The judge, by contrast, may or may not be referring to the moral soundness of the law's injunction against misrepresentations. The putative authority that lies behind the judge's categorization of certain conduct as tortious is the same authority that accompanies all legally justifiable statements within the legal system. Put differently, the "Says who?" challenge to the rule or norm against lying earns a different response in the moral case and the legal case. In the moral case, the speaker might plausibly take the challenge to be off point; the *provenance* of the "no lying" norm is not the issue—its soundness or truth is the issue. In the legal case, things are otherwise. The "Says who?" challenge is answered directly by reference to the *source* of the norm—the legal system. The norm's provenance (which, again, may not be determinable by a purely "positive" inquiry into "social facts") is critical to its authority.

To assert that some act is a legal wrong is to assert that it violates a legal directive. To be sure, many legal directives are legislative, constitutional, or regulatory, but not all are. Our common law system has built up precedents that identify certain kinds of acts as grounds for liability, and it has done so in a manner that conveys disdain for those acts and expresses an injunctive message that such acts are not to be performed. This is part of the force of leading opinions, such as Cardozo's memorable rejection of the privity rule in *MacPherson v. Buick Motor Co.*[163] Although it was of course intended to resolve the dispute before him, and to distinguish the domains of tort and

161. HART, *supra* note 157, at 81.

162. *Id.* at 86–87.

163. 111 N.E. 1050, 1053 (N.Y. 1916).

contract, his opinion was no less keen to establish two related ideas.[164] First, it announced unequivocally a duty of conduct embedded in the law of negligence.[165] Whatever the ambiguities of prior case law, thenceforth the manufacturer of a product posing substantial risks of physical harm that was to be sent into the stream of commerce without further inspection was unmistakably under a duty to product users to take care to prevent such harm. Injuring a consumer by releasing a product manufactured without sufficient "vigilance" against dangerous defects—a breach of the duty owed to users not to injure them by failing to "make [the product] carefully"—was clearly identified as a legal wrong that, when committed, would subject the manufacturer to liability.[166] Second, Cardozo was equally emphatic as to the source of this duty. It resided neither in promise nor in contract, but "in the law."[167] The failure of a manufacturer to take care to prevent its product from containing dangerous, harm-causing defects was thereby identified as one important instantiation of the broader legal wrong of negligence.

Many torts are moral wrongs, but to claim that an act is a tort is not to assert that it violates a moral directive but that it violates a legal directive. Even when a particular tort is not also a moral wrong, saying that it is a legal wrong is similar to saying that it is a moral wrong in at least the following respects: it asserts that the act in question is not to be done, and that it merits some form of accountability when done. However, the statements and principles categorizing torts as such are provided from within a legal system that has authority apart from whether its edicts all prove to be sound. Although in New York there was not and is no statutory code that lists these legal directives, there is, embedded in and discernible in New York law, properly interpreted, an answer to the question of which ways of treating others are torts for which an actor may be held legally accountable. In this sense, what is and is not a tort in any particular jurisdiction is a matter of law. Like Hart, we recognize that the normative aspect of a legal duty does not necessarily (or even typically) stem from the threat-like force of a legal command, and its provenance need not be statutory or legislative. Like Dworkin (and to a lesser, but still significant extent, Hart), we recognize that pinning down what the law actually says about some putative duty of conduct will be a process fettered within legal argumentation that uses both moral and nonmoral concepts.[168] Once one recognizes all three of these points, there is no

164. *See id.* ("If to the element of danger there is added knowledge that the thing will be used by persons other than the purchaser, and used without new tests, then, irrespective of contract, the manufacturer of this thing of danger is under a duty to make it carefully.").

165. *Id.*

166. *Id.*

167. *Id.*

168. For a methodological and jurisprudential approach that aims to sidestep debates between positivism and its critics but nevertheless takes both moral and nonmoral concepts to be central to the content of the common law of torts, see Benjamin C. Zipursky, *Pragmatic Conceptualism*, 6 LEGAL THEORY 457, 477–78 (2000).

obstacle to seeing tort law as a domain of duty-imposing legal directives. And then it is straightforward to understand torts—the violations of these directives—as legal wrongs.

Armed with this account of legal wrongs, we can readily explain the conundrum that seems to have skewed tort theory since Holmes's time: namely, how it is that torts are wrongs—i.e., violations of standards of how one must act (or not act) towards others—yet generate liability for conduct that, judged in abstraction from the legal system, is not necessarily blameworthy. The fact that an act falls under an authoritative legal directive that characterizes it as a legal wrong does not entail that such an act, in the circumstances it actually occurred, warrants categorization as morally wrongful. A faultless trespass or a "Menlovean" act of negligence still constitutes the breach of a norm set by tort law.[169] As the tort of trespass to land is defined by judicial decisions, one commits a wrong when one intentionally treads on property that is owned by another, just as one commits negligence by injuring someone through conduct that fails to meet the standard of ordinary prudence.[170] For a variety of reasons internal to the idea of a law of private wrongs, excuses that might carry the day in the domain of positive morality (and properly so)—"But I had no reason to believe the land was owned by another!"; "But I tried my very best to be careful!"—are not recognized in tort.[171] This feature of tort law does not somehow render the dictates of tort unconnected to notions of wrong. A tort, even a blameless trespass, is still a breach of a legal norm of conduct. One is required by law not to invade others' property.

Doctrinally, the sharpest challenge to a wrongs-based theory of torts is found in decisions such as *Rylands v. Fletcher*[172] and the doctrine of strict liability for harms caused by abnormally dangerous activities.[173] These forms of liability, after all, explicitly disavow having anything to do with wrongs. Liability is imposed even though the defendant has caused harm through conduct that the courts themselves are at pains to say is entirely permissible.[174] It is possible that these are true exceptions to the otherwise wrongs-based nature of tort law. To allow as such is hardly to make a substantial concession. For just as "strict liability" versions of the doctrine of statutory rape rest uneasily at the edges of criminal law, and just as certain

169. *See, e.g.*, McDonald v. Village of Winnetka, 371 F.3d 992, 1009 (7th Cir. 2004) (noting that "tortious conduct is by nature a departure from some norm").

170. *See supra* notes 74, 100.

171. *See* Goldberg & Zipursky, *supra* note 122, at 1143–63 (explaining how and why moral blameworthiness is often not a factor in determining tort liability).

172. (1868) 3 L.R.E. & I. App. 330 (appeal taken from Eng.) (U.K.).

173. *See* RESTATEMENT (SECOND) OF TORTS § 520 (1977) (articulating a test for when an activity is "abnormally dangerous" so as to generate strict liability for injuries caused by it).

174. *See, e.g.*, *Rylands*, 3 L.R.E. & I. App. at 340 ("If [a person accumulates lawfully anything on his land which] does escape, and cause damage, he is responsible, however careful he may have been, and whatever precautions he may have taken to prevent the damage.").

forms of promissory estoppel likewise barely belong within contract, *Rylands* and subsequent decisions identify a genre of case that sits at the margin of tort law. Each of these doctrines represents a precarious patching over of a vulnerable spot that marks the outer boundaries of the domain of law in which it is situated. To be sure, if *Rylands* were emblematic of a broad area of tort law, rather than a narrow exception, it would pose a challenge to the idea that liability for torts is wrongs based.[175] But so long as it remains sui generis, its existence does not count as evidence against our general interpretive account. In fact, courts have consistently beaten back efforts by plaintiffs' lawyers to expand the category of abnormally dangerous activities.[176] Thus, it remains entirely viable, from both a doctrinal and a jurisprudential perspective, to treat Torts as the domain of civilly actionable legal wrongs.[177]

Is the account of wrongs we have put forward jurisprudentially partisan, in that it requires embrace of legal positivism? The answer depends on what is meant by "positivism." To the extent that positivism is associated with the idea that the truth of legal claims is dependent on the content of principles, rules, and standards that are linked in an appropriate way to what certain institutionally entrenched sources say, we agree. However, to say this by no means entails rejecting the idea, associated most commonly with Dworkin,[178] that nonlegal moral reasoning—reasoning about rights, duties, and other concepts abstracted from the institutionalized setting of legal sources and legal claims—may or must be employed in interpreting what those sources say. Likewise, we join natural law theorists (and Hart himself) in recognizing that the concepts of wrong, duty, obligation, and right employed within tort law have roughly the same kind of normative force as they do when used outside of law.[179] In this sense, the concept of duty used in deliberating about the scope of duty within negligence law is a moral concept, just as is

175. Gilmore's argument for the "death of contract"—a manifestation of many of the same intellectual forces we are canvassing here—was of the same form. *See* GRANT GILMORE, THE DEATH OF CONTRACT 97–100 (1974) (arguing that the continued adoption of the promissory-estoppel doctrine could allow contract law to ignore the requirement of consideration in the form of reciprocal obligations). He inferred from what he took to be a trend toward greater judicial reliance on expansive forms of promissory estoppel that it no longer made sense (if it ever did) to think of contract law as being fundamentally about agreed-upon obligations. *Id.* Our sense is that Gilmore's argument was no less overstated than arguments alleging that the presence of certain forms of "strict" liability spell the "death" of torts *qua* wrongs.

176. *See* RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL & EMOTIONAL HARM § 20 reporters' note, at 296–328 (Proposed Final Draft No. 1, 2005) (identifying numerous instances in which courts have declined to apply the rule of strict liability for abnormally dangerous activities).

177. Likewise, we think it is erroneous to see in doctrines such as respondeat superior and "strict" products liability a judicial embrace of tort liability without regard to wrongdoing.

178. *See* RONALD DWORKIN, TAKING RIGHTS SERIOUSLY 149 (1978) (arguing for the fusion of constitutional law and moral philosophy).

179. *See* Goldberg & Zipursky, *supra* note 156, at 1575–77 (illustrating how Hart's concepts of duty within tort law have analogues outside the field of law within the realm of general social expectations).

the concept of duty used in deliberating about interpersonal and nonlegal questions such as whether I have a duty to offer to house my friend's family while they are waiting for their apartment to be repaired after a flood. When one describes an issue as a "moral question, rather than a legal question," the deliberator should ordinarily be understood to be drawing upon "duty" as a web of norms of conduct, not the institutionally entrenched web that we use the term "law" to refer to. Conversely, when a judge makes clear that she is talking about legal duties when she is deciding a case, not moral duties, she is indicating that she is identifying obligations within an institutionally entrenched web. As noted above, the articulation of the terms that constitute this web will often require the use of reasoning of a sort commonly deployed in discerning moral duties.[180]

V. Interpretive Failings of Loss-Based Accounts

It is one thing to show (as we have in Part IV) that a conception of Torts as wrongs is available; it is another to show that it is better than its main rival. In our Introduction and in the prior Part, we began to articulate reasons for selecting our account over others: understanding Torts as wrongs gives a sense of the breadth of the subject, of its continuity with many areas of law, and of our legal tradition's treatment of Torts as among a handful of fundamental legal categories such as Contracts, Property, and Criminal Law. The subject of wrongs and private redress for those wrongs satisfies these desiderata in a way that "accident-law-plus" simply does not.

Here we turn to more detailed interpretive work. Although detailed, we believe it cuts very deeply on the question of how Torts must be conceptualized. In prior work, we have argued for the interpretive superiority of a wrongs-and-recourse model on numerous grounds. These include its provision of illuminating accounts of leading cases old and new,[181] its explanation of the structure of tort law,[182] and its integration of

---

180. Our approach to tort law and law generally recognizes, as do almost all other approaches, that many of the acts categorized as "wrongs" by the law have been conceived as wrongs under the positive morality of the community and that their having been so conceived helps explain how and why they became part of the law of torts. Again like virtually everyone, we think that some wrongs of tort law are not necessarily wrongs of morality or of positive morality, and vice versa. Moreover, those which are often do not have exactly the same contours that they have in positive morality, or that would be appropriately identified as morally wrongful apart from positive morality. Finally, like virtually everyone, we believe that many of the wrongs of tort law are acts that merit description as wrongful quite apart from their entrenchment in a legal system that categorizes them as impermissible or wrongful; intentional physical assaults, injuring someone through careless driving, and deceiving someone out of a thing of value all fit this description.

181. *See, e.g.*, GOLDBERG, SEBOK & ZIPURSKY, *supra* note 155, at 783–90 (explaining Vincent v. Lake Erie Transp. Co., 124 N.W. 221 (Minn. 1910), without relying on the ad hoc incomplete privilege of private necessity); John C.P. Goldberg & Benjamin C. Zipursky, *Unrealized Torts*, 88 VA. L. REV. 1625, 1631–33, 1705–07 (2002) (reconciling the facially puzzling twin holdings of Metro-N. Commuter R.R. v. Buckley, 521 U.S. 424 (1997)); Benjamin C. Zipursky, *A Theory of Punitive Damages*, 84 TEXAS L. REV. 105, 105–07, 149–50 (2005) (interpreting BMW of N. Am., Inc. v. Gore, 517 U.S. 559 (1996), as articulating a line between punitive damages imposed as

core tort concepts in a manner that permits clear-eyed analysis of important contemporary issues of doctrine and policy.[183] In what follows, we build on these prior claims by addressing several basic features of tort law that are anomalous on loss-allocation models of tort, yet quite comfortably explained on a wrongs-and-recourse model.

A. Torts Without Losses

Allocative theories put losses at the center of tort law. A tort suit, they say, is always a plea to shift the plaintiff's loss to a tortfeasor.[184] It follows that a loss must be the predicate to a successful suit. Quite clearly, however, this is not the case. There are many torts that do not require a loss to be actionable. It is true that there is never a tort without an *injury*. But the concept of an injury is distinct from the idea of a loss that is capable of being shifted. This is a basic conceptual and theoretical point. In numerous instances, courts recognize that a tort has been committed and award a remedy to the victim, even though there is no loss to shift from the plaintiff to the defendant.

Trespass and nuisance, as well as battery and false imprisonment, do not set loss as a condition of liability.[185] For trespass to land, the question is whether the defendant physically invaded or occupied the plaintiff's property, thereby violating her right to exclude others.[186] In nuisance, the injury is an unreasonable interference with the plaintiff's right to enjoy her property.[187] Judgment can thus be entered for a trespass or nuisance plaintiff even if she and her land remain utterly unscathed. In its well-known *Jacque v. Steenberg Homes, Inc.* decision,[188] the Wisconsin Supreme Court ruled, soundly, that the defendant had trespassed by driving across the plaintiff's snow-covered field without permission—this even though the presence of the

---

redress for egregious wrongs and punitive damages imposed for regulatory purposes); Zipursky, *supra* note 8, at 13 (demonstrating the centrality of Palsgraf v. Long Island R.R., 162 N.E. 99 (N.Y. 1928), to tort law, understood as a law of wrongs).

182. *See generally* Benjamin C. Zipursky, *Civil Recourse, Not Corrective Justice*, 91 GEO. L.J. 695 (2003) (arguing that neither the corrective-justice model nor the law-and-economics model capture the structure of tort law, a system based on legal rights and wrongs).

183. *See generally* John C.P. Goldberg, *The Constitutional Status of Tort Law: Due Process and the Right to a Law for the Redress of Wrongs*, 115 YALE L.J. 524 (2005) (arguing for recognition of an individual's right to seek redress for wrongs as a means of limiting legislative tort reform); John C.P. Goldberg & Benjamin C. Zipursky, *Intervening Wrongdoing in Tort: The Third Restatement's Unfortunate Embrace of Negligent Enabling*, 44 WAKE FOREST L. REV. 1211 (2009) (applying the same framework to the question of negligence claims against remote actors).

184. *E.g.*, Goldberg, *supra* note 183, at 582.

185. Assault, libel, and slander per se also are defined so as not to require loss as a condition of recovery. *See* RESTATEMENT (SECOND) OF TORTS §§ 13, 21, 35, 158, 569–570, 822 (1965, 1977 & 1979) (specifying conditions of liability for, respectively, battery, assault, false imprisonment, trespass, libel, slander, and nuisance).

186. *Id.* § 158.

187. *Id.* § 822.

188. 563 N.W.2d 154 (Wis. 1997).

snow prevented harm even to a single blade of grass.[189] The injury consisted of the rights invasion, not a loss incurred as a result of it. Similarly, the tort of battery requires the plaintiff to have been the victim of a harmful or offensive touching.[190] A groping or kicking can be offensive and that is enough.[191] An intentional confinement that causes no physical harm, no pain, no anxiety, and no loss of opportunity can be a false imprisonment.[192] This much is clear from the chestnut of *Huckle v. Money*,[193] in which a confinement under pleasant conditions was deemed actionable.[194]

Needless to say, in many instances trespasses, nuisances, and batteries *do* cause losses for which the victim can obtain compensation.[195] But proof of losses is not necessary. In some cases without loss, nominal damages may be the only remedy available.[196] In others—including *Jacque* and *Huckle*—a court will permit an award of punitive damages.[197] In still others, the plaintiff may obtain declaratory or injunctive relief.[198]

The argument we have just made relies on features of tort doctrine that are obvious to anyone who studies the field. By the same token, it is not difficult to envision responses that loss-allocation theorists might offer.[199] One would be to suggest that the examples we have cited, and the torts that figure in them, are peripheral: that what really matters in tort is negligence, a tort for which loss is a component. This line of response fails to engage the criticism we are levying. A central point of this Article is to challenge the intellectual framework that treats tort law as coextensive with negligence law and to explain what has been lost because of scholars' attraction to such a

189. *Id.* at 160.

190. RESTATEMENT (SECOND) OF TORTS § 13.

191. *Id.*

192. *Id.* § 35.

193. [1763] 95 Eng. Rep. 768 (C.P.) (U.K.).

194. *Id.* at 769.

195. *See, e.g.*, Neal v. Miller, 778 F. Supp. 378, 387 (W.D. Mich. 1991) (granting the plaintiff compensatory damages for the physical consequences of a battery); Miller v. Cudahy Co., 592 F. Supp. 976, 1005 (D. Kan. 1984) (compensating plaintiffs for damages to crops caused by the defendant's continuing nuisance); Wilen v. Falkenstein, 191 S.W.3d 791, 799 (Tex. App.—Fort Worth 2006, pet. denied) (upholding actual damages based upon diminution in the value of land due to trespass).

196. *See* RESTATEMENT (SECOND) OF TORTS § 218, cmt. d (suggesting that only nominal damages may be available for some trespasses).

197. *See Jacque*, 563 N.W.2d at 161 ("We conclude that both the private landowner and society have much more than a nominal interest in excluding others from private land. Intentional trespass to land causes actual harm to the individual, regardless of whether that harm can be measured in mere dollars. . . . Accordingly, . . . we hold that nominal damages may support a punitive damage award in an action for intentional trespass to land."); *Huckle*, 95 Eng. Rep. at 769 (upholding an award of "exemplary damages" by the jury for an unlawful detainment by the government, despite no evidence of physical harm).

198. *See, e.g.*, RESTATEMENT (SECOND) OF TORTS § 933 (setting forth the criteria for enjoining a nuisance).

199. Of course it is open to a purely prescriptive theorist to argue that there should be no torts without losses, but that is a different question.

framework. In this context, it is nonresponsive simply to assert that negligence is the essence of Torts. Relatedly, loss-allocation theorists might concede that there are viable no-loss tort claims but insist that these are special exceptions to the general requirement of loss. As to a certain kind of boundary-pushing decision, this response might have some plausibility.[200] But, again, there is no justification for treating trespass, nuisance, and battery as outliers.

A different response to the presence of torts without losses is to insist that they really do involve losses. After all, the argument might proceed, successful trespass and false-imprisonment plaintiffs stand to recover a monetary payment. And if they are being paid damages, it must be because they have suffered a loss. What else can it mean for a plaintiff to obtain compensation? As we suggested in Part IV, there is an alternative and cogent account of compensatory-damage payments that does not treat them as having a logical or definitional connection to losses. According to that account, a tort award is compensation *for the wrong done to the plaintiff*—damages are what the victim of a tort is entitled to exact from the defendant in light of what the defendant has done to him.[201] This alternative conception of compensation as fair redress does not suggest that redress is or should be determined without regard to whether the plaintiff has suffered losses in connection with having been wronged. Plaintiffs whose injuries are accompanied by losses ordinarily are entitled to reimbursement for those losses.[202] But to say that reimbursement will tend to figure in the determination of what counts as redress is not to say that tort damages are just reimbursements, nor that reimbursable losses must be incurred before a tort can be committed. The governing concept is redress for a wrong done, where appropriate redress will typically include compensation for losses suffered *when there are such losses*.

Lastly, and in a related vein, some allocation theorists might argue that our critique of the centrality of loss presupposes an unduly narrow conception of what can count as a loss. A rights violation, they might say, really is a loss—a "normative" loss, or a debit on the balance sheet of life. The problem with this move is a familiar one. Like a rational-actor model of human behavior in which altruistic acts are accommodated as *really in one's self-interest*, it salvages the descriptive plausibility of loss-allocation theory only by abandoning what makes the theory distinctive.

When tort law is viewed as a law of private wrongs, the theoretical stresses and strains faced by loss-allocation theory in dealing with garden-

200. *See In re* Simon II Litig., 211 F.R.D. 86, 96 (E.D.N.Y. 2002) (Weinstein, J.) (certifying a punitive-damages-only class action for plaintiffs unable to prove that they had suffered losses by virtue of certain tobacco-company misrepresentations), *vacated* 407 F.3d 125 (2d Cir. 2005).

201. Goldberg, *Two Conceptions*, *supra* note 67, at 438–45.

202. *See generally* DOUGLAS LAYCOCK, MODERN AMERICAN REMEDIES 16 (3d ed. 2002) (arguing that the basic principle for damages is to restore the plaintiff to "the position he rightfully would have come to but for defendant's wrong").

variety torts such as trespass disappear. A private right of action is made available to the victim of one of these legal wrongs because she was wronged, not because she has incurred a loss. There may be some wrongs that are defined in such a way that the doing of the wrong involves bringing about an injury in another person of a sort aptly described as a loss. Negligence is perhaps one such tort; products liability may be another. But many other torts are not defined this way. If tort law is all about shifting losses, these torts become anomalies. But if tort law is about privately actionable wrongs, they do not. It is perfectly understandable why the law might count a certain way of treating another person as a wrong, and might conceive of the plaintiff as having been injured, even if there is no loss that stands in need of being shifted.

B. Foreseeably Caused Losses Without Torts

Allocative theories also face a problem converse to the one just described. Their focus on losses rather than wrongs renders them incapable of explaining why it is that a plaintiff must establish that *she has been wronged*, as opposed to proving that the defendant acted wrongfully toward another.

In prior work we have explained that, for each tort, it is a condition of liability that the defendant's tortious conduct be a wrong relative to the plaintiff. We have sometimes referred to this condition as a "substantive standing" requirement.[203] It is a "standing" requirement because it goes to the issue of whether the plaintiff is an appropriate person to assert a claim against the defendant. It is "substantive" because the rules that determine tort standing are among those that define the wrong(s) for which a plaintiff is suing.

The requirement of substantive standing comes under various guises and names. In trespass it is found in the rule that the plaintiff must have a possessory interest in the land trespassed upon.[204] Absent a possessory interest, there can be no recovery, even if the defendant's conduct constitutes trespass to another, and even if that trespass causes foreseeable losses to the victim.[205] Similarly, a plaintiff in a common law fraud case must prove that

---

203. This is Zipursky's original label for this concept. Zipursky, *supra* note 8, at 3–5.

204. *See id.* at 25 (showing that "in defining the contours of the right of exclusive possession, the courts are in fact defining who has substantive standing to sue for wrongs incurred through trespass").

205. *See id.* (expounding the rule that "[o]nly those who have a right of possession in the property trespassed upon have a cause of action for trespass"). Suppose *D* knowingly drives his car across *A*'s land without permission but for a good reason. In doing so, *D* is mindful that he has seen hikers on *A*'s land, and thus drives carefully. Now suppose that *P* is a hiker who, while conscientiously following a trail map that she reasonably believes is accurate, unintentionally strays onto *A*'s land. Even if *D* were to run down and injure *P* while both are on *A*'s property, *P* will not have a trespass claim against *D*. (And this is not because *P* can instead sue for negligence. We have assumed *D* was acting reasonably.) Although *P* was a perfectly foreseeable victim of *D*'s

she relied upon the defendant's misrepresentation, or at least its content.[206] A loss flowing purely from others' reliance on a misstatement will not support a common law fraud claim. A plaintiff in a libel case must prove that the defendant's libelous statement was "of and concerning" her.[207] A loss caused by a libelous statement made exclusively about others does not constitute a wrong that supports a libel claim. A negligence plaintiff must prove that the defendant breached a duty of care owed to the plaintiff.[208] A loss from a breach of a duty of care owed to someone else (and not to the plaintiff) does not constitute an injury that supports a negligence claim.[209]

In sum, a tort plaintiff cannot prevail merely by establishing that the defendant has acted *in some sense* wrongfully so as to cause her a loss under conditions where the causation of such a loss was reasonably foreseeable. As Cardozo memorably put it in *Palsgraf*, she must show "'a wrong' to herself; i.e., a violation of her own right, and not merely a wrong to someone else."[210] To say the same thing, a tort plaintiff "sues in her own right for a wrong personal to her, and not as the vicarious beneficiary of a breach of duty to another."[211]

Allocationists of all stripes often have ignored the substantive standing requirement.[212] Others have criticized it on normative grounds.[213] Apart from denial or critical concession, two other strategies might be tried. One aims to make sense of substantive standing as a fairness-based limit on the instances in which a court will shift losses from a plaintiff to a defendant. The problem in a case like *Palsgraf*, on this view, is that the plaintiff was asking the court to impose liability on the defendant even though its employees could not have reasonably foreseen that their conduct might injure her. Liability imposed on these terms is unfair.

The most serious problem with this response is that it does not come close to making sense of the substantive standing rules actually in the law. For within each tort, these rules cut off recovery *even when loss to the victim is entirely foreseeable*. A hustler who lies shamelessly to her elderly victim as part of a successful scheme to be named the sole beneficiary of his will, thus depriving his loving children of their inheritances, may know that her

---

trespass, *P* has no claim because *D*'s conduct was not a trespass with respect to property that *P* owned, leased, etc.

206. *Id.* at 18.

207. *Id.* at 17.

208. *Id.* at 8.

209. *Id.*

210. Palsgraf v. Long Island R.R., 162 N.E. 99, 100 (N.Y. 1928).

211. *Id.*

212. *See* Zipursky, *supra* note 8, at 4 ("Proponents of the most prominent theoretical approaches to tort law, law and economics and corrective justice theory, have generally neglected the substantive standing rule.").

213. *See, e.g.*, Dilan A. Esper & Gregory C. Keating, *Putting "Duty" in Its Place: A Reply to Professors Goldberg and Zipursky*, 41 LOY. L.A. L. REV. 1225, 1253–54 (2008) (criticizing *Palsgraf*'s relationality requirement on conceptual and pragmatic grounds).

misrepresentations will cause the children to suffer economic loss. Yet she has not defrauded the children, and they will have no claim for fraud against her.[214] Likewise, a physician might readily foresee that his failure properly to treat a potentially fatal illness suffered by the charismatic CEO of a large company will cause economic losses to the company's employees and shareholders. (Perhaps stakeholders will tell him as much.) Yet if he treats the patient incompetently, none of those who suffer economic loss will have an action against him. His duty to provide competent medical care is a duty owed to his patient; its breach is not a breach as to those suffering economic losses. By virtue of the substantive standing rule of negligence, they are therefore barred from recovering: There was only a breach of a duty owed to another, not a breach of any duty owed to them.

This last example may invite a different kind of counterargument from allocationists. It claims that tort law's substantive standing rules exist not to ensure costs are shifted only when it is fair to do so, but to address the administrative concern of preventing what would otherwise be a flood of litigation and liability.[215] It is not difficult to see the appeal of this argument. It matches what courts sometimes say in explaining why they would deny liability in a case such as the malpractice example just provided. It also fits with other tort rules or principles (such as certain proximate-cause limitations) that appear to have a similar rationale. And it creates a set of criteria by which to evaluate whether the rules should be maintained or modified.

The question, however, is whether floodgate rationales succeed in explaining substantive standing rules. They do not, for several reasons. First, while it is true that courts sometimes back their invocation of these rules with floodgates language,[216] such an explanation is often lacking and would not make sense. A judge who reasons that a private figure cannot recover for a libel contained in a club newsletter because she was harmed but not herself defamed is not worried about floods of litigation or excessive

---

214. Some courts would on these facts permit a claim for tortious interference with expectancy. *See generally* Diane J. Klein, *"Go West, Disappointed Heir": Tortious Interference with Expectation of Inheritance—A Survey with Analysis of State Approaches in the Pacific States*, 13 LEWIS & CLARK L. REV. 209, 210–11 (2009).

215. *See* Zipursky, *supra* note 8, at 53 (outlining the argument that "defendants will face crushing liability and courts will be flooded with cases" unless a tortious defendant's liability is limited and "recovery is denied . . . where substantive standing is lacking").

216. *See, e.g.*, Pruitt v. Allied Chem. Corp., 523 F. Supp. 975, 979–80 (E.D. Va. 1981) ("Perhaps because of the large set of potential plaintiffs, even the commentators most critical of the general rule on indirect damages have acknowledged that some limitation to liability, even when damages are foreseeable, is advisable. Rather than allowing plaintiffs to risk a failure of proof as damages become increasingly remote and diffuse, courts have, in many cases, raised an absolute bar to recovery. The Court thus finds itself with a perceived need to limit liability, without any articulable reason for excluding any particular set of plaintiffs."); Strauss v. Belle Realty Co., 482 N.E.2d 34, 38 (N.Y. 1985) (declining to impose liability on the defendant for the plaintiff's injuries on the basis that finding that the defendant owed the plaintiff a duty of care would "violate the court's responsibility to define an orbit of duty that places controllable limits on liability").

liability. The requirement of substantive standing is not simply tacked onto an otherwise indeterminate system for shifting losses. It is integral to the definition of tortious wrongdoing. Second, the idea that courts would set up floodgates precisely at the point at which substantive standing rules block tort claims is inexplicable on the terms of the many allocation theories that deem the foreseeable causing of harm to another through substandard conduct as the appropriate trigger for loss shifting. As we have seen, substantive standing rules exclude liability even for foreseeably caused losses. Third, there are many options for floodgate devices that are clearer and better motivated than substantive standing requirements. Why not instead screen out claims for minor harms? Why not set damage caps for some or all classes of claims?

We saw above that the abandonment of loss-allocation theories in favor of a wrongs-and-recourse view of tort explained the apparent anomaly of torts without losses. The same is true for instances of losses that do not generate successful tort claims because of substantive standing rules. Each such rule is a requirement that the defendant's conduct constitute not merely a wrong in the sense of antisocial conduct, nor a wrong to someone else, but a wrong relative to the plaintiff. To demand that a trespass plaintiff have the requisite possessory interest is to demand that she prove that the defendant *trespassed against her*. The rules of fraud, libel, and malpractice law similarly require the plaintiff to prove, respectively, that the defendant *defrauded her*, *libeled her*, or *committed malpractice on her*.[217] The duty-imposing norms of tort law are relational norms: they enjoin persons from acting toward certain other persons in certain ways. Substantive standing rules ensure that rights of action are generated only in those who have been wronged.

C. *The Diversity of Tort Remedies*

Allocative views treat tort suits as claims by loss-sufferers to be entitled to off-load their losses onto others. In so doing, they conflate the issue of the plaintiff's *remedy*—to what relief is a successful plaintiff entitled—with the issue of the plaintiff's *right of action*—under what circumstances does a tort plaintiff have a valid claim, such that she is in a position to obtain some sort of relief? Just as many tort plaintiffs can prevail without proving losses, many obtain a remedy that cannot be cogently depicted as one that shifts losses. Even compensatory damages—the cornerstone of loss-allocation theories—often involve something other than the transfer of a loss from plaintiff to defendant.[218] Moreover, tort law operates in conjunction with

217. *See supra* notes 110–21, 140, 143 and accompanying text.

218. *See* Zipursky, *supra* note 8, at 88 ("The availability of compensatory damages to one who can establish a rights invasion does not necessarily indicate that harm is the basis of a right to recourse. Rather, it merely reflects acceptance of the view that compensation for the harm caused is typically an appropriate form of recourse for those whose rights have been invaded.").

remedies apart from money damages. The question of whether a plaintiff has a valid tort claim is distinct from the question of what sort of remedy she is entitled to if she does have a claim.

In tort suits, fact finders are asked to award fair or reasonable compensation in light of the harm suffered by plaintiff, and in so doing, they are guided by the idea of making the plaintiff whole.[219] Thus, they are being asked to select a financial sum that is *in some sense equivalent* to the loss suffered by the plaintiff. But in what sense? Imagine a successful plaintiff who proves she has suffered a broken leg and therefore incurred medical expenses, lost wages, pain and suffering, and lost enjoyment of life. An award of money damages to this plaintiff in principle should cover any monetary debts incurred by the plaintiff in connection with her injuries. She may also have suffered other forms of setback—for example, lost economic opportunities or property damage—that can and ought to be rectified with money. Finally, she will have experienced setbacks that cannot be rectified as such but the impact of which money can help ameliorate. While the first two aspects of compensatory damages are arguably characterized as allocative, the third is not. And yet the third is pervasive in tort law. A lost limb, a damaged reputation, being rendered paraplegic in a car accident—all of these support payments that *compensate for* a kind of harm rather than make good on a debt or loss.

Quite apart from the question of whether allocative theories can accommodate the concept of compensatory damages, they plainly cannot accommodate punitive damages. The standard "under-deterrence" explanation provided by deterrence theorists fails entirely to explain the rules for when punitive damages will be awarded, as well as the amounts in which they are awarded.[220] The same goes for accounts of punitive damages that cast them as compensatory of losses suffered by persons not before the court.[221] Finally, tort law of course makes available other remedies,

---

219. *See supra* notes 155, 202 and accompanying text.

220. *See* Mathias v. Accor Econ. Lodging, Inc., 347 F.3d 672, 677 (7th Cir. 2003) (Posner, J.) (suggesting that punitive damages are awarded to induce plaintiffs with modest compensatory claims to sue, and to encourage litigants to uncover hidden wrongs, thereby promoting the private prosecution of conduct that would otherwise go unsanctioned). On this theory, one should never see an award of punitive damages in cases of tortious conduct causing substantial harms, nor should courts permit punitive damages in cases of open and obvious misconduct. The law allows punitive awards in both kinds of cases. *See generally* Zipursky, *supra* note 181, at 106–07 (criticizing allocative theories of punitive damages).

221. *See* Catherine M. Sharkey, *Punitive Damages as Societal Damages*, 113 YALE L.J. 347, 392 (2003) (suggesting that punitive damages can help compensate persons injured by the defendant's conduct but not before the court). If this sort of approach were true to the law, courts would not insist on grave misconduct as a threshold for a punitive award. After all, merely careless conduct that causes harm to others who are not pursuing a claim creates a basis for awarding extracompensatory damages.

including declaratory and injunctive relief, which allocative theories simply overlook.[222]

Once again, features of tort law that cause headaches for loss-allocation theories present no problem for wrongs-based theories. To understand a tort as conduct causing a loss to another under circumstances that call for the loss to be shifted is to draw a definitional linkage between substantive and remedial law: The wrongful causing of a loss entails relief in the form of loss shifting. By contrast, to understand a tort as a wrong that generates a right of action in its victim leaves the issue of remedies open. To be sure, the remedy of compensatory damages is perfectly explicable as the standard way in which the law allows a plaintiff to respond to a wrongdoer. Under tort law, recourse typically takes the form of the plaintiff being permitted to exact some money from the defendant because of what the defendant did to her,[223] and in many cases, the plaintiff stands to exact a quantity of money from the defendant in an amount equal to the financial losses or debts she has incurred or will incur because of the wrong done to her.[224] Yet this is only one aspect, or head, of compensatory damages, and even for negligence cases arising from car accidents and medical malpractice—let alone cases of battery, nuisance, and libel—this purely financial sense of compensation is often secondary to compensation designed to ameliorate pain and suffering and emotional distress.[225]

If compensatory damages make sense as a form of redress, so too do remedies such as punitive damages and injunctive relief. Indeed, there is nothing remotely surprising about the idea that a victim of a particularly malicious or willful wrong would be entitled to ask the court for permission to *be punitive* in her response to the defendant. This is why punitive damages are also called "vindictive damages."[226] Although "make whole" is the default measure for monetary damages, courts in cases like *Jacque* and *Huckle* have seen fit to relax the default rule and permit the plaintiff to go

222. *See, e.g.*, Pardee v. Camden Lumber Co., 73 S.E. 82, 85 (W. Va. 1911) (granting an injunction against the defendant's trespass).

223. *See supra* note 150 and accompanying text. One could conjure up a different default conception of recourse. For example, the law might allow successful tort plaintiffs to demand prison time for tortfeasors even in the absence of a criminal prosecution. On the other hand, it is not difficult to see why the law has tended to favor recourse in the form of monetary compensation. Among other things, it better suits the generally lower threshold for wrongdoing found in tort as compared to criminal law. It also lowers the stakes (as compared to more visceral forms of punishment) associated with the provision of recourse, thereby discouraging further cycles of dispute among tortfeasor and victim.

224. *See supra* note 155 and accompanying text.

225. Goldberg & Zipursky, *supra* note 122, at 1140–41.

226. *See, e.g.*, BRYAN A. GARNER, A DICTIONARY OF MODERN LEGAL USAGE 718 (2d ed. 1995) (naming "vindictive damages" as one of several variants of "punitive damages").

beyond that limit because of the egregious way in which the defendant wronged her.[227]

In nuisance, trespass, libel, and fraud, courts often grant injunctive relief.[228] In doing so, they are empowering the plaintiff to exact some performance from the defendant through the legal system. If, for example, a defendant has wronged the plaintiff—and continues to wrong the plaintiff—by unreasonably interfering with her use and enjoyment of her land, then the plaintiff can ask for the state's assistance in forcing the defendant to shut down its interfering activity.[229] Here, the remedy is a stopping of the wrong and has nothing to do with the allocation of a loss.

D. Predicate Injuries and Parasitic Damages

Tort law has long drawn, and continues to draw, a distinction between predicate injuries and parasitic damages. Consider the Supreme Court's *Buckley*[230] and *Ayers*[231] decisions. *Ayers* holds that a railroad worker who suffers a physical injury such as pleural thickening from exposure to asbestos caused by his employer's negligence can recover from the employer for the injury itself and also for the fear that she will develop cancer from the same asbestos exposure that caused the physical injury.[232] However, *Buckley* holds that a person who has been negligently exposed to asbestos and consequently has fear of cancer without any present physical ailment cannot recover for that fear.[233] The two cases are distinguished as follows: In *Ayers*, the defendant is being held liable for negligently causing pleural thickening, and the measure of compensatory damages is make-whole, which includes compensating for related emotional harm, including fear of cancer.[234] In *Buckley*, the plaintiff suffered no injury that the defendant had a duty not to cause because there is no general duty to take care against causing foreseeable emotional harm.[235] Hence, there is no cause of action for negligence, and no recovery

---

227. Jacque v. Steenberg Homes, Inc., 563 N.W.2d 154, 160–61 (Wis. 1997); Huckle v. Money, (1763) 95 Eng. Rep. 768, 768–69 (C.P.) (U.K.).

228. *See* DOUGLAS LAYCOCK, THE DEATH OF THE IRREPARABLE INJURY RULE 4–5, 38, 80, 166–67 (1991) (describing courts' increasing openness to equitable remedies and listing areas of the law where this trend is noticeable).

229. KEETON ET AL., *supra* note 12, § 89, at 640–41.

230. Metro-N. Commuter R.R. v. Buckley, 521 U.S. 424 (1997).

231. Norfolk & W. Ry. v. Ayers, 538 U.S. 135 (2003). Of course *Buckley* and *Ayers* are applications of a statute—the Federal Employers Liability Act (FELA)—rather than the common law of tort. But the Supreme Court has long emphasized that common law rules inform its interpretation of FELA, and in fact both decisions sit comfortably with state court decisions. *See, e.g.*, Atchison, Topeka & Santa Fe Ry. v. Buell, 480 U.S. 557, 568 (1987) ("Assuming, as we have [in prior cases], that FELA jurisprudence gleans guidance from common-law developments . . . .").

232. *Ayers*, 538 U.S. at 158.

233. *Buckley*, 521 U.S. at 436.

234. *Ayers*, 538 U.S. at 135.

235. *See Buckley*, 521 U.S. at 428–30 (explaining that FELA does not recognize a general duty to avoid causing emotional distress, only a limited duty associated with the common law zone-of-danger rule).

at all. In the terminology of tort doctrine, *Ayers* involves recovery for fear of cancer as damages "parasitic" on the "predicate injury" of pleural thickening.[236] Given the absence of a predicate injury in *Buckley*, there is no recovery.

If tort law is about shifting losses caused to innocent (or less culpable) plaintiffs by tortious actors, this familiar distinction is puzzling. In each case, the defendant's carelessness toward the victim has caused foreseeable losses. The question thus arises: If the fear of cancer in and of itself is a loss that the law appropriately declines to shift, why does tacking this loss on to a distinct compensable loss suddenly render it appropriate for transfer?[237]

Ayers permits us to look at various allocative theories in some detail. If we look at tort law as fundamentally serving the needs of plaintiffs who require compensation, it is troubling. The Supreme Court plausibly rejected pure fear-of-cancer claims in *Buckley* in part because there are many potential asbestos claimants whose latent cancer will eventually become actual cancer.[238] Permitting fear-of-cancer claims would threaten to bankrupt defendants and leave persons who later develop serious physical illnesses with no possibility of recovery.[239] This is a powerful policy argument against recovery for fear-of-cancer in this context. Yet, if adopted, it implies there should be no recovery for fear in *Ayers* either: the presence of pleural thickening does nothing to undercut the limited-fund rationale for declining to shift the costs of that fear to the defendants.

Suppose, instead, we look at tort law as fundamentally about forcing actors to internalize the costs of their activities in cases where care should have been taken. This would seem to imply there should be recovery in a case like *Ayers* because fear of cancer is a genuine social cost and the employer's conduct was careless. But then the same reasons would apply

---

236. *Ayers*, 538 U.S. at 148–49.

237. The same distinction is at work in tort law's embrace of the "eggshell skull rule." If an actor commits an offensive contact battery against a victim, he is potentially on the hook for all the harms that flow from the battery, even those that could not have been foreseen. Vosburg v. Putney, 50 N.W. 403, 404 (Wis. 1891). By contrast, a touching that does not amount to a battery—say, an ordinary tap on the shoulder that happens to cause catastrophic harm—is entirely nonactionable, being neither an offensive-contact battery nor an instance of negligence. *See* KEETON ET AL., *supra* note 12, § 9, at 39–42 ("Absent expression to the contrary, consent is assumed to all those ordinary contacts which are customary and reasonably necessary to the common intercourse of life, such as a tap on the shoulder to attract attention . . . ."). Why does a loss that is too unexpected to be one for which an actor can be held responsible suddenly become appropriate for transfer because it is connected with a distinct and actionable loss? None of the principles or policies that determine which losses should be shifted to defendants—compensation, deterrence, or justice—seems to carry with it grounds for different treatment based on the linkage of an unexpected harm to a less unexpected one.

238. *See Buckley*, 521 U.S. at 435 (suggesting that a different outcome could, given "the large number of those exposed and the uncertainties that may surround recovery," lead to "unlimited and unpredictable liability").

239. *See id.* at 435–36 ("We do not raise these questions to answer them (for we do not have the answers), but rather to show that general policy concerns of a kind that have led common-law courts to deny recovery for certain classes of negligently caused harms are present in this case as well.").

equally to *Buckley*. Within an efficient-deterrence framework, the answer to the question of whether the cost of fear of cancer should be shifted to the negligent actor cannot hinge on whether there was or was not pleural thickening. Of course, there might be reasons for thinking that this kind of cost should not be among the accident costs the law should aim to minimize. And so perhaps there should not be recovery in cases like *Buckley*. But then there is no reason to believe that this sort of cost should count when the defendant is being held liable for it on the basis of having caused pleural thickening.

The same problem would seem to arise even for a noninstrumental allocative view, such as Ripstein's variation on corrective-justice theory. He argues that tort law sets a standard of conduct that reconciles liberty and security, imposing upon individuals a duty to take care not to interfere with others' well-being in a manner that deprives those others of primary goods.[240] Tortiously harming someone is interfering with a primary good, and that is why one is obligated not to do it, and why there is a rights invasion if it is intruded upon.[241] Ripstein suggestively asserts that a loss that transpires when a defendant takes an unjustifiably large risk of harming the plaintiff is a loss "owned" by the defendant.[242] Under this approach, to determine whether a defendant "owns" a plaintiff's fear of cancer would require a judgment as to whether the freedom from fear of disease caused by conduct heedless of such fear falls inside or outside the category of primary goods. If it is outside, then *Buckley* is explicable, and the loss should not be shifted because it is not the defendant's responsibility. But then *Ayers* should come out the same way. Conversely, if *Ayers* is rightly decided, that must be because the freedom from such setbacks is inside the package of primary goods, in which case *Buckley* should come out for the plaintiff.

Where loss-allocation theories stumble, wrongs-based theories do not. Indeed, the distinction between predicate injury and parasitic damages flows quite easily from the analysis provided in the prior Part. Whether the plaintiff enjoys a right of action against the defendant in light of what the defendant has done to her is distinct from the question of the remedy to which she is entitled should there be such a right. The former turns on whether the defendant *wronged the plaintiff*. If so, there is a right of action absent any affirmative defense. If not, there is no right of action.

A court can plausibly decide that a duty not to harm someone physically through failure to take care as to their physical well-being is breached by a defendant whose negligent exposure of its employees to asbestos causes one

240. RIPSTEIN, *supra* note 50, at 53–58.

241. *See id.* at 273 ("[A] focus on primary goods gives priority to protecting the capacities for exercising important liberties. It also demarcates risks that are taken from those that merely arise, by determining which risks are always to be held in common.").

242. *See id.* at 53–58 ("Unreasonable risks belong to those who create them; as a result, the injuries that result from unreasonable risk imposition belong to the injurers. Since they are the injurer's problem, the injurer must make them up.").

of them to suffer pleural thickening. This is because pleural thickening can fairly be regarded as genuine physical harm and as such falls within the scope of the duty of care owed by the employer to employees. That injury therefore qualifies as the ground for a right of action. Once there is such a ground, the question arises as to what remedy should be available, and if the answer is compensatory damages, then typically make-whole is the measure. This means that damages for emotional harm are also available. There is a difference between *the kind of impact upon plaintiff, the causing of which by a defendant is a wrong*, and *the kind of harm that figures into the goal of making whole*. That is precisely the difference between predicate injury and parasitic damage.

To be sure, this explanation presupposes that parasitic losses, standing on their own, are not treated as harms that defendants have a duty of care to avoid causing. But, at least as applied to a case like *Ayers*, this presupposition is doctrinally uncontentious. No one can dispute that tort law as it presently stands has adopted and maintained limited duty rules for negligently caused emotional distress—rules that, among other things, relieve employers from any general duty to take care against causing employees emotional distress. Moreover, it is not as if these rules are unmotivated.[243]

E. The Relevance of Action and Agency

Among the most difficult topics in negligence law is the misfeasance–nonfeasance distinction. A defendant who carelessly causes a baby to drown by accidentally knocking her out of a boat will be liable in tort for having done so. A bystander who fails to save a drowning baby because he does not want to get his sleeve wet has committed no tort at all. Negligence doctrine deems this result to follow from certain principles. A very broad duty exists to be careful not to cause others reasonably foreseeable physical harm through one's own conduct, yet there is no general duty to be careful to protect or save someone from harm that comes from another source, even reasonably foreseeable harm.[244] In fact, the misfeasance–nonfeasance distinction is but one example of tort law's attribution of great significance to the connection between the exercise by the defendant (or another person) of his agency and a victim's injury. The distinct treatment accorded to intentional wrongdoing is another. Anxiety about causes that are too *indirect*,

243. *See* Goldberg & Zipursky, *supra* note 181, at 1676–85 (defending limited duty rules in part on the ground that victims can be expected to handle ordinary stresses without "falling apart"). The eggshell skull rule admits of a similar explanation. So long as the defendant has committed an actionable wrong against the plaintiff, a right of action exists. The extent of damages may go far beyond what was necessary to generate an actionable wrong, but that is because there is a remedial rule calling for victims of completed wrongs to be made whole. The causing of this additional increment of losses is not itself a wrong.

244. *See, e.g.*, PROSSER, *supra* note 12, at 190 ("The law has not recognized any general duty to aid a person who is in peril.").

often stationed under the proximate-cause and superseding-cause doctrines within negligence, is yet another.

The significance of the misfeasance–nonfeasance distinction itself is not simply expressed in the law under the heading of "duty." Even where there is liability in a case that does not involve misfeasance, the treatment of the case is different from that of an instance of misfeasance. When the basis of liability is a breach of an affirmative duty to protect the plaintiff, distinctive rules on causation, damages, and rights as to third parties can come into play. For example, victims of breaches of affirmative duties are often treated generously on the issue of causation—a pro-plaintiff feature of the law that, paradoxically, makes courts justifiably wary of recognizing new affirmative duties.[245]

Intentional wrongdoing is likewise treated as categorically distinct from carelessness. When contributory negligence was a complete defense to negligence, it provided no defense to intentional torts.[246] Even with the switch to comparative fault, in most jurisdictions plaintiff carelessness does not provide a ground for reducing the damages payable by an intentional tortfeasor.[247] Punitive damages are generally available against intentional tortfeasors but not those who are merely negligent.[248] Emotional harm and economic harm arising from carelessness are actionable only in special circumstances, yet when the plaintiff depicts the defendant as having acted with an intention to bring about these sorts of harm, courts are much more receptive.[249] Historically, in a case in which a tortfeasor's carelessness toward the plaintiff combined with another's intentional mistreatment of the plaintiff, the former could seek contribution and perhaps indemnification from the latter, but the latter was barred from seeking contribution from the former.[250]

The contribution of actors other than the defendant herself to a plaintiff's injury—including the plaintiff—has also received special

---

245. *See, e.g.*, Affiliated Ute Citizens v. United States, 406 U.S. 128, 153–54 (1972) (holding that in a suit for breach of an affirmative duty to disclose information, it can be presumed that the plaintiff would have relied on the information had it been disclosed); RESTATEMENT (SECOND) OF TORTS § 402A cmt. j (1965) (adopting a rebuttable presumption that a consumer suing for a failure to warn would have heeded the warning had it been given).

246. RESTATEMENT (SECOND) OF TORTS § 481.

247. *See* RESTATEMENT (THIRD) OF TORTS: APPORTIONMENT OF LIAB. § 1 reporters' notes, at 13 (2000) (describing the majority rule). *But see id.* (arguing for a departure from the majority rule); *id.* § 3 reporters' notes, at 43 (treating the question of whether to recognize comparative fault as a defense to intentional torts as a "policy" question for courts to decide on a case-by-case basis).

248. *See* RESTATEMENT (SECOND) OF TORTS § 908(2) (1977) (requiring conduct that is "outrageous, because of the defendant's evil motive or his reckless indifference to the rights of others"); *id.* § 908 cmt. b (acknowledging no punitive damages for mere inadvertence).

249. *See, e.g.*, Ultramares Corp. v. Touche, 174 N.E. 441, 447 (N.Y. 1931) (holding that accountant malpractice causing economic loss is not actionable absent privity or near-privity between plaintiff and defendant, whereas fraud would be actionable without privity or near-privity).

250. *See* RESTATEMENT (SECOND) OF TORTS § 886A(3) (asserting no right of contribution for intentional tortfeasors).

treatment in tort law. In some instances, it has been prepared to treat multiple wrongdoers as "concurrent" causes of a victim's injuries.[251] In others, however, it deems the intervention of a wrongdoer to relieve the more remote actor of any responsibility for a victim's injury, even granted that the remote actor's own wrongful conduct was a necessary condition for the happening of that injury.[252]

Different allocation theorists have different accounts of the misfeasance–nonfeasance distinction, of the distinct treatment accorded intentional torts, and of the role of directness. But for all of them there is a basic problem. Once one decides that a loss is sufficiently foreseeable that a defendant could reasonably have anticipated and avoided or prevented it, it is not clear why features such as the absence of misfeasance or the presence of intentionality on the part of the tortfeasor or some other actor should affect the decision to hold the defendant responsible for the loss. For example, on the issue of superseding cause, the fact that another actor's intervening misconduct has contributed, along with the defendant's earlier carelessness, to the plaintiff's loss seems as if it should merely add a name to the list of potentially liable parties and thereby alter comparative fault allocations. It should not provide a reason to *subtract* from the list of potential loss-bearers the initial tortfeasor. Yet this is precisely what the superseding-cause doctrine calls for, where applicable.[253] Similarly, why should it make any difference to a defendant's ability to invoke the plaintiff's fault to diminish recovery that the defendant's conduct was intentional? If the plaintiff is partly at fault, it seems both equitable and efficient to permit the intentional tortfeasor to diminish her liability by proving that the plaintiff's fault contributed to his injury.

The obvious reply of allocationists on this last point is that intentionality raises the degree of defendant fault and therefore stands as an equitable reason to place liability on the defendant and not the plaintiff. A similar line could be run on affirmative duties and misfeasance; while negligent misfeasance is less blameworthy than intentional harming, it is more blameworthy than the breach of an affirmative duty to protect someone. The common law, the allocationist might argue, simply took these views to an extreme for reasons of administrative simplicity. If it had operated with our nuanced system of comparative fault, rather than with simplistic all-or-nothing rules such as contributory negligence and superseding cause, courts would have permitted all of these issues to go to the fact finder, which could then have apportioned greater liability to intentional tortfeasors, somewhat less to

251. *See* KEETON ET AL., *supra* note 12, at 268 (recognizing that under the law of joint tortfeasors, multiple wrongdoers may be held jointly liable for injuring a victim).

252. *See id.* at 302 (explaining that a defendant may be relieved of responsibility if his tortious conduct is superseded by the subsequent misconduct of an independent actor).

253. KEETON ET AL., *supra* note 12, at 301–02; *see also infra* notes 259–62 and accompanying text.

negligent actors, and less still to those who merely breach affirmative duties. A version of this approach is advocated by the soon-to-be-adopted Restatement (Third) of Torts, which proposes that courts use apportionment to evade issues of directness, superseding cause, and (in many cases) affirmative duties.[254]

Unfortunately, this line of argument amounts to a concession on the part of loss-allocation theorists of their inability to grasp the reasons for the common law to be organized the way it has been organized—and continues to be organized in most jurisdictions—on issues relating to the link between a defendant's agency and his liability. It is clear that the law would require quite dramatic change to conform to the allocationists' views on these issues. Perhaps it should, but that is not our question. The question is whether the allocationists can explain this terrain.

A wrongs view, by contrast, carries with it the analytical resources to understand the ways in which tort law has tended to address these issues. The tort, in negligent-misfeasance cases, is the *doing* of the physical harm through careless conduct. The driver who runs into my car negligently has damaged my car—the wrong is the negligent damaging. The driving instructor who negligently fails to stop his student from crashing his car into mine has *not damaged my car*—he has failed to stop his student from damaging it. He may or may not have a duty to protect me from injury at the hands of the student, but even if he does, he has not inflicted damage on me in the way that his student has. The nature of each of the two wrongs is quite different. The student has done something wrong to me; the instructor has failed to protect me from being wrongfully injured by another.

The significance of indirectness can similarly be accounted for on a wrongs view. Consider the *Allbritton*[255] decision from the Texas Supreme Court. A manufacturer's carelessly made product started a fire at an industrial facility.[256] After the fire was extinguished, the plaintiff's supervisor was directed to block off a certain valve. The plaintiff asked to accompany him. Upon reaching the site of the valve, however, the two were informed that it no longer needed to be blocked off. They then left the scene by means of a shortcut along a slippery elevated pipe rack. The plaintiff fell off the rack and was injured. The majority ruled that so much had happened between the realization of the risk of fire contained in the defendant's carelessness and the plaintiff's harm that it no longer made sense to treat the defendant's conduct as *inflicting* a harm upon the plaintiff.[257] The defendant's

254. *See* Goldberg & Zipursky, *supra* note 183, at 1211–12 (explaining and criticizing the Restatement's position).

255. Union Pump Co. v. Allbritton, 898 S.W.2d 773 (Tex. 1995), *abrogated in part by* Ford Motor Co. v. Ledesma, 242 S.W.3d 32 (Tex. 2007).

256. *Id.* at 774.

257. *Id.* at 776.

carelessness was a necessary condition of her harm but not one of its proximate (or "legal") causes.[258]

Superseding cause is a variation on the same idea. Sometimes, the agency of a third party alters the sequence of events running from the defendant's conduct to the plaintiff's injury in such a way as to render it implausible to see the defendant as having wronged the plaintiff, even if the defendant acted negligently and even if such conduct was a cause-in-fact of the plaintiff's foreseeable injury. The question is not whether the plaintiff has a loss and the loss was foreseeably caused by the tortious conduct of the defendant. The question is whether the conduct–consequence–injury sequence hangs together as a wronging by the defendant of the plaintiff. Intentional intrusions by a third party can sometimes destroy this sequence.[259] Such was the case, according to the Third Circuit, when the owner of the World Trade Center sued the manufacturer of the fertilizer used by the terrorists who bombed it in 1993.[260] Even if it was foreseeable to the manufacturer that its product might be converted by determined terrorists into a bomb, the bombing was their doing, not the manufacturer's.[261]

As we noted above, tort law today is often prepared to apportion liability among two or more parties, each of whom contributed to a plaintiff's injury. For an important realm of negligence and products-liability cases—typically accident cases—there is a synergy between independent risks of harm created by multiple defendants. In these cases, the injuring of the plaintiff may be a concurrence of independent wrongs committed by multiple parties. This is the way that maritime law has tended to conceive of

---

258. *See id.* ("[T]he forces generated by the fire had come to rest when she [Allbritton] fell off the pipe rack. The fire had been extinguished, and Allbritton was walking away from the scene. . . . [T]he pump fire did no more than create the condition that made Allbritton's injuries possible. We conclude that the circumstances surrounding her injuries are too remotely connected with Union Pump's conduct or pump to constitute a legal cause of her injuries.").

Even one who agrees with the *Allbritton* dissent that the defendant's carelessness was a proximate cause of the plaintiff's injury, *see id.* at 785 (Spector, J., dissenting), presumably would allow that some weaker causal linkage of carelessness to injury would defeat the notion that the defendant had done wrong to the plaintiff. Such might be the case, for example, if the plaintiff had safely returned from the attempt to fix the valve to a staff break room, then, because she was tired from fighting the fire, accidentally burned her hand on a hot plate.

259. *See, e.g.*, GOLDBERG, SEBOK & ZIPURSKY, *supra* note 155, at 296–97 (explaining the doctrine of superseding cause and giving examples of situations "in which third-party misconduct intervenes as a necessary step in the sequence of events leading from the defendant's breach to plaintiff's injury").

260. Port Authority of N.Y. & N.J. v. Arcadian Corp., 189 F.3d 305 (3d Cir. 1999).

261. *See id.* at 314 ("[T]he raw ammonium nitrate and urea sold by defendants were not explosive until the terrorists purposefully manipulated and adulterated them by mixing them together with additional chemicals such that they were transformed into energized materials that could be incorporated into an explosive charge. The danger to plaintiff was presented not by the raw materials, but by a bomb that incorporated the raw materials after they had been substantially altered.").

accidents at sea; *Kinsman* offers a famous contemporary example.[262] The rejection of all-or-nothing liability rules and the application of comparative-fault principles makes sense in these sorts of cases, not because intervening wrongdoing is never a reason to block the imposition of liability on a remote wrongdoer, but because they present a special situation in which two or more actors engage in conduct that is already careless as to the victim irrespective of the prospect of wrongdoing by another.[263] When *D1*'s conduct is careless toward a victim, but just happens to cause injury to the victim by virtue of intervening wrongful conduct by *D2*, the intervention of *D2*'s careless conduct toward the victim provides no reason to deny that *D1* has wronged the victim. In the same vein, the replacement of contributory negligence by comparative negligence can be understood as a recognition that an injuring of a plaintiff can sometimes be simultaneously an instance of the plaintiff negligently injuring herself and the defendant negligently injuring her.

Contributory negligence and comparative fault do not apply to intentional torts for the same reason that superseding cause has been and continues to be a significant force in negligence doctrine. Tort law does not depict the event of a plaintiff's being injured as a careless wronging if there was someone who deliberately set that injury as a target of her plans, set out to reach that target by intentional conduct, and reached that target. In this situation, the victim's injuries are the actor's responsibility and no one else's. Others' fault—including the plaintiff's—may have been a necessary condition for the success of the intervening actor's plan. But the law is not concerned here merely to identify grounds that warrant or cut against the shifting of losses from a plaintiff to others. It instead is concerned with determining when one actor has wronged the plaintiff. In so doing, it plausibly identifies different forms or classes of wrongings, subject to different definitions and defenses.

VI. Wrongs and Recourse

Part IV argued that it is possible for tort law to be understood in terms of a meaningful concept of private wrongs without falling into various fallacies that jurisprudential scholars had taken to be fatal. Part V went further, arguing that, whereas loss-based accounts of tort law fail to explain many basic tort doctrines, wrongs-based accounts can make sense of them. And yet we suspect that, for many readers, loss-based accounts will retain some appeal. Why?

As we noted at the outset of this Article, a virtue of such accounts is that they have something concrete to say about the point of having tort law. It is, they say, law for shifting losses from persons who should not have to bear

---

262. *See In re* Kinsman Transit Co., 338 F.2d 708, 726–27 (2d Cir. 1964) (applying comparative-fault principles in attributing liability to multiple defendants).

263. *See* Goldberg & Zipursky, *supra* note 183, at 1236–37 (explaining the distinctive nature of "concurrent negligence" cases).

them to those who (for whatever reason) should. The attraction of this account of tort law's usefulness has been sufficiently strong to encourage aggressive theorists—James and Calabresi, for example—to abandon any effort to defend their theories as interpretively plausible and to shift instead to a forthright call for the revision or elimination of doctrines that prevent tort from operating more satisfactorily as a scheme of loss allocation. But even more interpretively oriented theorists tout as a virtue of loss-allocation theories that they have a story to tell about the role tort law plays in our legal system. By the same token, they will insist or suppose that there is no such case to be made for a law of wrongs. In the absence of a modern bureaucratic state, a law of wrongs perhaps was useful in channeling the passions of those keen to act on their vengeful dispositions and in thereby keeping the peace. But we moderns are well past the point of needing a body of law that indulges base instincts of this sort.[264] A modern state should concern itself with "real" problems, such as the delivery of compensation to those in need, or the deterrence of antisocial conduct, or the shifting of losses to those who are responsible for them.

In light of these concerns and arguments, it surely has not helped the cause of wrongs-based views that their most visible modern proponent—Professor Weinrib—has sometimes tied his particular account of torts as wrongs to a starkly formalist jurisprudence that rejects as off-point any inquiry into tort law's worth.[265] Tort law, he has said, can only be understood for what it is, not what it does.[266] That the leading wrongs-based theory claims as one of its chief virtues indifference to the point of having tort law suggests that there really is nothing much to be said for any such theory.[267]

In contrast to Weinrib, as well as to loss-allocation theorists who criticize him, we believe that a wrongs-based account of Torts connects elegantly to a plausible and appealing account of tort law's place in our legal system. Simply put, it is legitimate and useful for a modern liberal-democratic state to afford the victims of certain wrongs an avenue of recourse against those who have wronged them. *Civil recourse* is what the state delivers by having tort law.[268] Moreover, tort law as a law of wrongs guides conduct and protects individuals against mistreatment by others.[269]

It is perhaps uncontentious to assert that many victims of libel, battery, or negligence wish to have some means of responding to those who have

---

264. *See, e.g.*, Emily Sherwin, *Compensation and Revenge*, 40 SAN DIEGO L. REV. 1387, 1400 (2003) (articulating this view).

265. *E.g.*, WEINRIB, *supra* note 51, at 45.

266. *Id.* at 45–46.

267. *See, e.g.*, Robert L. Rabin, *Law for Law's Sake*, 105 YALE L.J. 2261, 2263 (1996) (reviewing ERNEST J. WEINRIB, THE IDEA OF PRIVATE LAW (1995)) (dismissing Weinrib's treatment of tort law because of its formalism).

268. *See supra* text accompanying notes 125–29.

269. *See supra* note 7 and accompanying text.

defamed, assaulted, or carelessly injured them. But the issue is not whether some people desire redress against those they perceived as having wronged them. Rather, the question is whether the state does well to provide an avenue for such recourse. Moreover, the explanation for why there is value in providing recourse cannot reside principally in the idea that those who suffer injuries are entitled to be made whole. This would just be a reworking of a compensation-driven allocationist view in the language of recourse theory. Nor can the case for the value of providing recourse depend primarily on the idea that it will help deter accidents and injuries. This would be mere rehash of a deterrence-based allocation theory. The same goes for the idea that a system of recourse will allow for a fairness-based or responsibility-based shifting of costs from victims to wrongdoers—a rendition of recourse theory that reduces it to justice-oriented allocation theories.

In fact, the notion that there is value to the state's provision of civil recourse rests on a different idea than any of these, one that comes near to being captured by the hoary common law maxim: "[W]here there's a right, there's a remedy."[270] By recognizing relational duties of noninjury, tort law identifies and enjoins actions that constitute mistreatments of others. In turn, it identifies and confers on each of us a set of rights not to be mistreated. When one of these directives is violated—when a tort is committed—the victim of the mistreatment not only has suffered a setback in the eyes of law but is also recognized as having a legitimate grievance against the wrongdoer. The defendant has violated her legal rights and that violation entitles her to a remedy as against the wrongdoer.

One can imagine this remedy taking various forms, including a legal privilege of self-help. However, self-help is for the most part forbidden by the modern state.[271] Nevertheless, a victim's awareness that responsive aggression is prohibited does not put to rest the grievance. A person rendered paraplegic by the negligent driving of another, beaten by another, or humiliated by a libel has endured a rights violation; the injury and the grievance are real. The principle of civil recourse states that the victim of a legal wrong is entitled to some official avenue of recourse against the wrongdoer.

When courts embrace the *ubi jus ibi remedium* maxim, they tend to be articulating the gist of the principle of civil recourse and articulating it in a performative manner. To assert it is to say something to the following effect: "The plaintiff, having shown that what was done to her was a violation of her right not to be treated in a certain way, is entitled to, and therefore shall have, a remedy against the wrongdoer." Courts providing rights of action to

---

270. *See, e.g.*, Douglas Laycock, *How Remedies Became a Field: A History*, 27 REV. LITIG. 161, 168 (2008).

271. *See* Douglas I. Brandon et al., *Self-Help: Extrajudicial Rights, Privileges and Remedies in Contemporary American Society*, 37 VAND. L. REV. 845, 853 (1984) ("Modern courts generally are dubious of or hostile to self-help.").

victims of legal wrongdoing are recognizing themselves as fulfilling a political obligation to provide victims some means of civil response to having been wronged.[272] A plaintiff's entitlement to a right of action against a tortfeasor thus involves obligations of both the tortfeasor and the state. The state recognizes itself as obliged to empower the plaintiff to act in some manner against the defendant and acts on that obligation by permitting the plaintiff to exact damages or have the defendant enjoined against performing certain acts. Once the state has so acted (by entering a judgment), the defendant incurs a legal obligation to the plaintiff.

The normative idea at the root of the principle of civil recourse is not dependent upon social-contract theory, but it is illuminated by it. One can understand the state's obligation to empower the plaintiff who chooses to sue as part and parcel of a larger bargain that exists between the individual and the state. An individual relinquishes the raw liberty to respond aggressively to having been wronged and receives in return a certain level of security against responsive aggression by others, plus the assurance that a civil avenue of redress against wrongdoers will be supplied. At a higher level, this can be depicted as a bargain citizens make among themselves in choosing to have a state at all and reciprocally agreeing to have a state so long as it conforms its exercises of power to certain domains, preserves various domains of liberty, and recognizes certain forms of rights.

When the social-contract metaphor is stripped away, the idea of civil recourse becomes clearer. It is a political commitment to the following effect: Individuals who are able to prove that someone has treated them in a manner that the legal system counts as a relational, injurious wrong shall have the authority to hold the wrongdoer accountable to him. This commitment is not founded, in the first instance, on instrumental concerns but on political and moral ones. Part of the state's treating individuals with respect and respecting their equality with others consists of its being committed to empowering them to act against others who have wronged them. Relatedly, a legal and political order that respects an individual's right not to be treated in a certain manner cannot permit persons to invade such rights with impunity; forbidding responsive aggression without providing any avenue of private redress is a way of permitting rights invasions with impunity. This is particularly true with regard to wrongs that are not crimes or regulatory infractions, but it is even true of wrongs that are crimes or infractions, given that enforcement by the state of criminal and regulatory law is discretionary. Our system affords a victim a civil right to hold a wrongdoer answerable to her. A legal right of action in tort against the wrongdoer is that right.

---

272. *See* Goldberg, *supra* note 183, at 563 (connecting Justice Marshall's invocation of the *ubi jus* maxim in Marbury v. Madison, 5 U.S. (1 Cranch) 137, 163 (1803), to the longstanding notion that a constitutional government is under a duty to afford its citizens the protection of the laws, including laws that provide recourse for wrongs done to them).

By focusing on the significance of civil recourse, we do not mean to overlook another valuable aspect of tort law. As a law of wrongs, it guides conduct by reference to, gives recognition to, and enforces duties not to mistreat others. While the idea of civil recourse helps to explain in what sense tort law is committed to holding actors responsible or accountable to those whom they have wronged (when such a demand is made by the victim), the idea of legal wrongs helps to explain in what sense tort law recognizes responsibilities not to mistreat others in various ways.

We have elsewhere explained why the deterrence model that treats tort law as a set of liability rules does not fully capture the conduct-guiding capacity of the common law of torts.[273] The legal directives of tort law work indirectly. They spring from social understandings, where norms of conduct are to different degrees inchoate. Consider the following three levels at which there are various levels of knowledge of the wrongs of tort and there exists pressure to give definition to the wrongs of tort.

Pushed to adjudicate a plaintiff's demand for redress in a tort action, a judge or jury is often required to determine whether what the defendant did to the plaintiff was a wrong. Members of a community understand that the question answered in such circumstances is a question about whether the defendant's conduct was a wrong of a sort the legal system recognizes—a legal wrong.[274] If the issue is forced through appellate courts, a potentially more general and enduring announcement is made as to whether the conduct in question is a legal wrong, or at least whether it is plausibly regarded as such by a jury. Outside the context of litigation, individuals and companies will ask their lawyers whether a proposed course of conduct is permissible or whether the legal system counts it as an impermissible way of treating others. Can a certain newspaper article be published? Must patients be given certain warnings? If a product with a certain sort of risk inherent in its design were to injure a consumer, would that count as a tort? Must I reveal all the problems of my car or house before I sell it?

Most individuals do not have lawyers for their daily lives, and most decisions of businesspersons and professionals are made without the advice of lawyers, of course. Yet individuals and businesses know a great deal about what they may and may not do and what they can and cannot reasonably expect others to refrain from doing to them, and they use such knowledge

273. *See* John C.P. Goldberg & Benjamin C. Zipursky, *Accidents of the Great Society*, 64 MD. L. REV. 364, 395–98 (2005) [hereinafter Goldberg & Zipursky, *Accidents of the Great Society*] (arguing that deterrence models that fail to recognize the influence of social norms and institutions are inadequate); John C.P. Goldberg & Benjamin C. Zipursky, *The Moral of* MacPherson, 146 U. PA. L. REV. 1733, 1841–42 (1998) [hereinafter Goldberg & Zipursky, *The Moral of* MacPherson] (arguing that a relational conception of duty—one that incorporates an actor's moral responsibilities and legal duties—will permit law to build on and reinforce ordinary judgments about obligations of care).

274. GOLDBERG, SEBOK & ZIPURSKY, *supra* note 155, at 3.

in deciding what to do.[275] People know not only that it would be wrong to advertise a car for sale as having only 22,000 miles on it when it has 122,000; they know not only that lying is considered a moral wrong under the widely shared morality of contemporary society, but that selling something through deceitful representations is a legal wrong. Companies know that if there is a latent defect in a product that injures someone and that their sale of the product caused the injury, they can be held responsible for the injuring. Individuals know that if they have been subjected to excessive force by a police officer or beaten up by a nightclub bouncer or molested by a boss at work, their legal rights and not just their moral rights have been violated; police officers, bouncers, and bosses know that these are legal wrongs.

It is impossible to articulate how and why a law of private wrongs is of value without taking seriously that the duties and rights that exist under the law of torts figure in the minds of actors both inside and outside the legal system. Although many suppose that it clarifies to say that tort law has deterrent value (we are skeptical of this), it does so at the cost of oversimplification. It certainly does not follow from the fact that tort law has the capacity to influence conduct that a judge deciding a tort case should figure out how she thinks conduct should be deterred and then shape the rule before her to deter maximally or efficiently. The judge's first task is to ascertain whether the system of legal wrongs articulated in the precedent of the relevant area already speaks to the question at hand, and if so, what it says. To the extent that the dimensions of the tort in question remain underspecified, an application of the concepts and principles that are embedded in the precedent is called for. It will usually require a context-sensitive judgment that is attuned to both conceptual and pragmatic considerations.

Our point in this Article is not to set forth a normative theory of adjudication in the common law or to defend a jurisprudential view about how much is already "in" the common law; we have addressed these questions elsewhere.[276] The point is that part of what gives tort law value is that it is a system of rules contained in common law that articulates legally enforceable norms about how one is obligated to treat others. These norms are, at a number of different levels, grasped by members of the community in such a manner as to guide conduct and generate expectations, both directly

275. Goldberg & Zipursky, *Accidents of the Great Society*, *supra* note 273, at 389.

276. *See, e.g.*, Benjamin C. Zipursky, *Practical Positivism Versus Practical Perfectionism: The Hart-Fuller Debate at Fifty*, 83 N.Y.U. L. REV. 1170, 1211 (2008) ("[T]he conclusion that more just results can be reached . . . by practical perfectionism suffers from an artificial limitation of vision; an advocate of practical perfectionism must be willing to take what he . . . regards as the bad results along with what he . . . regards as the good ones, and must be willing to accept the corresponding alterations in judicial and nonjudicial power.").

and indirectly.[277] We recognize that the point is obvious; the problem is that it is almost blindingly obvious. It has been too easy to overlook the fact that tort law really is about wrongs.

Ironically, it is the pervasiveness and breadth of tort law that partly explains why many tort scholars have failed to countenance the centrality of wrongs to tort theory. When one looks for a theory, one typically wants an account that is unifying and simplifying in various ways; when one looks for a legal theory, one typically wants an account that is able to explain a variety of legal phenomena on the basis of a relatively limited set of values or principles or structures. For this reason, it has been attractive to understand tort law as a system aimed at protecting individuals' bodily integrity and at securing their freedom from accidental physical harm and the enormous financial toll inflicted by such harm. Seeing tort law as accident law permits such an approach. When we add to that—as a vitally important constraint—the need for the law to preserve individuals' liberty to engage in a variety of activities, then (one might think) we have the beginnings of a normative framework for understanding why the law of accidents—primarily negligence law—contains duties not to injure others by taking certain kinds of risks; one begins to understand why damage is an element of the tort of negligence; one starts to generate interesting and (one hopes) fruitful question about how activity levels and bodily security levels are to be accommodated, and how different parts of the law can do so in different ways.

The problem is that the law of negligence pertaining to accidental physical injuries is only one part of tort law. The defamation and privacy torts have nothing to do with this. The same is true of fraud; perhaps even of legal malpractice and the law of nuisance and many other torts. Moreover, the wrong of deceiving (fraud) seems to have little to do with the wrong of interfering with someone's use and enjoyment of property (nuisance), and it is hard to see why the values underlying the shape of the tort of fraud are going to help with nuisance. In other words, it is not as if all one needs to do is handle some other set of torts. Given the diverse array of torts, it is perhaps understandable that many have retreated to an "accidents-law-plus" conception of tort law. The alternative, it would seem, would be along the lines of Jules Coleman's plausible assertion that what counts as "wrongful" in tort law is a matter of convention[278] or Ernest Weinrib's tantalizing (but perhaps unfulfilling) suggestion that the substance of tort law cannot be understood except on its own terms.[279]

---

277. *See* Goldberg & Zipursky, *Accidents of the Great Society*, *supra* note 273, at 395–96 (contending that social norms and institutions play a role in an actors' behavior and expectations).

278. *See* COLEMAN, *supra* note 48, at 334 (explaining that tort law concerns wrongdoing as defined by the relevant and appropriate norms of conduct).

279. *See* WEINRIB, *supra* note 51, at 14–15 (arguing that all private law, including Torts, can only be understood internally).

We have offered the idea of civil recourse and the ideas of relational, legal, injury-inclusive wrongs as unifying features of tort law and tort theory, but in doing so we do not suppose that we have relieved ourselves of the need to discuss the substance of tort law. To the contrary, both for torts within the negligence family, for products liability, and for torts from fraud to libel to trespass, we have taken the view that substantive tort theory is possible and valuable. Scholars like Posner on the one hand and Ripstein and Keating, on the other, have said a great deal about the theoretical underpinnings of the wrong of negligence by taking seriously what "reasonable care" means, and why. To a great extent, their accounts of that tort have focused on a set of activities risking physical injury and property damage to others.[280] In numerous places, we have offered accounts of the wrong of negligence pertaining to physical injury, as well as reasons for thinking differently about the wrongfulness of negligent conduct causing emotional harm. What is needed in tort theory is an appreciation that this sort of interpretive account of the wrong of negligence should be developed for all of the wrongs of tort law, at least if the goal is to understand the law of torts. The fact that some torts pertain to reputational attacks, others to property rights, and others still to protection against false imprisonment and malicious prosecution does not show that tort law is incoherent. It shows that there are many kinds of wronging that the legal system has chosen to recognize as legal wrongs.

VII. Implications

In this final Part, we briefly sketch some of the ways in which a wrongs-and-recourse view can illuminate contemporary debates about tort law.

A. Accident Law Revisited

We have argued that it is a huge mistake to depict tort law as law for allocating accidentally caused losses. The mistake is academic, but not merely so. A case can be made that our society has lost a great deal by wrongly supposing Torts is our legal system's first line of response to accidents and its first line of prevention for accident-inducing conduct. If achieving important compensatory and regulatory goals is really what a government wants to do, it would do best to give up the presumption that tort law stands ready to deliver on these goals. While tort law does permit injured victims to gain compensation and does provide financial incentives for actors to address the potential harmfulness of their conduct,[281] it is a

280. *See, e.g.*, RIPSTEIN, *supra* note 50, at 48–64; Gregory C. Keating, *Reasonableness and Rationality in Negligence Theory*, 48 STAN. L. REV. 311, 311–28 (1996); Posner, *supra* note 60, at 29–44 (all discussing negligence through analysis of "reasonable care" with discussion of activities posing risks to physical injury and property damage).

281. *See, e.g.*, LANDES & POSNER, *supra* note 57, at 312 (asserting that tort law provides incentives that generally promote economically efficient behavior).

remarkably inconsistent, blunt, and expensive tool for these tasks. Other forms of public and private legal arrangements are demonstrably superior in a wide range of cases.

As many before us have observed, products-liability law serves as a good example of the weakness of tort law as a compensatory system.[282] California's early embrace of strict liability for products defects was in part motivated by an expressly allocationist argument that manufacturers were better able than injured consumers to bear the burden of the losses suffered by the victims of their defective products.[283] Not only do they have deeper pockets, they also have the capacity to pass on the costs of injuries through higher product prices.[284] However, products liability law will often make for a poor form of insurance given how expensive and slow the tort system is in delivering compensation.[285]

It is also far from clear that products-liability law can be defended simply by virtue of the marginal contribution to safety it provides above and beyond negligence law, although it surely at times helps make up for spotty safety regulation. As manufacturers understandably complain, juries on the whole tend to be less capable regulatory decision makers than expert administrative agencies, and it is unclear why *ex post*, ad hoc, judgments that a particular product design was too dangerous should form the centerpiece of a rational regime by which to protect consumers from dangerous products.[286] In addition, the randomness as to which injured persons choose to sue, the differing abilities of different plaintiffs to endure and succeed at litigation, and the unpredictability of juries' damage awards significantly muddy any deterrent message products-liability law may be sending.

It would be unfair to blame allocationists for the failings of products-liability law as a system of accident compensation and safety regulation, and that is not what we are doing. But it is fair to suggest, as we are doing, that our society's failure to establish better systems of accident compensation and risk regulation stems partly from a misplaced reliance upon tort law as loss-

---

282. *See, e.g.*, MARK A. GEISTFELD, PRINCIPLES OF PRODUCTS LIABILITY 53–58 (2006) (explaining the inefficiencies of tort compensation compared to first-party insurance in the context of products liability and that the problem of underinsurance does not justify an expansion of tort liability).

283. *See* Escola v. Coca-Cola Bottling Co., 150 P.2d 436, 441 (Cal. 1944) (Traynor, J., concurring) (contending that victims suffering injuries from defective products are typically ill-prepared to meet the consequences of injury, while manufacturers are best positioned to prevent and absorb those risks and costs).

284. *See* RICHARD A. EPSTEIN, MODERN PRODUCTS LIABILITY LAW: A LEGAL REVOLUTION 46 (1980) (noting the argument that the manufacturers can pass premium costs of insurance on to purchasers as part of their general cost of doing business).

285. *See* LANDES & POSNER, *supra* note 57, at 57–58 (asserting that the tort system functions as an "exceedingly costly insurance mechanism" with substantial administrative costs imposed by the time and money spent on litigation).

286. *See* EPSTEIN, *supra* note 284, at 87–88 (arguing that the judicial process involving layperson jurors provides a particularly poor forum for making the difficult choices of appropriate product design and regulations to protect consumers).

allocation law. More importantly, it is valuable to take stock of the radical shortcomings of tort law *as compensatory and regulatory* when thinking about what forms of law might meet our compensatory and regulatory goals as we move forward. In this respect, the critique of allocationism serves as a springboard for a more progressive and flexible approach toward our system's method of dealing with accidents.

The critique of allocative theories is progressive because it punctures the idea that tort law should be a default for dealing with accidents, and thereby directs attention to alternative schemes for dealing with accidents. Conversely, however, our account counsels a more restrained approach to the reform of tort doctrine itself. For insofar as the arguments for revision of tort doctrine are predicated on the idea that certain features of tort law are ill-suited to the job of shifting losses, our view gives reasons for backing off of such revisions. It is no accident that Part IV's list of doctrinal conundrums generates something of a to-do list for those keen to make tort law fit their misguided sense of what it is. If Torts is going to be a decent system of loss allocation, rules limiting liability for economic and emotional harm will need to be relaxed, and doctrines such as superseding cause and punitive damages would have to be abolished. We have not provided prescriptive arguments demonstrating that, all told, these and other features of tort doctrine are worth keeping. However, we have shown that they are well-motivated features of a law of civil wrongs and recourse.

B. Evaluating Tort Reform

Although they are entering their fourth decade, debates over contemporary tort reform have for the most part lacked an important dimension. Pro-reform forces argue that tort law does not do a good job of providing compensation and deterrence, and that whatever it does do is accomplished at too high a price in terms of over-deterrence, wasteful expenditures, and suppression of productive activity.[287] Thus, the argument concludes, legislation is needed to eliminate joint and several liability, cap damages, and immunize certain actors from certain kinds of tort liability.[288] Meanwhile, anti-reform forces argue that tort law is not significantly affecting the cost or availability of goods and services,[289] and is a needed mechanism by which citizens can invite courts to regulate and tax powerful industries that the other branches of government have failed to control. It

287. *See* Joanna M. Shepherd, *Tort Reforms' Winners and Losers: The Competing Effects of Care and Activity Levels*, 55 UCLA L. REV. 905, 914–19 (2008) (summarizing tort-reform proponents' arguments in favor of reform).

288. *Id.*

289. *See, e.g.*, Lucinda M. Finley, *The Hidden Victims of Tort Reform: Women, Children, and the Elderly*, 53 EMORY L.J. 1263, 1270–71 (2004) (arguing that medical litigation has not caused an increase in medical-malpractice-insurance costs).

follows, they say, that there is no need for legislatures to adopt curbs or caps on liability.[290]

Ignored in these arguments is a question brought to the fore by a wrongs-and-recourse view of Torts: the extent to which it is important and valuable for our legal system to provide to claimants, as a matter of right, the ability to bring another person into court to answer for an alleged wrong done to her. By focusing on this issue, we may come to appreciate that our system needs both less and more reform than it has been getting. It may need less reform because the value of providing recourse law to the citizenry, once fully appreciated, may counsel caution in the removal or limitation of tort claims. It may need more reform insofar as the tort system fails effectively to identify wrongs and provide victim recourse for them.

One of the most basic themes in the Anglo-American constitutional tradition—one that dates back to the efforts of the common lawyers to resist Stuart absolutism—is the idea that institutional complexity is both a vital bulwark against oppression and a necessary means by which government can accomplish what it is obligated to accomplish.[291] A liberal-democratic government, the thinking goes, has the best chance of functioning well when it is comprised of distinct branches and offices, and when citizens have multiple access points through which to engage government and each other.[292] Membership in a national or local legislative body, the holding of an executive or judicial office, participation on a jury, voting in fair elections, petitioning and freely speaking on matters of public interest, and—yes—suing in courts: Each provides a distinctive form of political participation and with it political power.[293]

The provision by a government to its citizens of a law of wrongs and recourse embodies and furthers several related liberal-democratic values. In multiple ways, it affirms the significance of the individual citizen. It identifies protected interests that each of us possesses—such as the interest in bodily integrity—and with which each of us must refrain from interfering; to say the same thing, it confers upon each of us duties not to mistreat others in various ways and rights not to be so mistreated.[294] It enables individuals to assert claims as a matter of right without first obtaining the permission or blessing of government officials. It renders wrongdoers specifically answerable to victims rather than to a government prosecutor acting on behalf of the

290. *Id.* at 1313–14.

291. *See* Goldberg, *supra* note 183, at 535 ("The common lawyers argued that each of these complexities was vital to the health of the polity, just as the health of each organ in a complex organism ensures its well-being.").

292. *See id.* at 538–39 (summarizing the common lawyers' arguments in favor of maintaining a complex legal system and refraining from consolidating authority in the executive).

293. *See id.* at 601–02 (discussing the empowerment of a litigant who sues for the redress of a wrong).

294. *See id.* at 607–08 (conceptualizing tort law as conferring a right against injury and a corresponding duty not to injure).

state or the people. In holding individuals accountable based on what they have done, irrespective (in principle) of who they are, it embodies and reinforces a notion of democratic equality—the idea that there is not a class or group of persons who are somehow entitled to mistreat another, "lower" class or group.

Of course there are other institutions within our system that, in different ways, embody and further these values. Our point is that once one sees tort as a law of wrongs and recourse, one sees that it is not merely a regulatory or benefits program but part of the architecture of constitutional government. It is no accident that seminal figures in our constitutional tradition, including Coke, Locke, and Blackstone, deemed individuals to enjoy a *right of recourse* against those who wronged them and deemed governments to be obligated to provide an avenue by which to exercise this right.[295] Because it is only natural for individuals to abuse this right—to see themselves as having been wronged when they have not been, or to seek unadulterated vengeance rather than fair recourse[296]—and because it is vital to protect every member of society against the destabilizing effects of individual and family feuds, government can fairly deny individuals the ability to exercise such a liberty simply as they might wish. But in doing so, government is at the same time obligated to recognize some right of response. The most straightforward way of doing so—and the way in which our system for centuries has actually done so—is by providing a law of wrongs and recourse.

We said above that the adoption of a wrongs-and-recourse view was a two-way rather than a one-way ratchet. To note the political significance of tort law in certain ways rather obviously counsels less in the way of modern tort reform. By contrast, procedural and remedial law are areas for which our view might support more reform. If tort law is for the recourse of wrongs, then it will be important to know whether its definitions of legal wrongs are plausible, and if it is providing meaningful recourse to victims of such wrongs. Where necessary to ward off suits that allege nominal wrongs and injuries that seem unlikely actually to be wrongs and injuries, there is room for judges and legislatures to block or raise barriers to suit. For example, the enactment by the Michigan legislature of a bar to claims for "loss-of-a-chance,"[297] whether wise or unwise from a policy perspective, was probably justified given the dubiousness of the idea that a lost chance for health is really an injury.

Finally, note that the cause of more reform need not be limited to defendant-friendly reform. If tort litigants, or some portion of them, consistently report significant frustration or dissatisfaction with the way in which their claims are handled, then we will have occasion to study whether

295. *Id.* at 534–35, 541, 545–46.

296. *Id.* at 602.

297. MICH. COMP. LAWS § 600.2912a (2000).

streamlined procedures ought to be implemented, perhaps even though it might mean reducing the likelihood that every last piece of relevant information pertaining to the lawsuit will be uncovered during the course of discovery. Likewise, if the expense to plaintiffs of litigation works systematically to make a certain category of claims economically infeasible, then it might be an occasion for legislatures or courts to consider the desirability of measures such as fee-shifting rules.

C. Responsibility and Institutional Questions About Wrongs

Tort has long been thought to raise issues of "institutional design."[298] Usually this question is framed in terms of a question as to whether lay judges and jurors are better or more appropriate policy makers than legislatures or bureaucrats or whether and when an *ex post* form of regulation is superior to *ex ante* mechanisms.[299] Buried or lost in these questions is the central role played by social norms and their interaction with legal standards of conduct. A conception of tort law as a law of wrongs and recourse brings them back to the fore.

Economists have been enthralled with the "discovery" that norms shape behavior in ways that depart from rational-actor models.[300] For students of the common law, this is old news. Tort law has always been about a victim's right to have the state's assistance in holding a wrongdoer accountable, or responsible, for what he did, and the wrongs of tort law, as a matter of formal law and informal legal practice, have tended to track social norms of acceptable and unacceptable conduct. Negligence law instructs jurors to determine whether defendant and plaintiff have exercised ordinary care—i.e., acted in the manner that one would expect a person of reasonable prudence to act under the circumstances.[301] Products-liability law, in some iterations at least, keys the finding of a product defect to consumer expectations.[302] Certain nonharmful touchings will count as batteries—namely those that are widely

298. *See, e.g.*, Robert L. Rabin, *Reassessing Regulatory Compliance*, 88 GEO. L.J. 2049, 2061 (2000) (discussing current controversy over the functioning of the tort system).

299. *See generally id.* (discussing current controversies involving the role of judges and juries and of tort law's approach to the defense of regulatory compliance).

300. *See* Robert C. Ellickson, *Law and Economics Discovers Social Norms*, 27 J. LEGAL STUD. 537, 540–43 (1998) (discussing the recent move in law and economics from studying hypothesized rational actors to studying individuals guided by social norms); Christine Jolls et al., *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1471 (1998) ("Empirical evidence gives much reason to doubt [the] assumptions [of neo-classical economics]; people exhibit bounded rationality, bounded self-interest, and bounded willpower. This article offers a broad vision of how law and economics analysis may be improved by increased attention to insights about actual human behavior.").

301. RESTATEMENT (SECOND) OF TORTS §§ 281–283, 463–464 (1965) (defining the cause of action for negligence and the standards for negligence and contributory negligence).

302. RESTATEMENT (THIRD) OF TORTS: PRODS. LIAB. § 2 cmt. c (1997) ("More distinctly than any other type of defect, manufacturing defects disappoint consumer expectations.").

understood to be inappropriate or offensive, such as an uninvited grope.[303] The wrong of intentional infliction of emotional distress requires an outrageous departure from ordinary standards of acceptable conduct.[304] As these and many other examples attest, the wrongs of tort are definitionally connected to social norms.

And yet the wrongs that count as torts are also positivistically defined by legislatures, courts, and jurors: Tort law does not simply incorporate extralegal standards in an unadulterated form. Jurors can deem conduct careless even if it meets customary expectations as to the care one ought to take—an entire practice or calling can be declared substandard.[305] As negligence law's "objective" fault standard attests, judges can push to identify as wrongful forms of conduct that may often be deemed acceptable or at least not blameworthy in ordinary morality.[306] Courts and legislatures can declare new wrongs (gender discrimination in the workplace)[307] or refuse to recognize as tortious long-recognized wrongs (seduction of another's spouse).[308]

Are courts relatively good at walking the line between articulating and flouting social norms? Would judges who presently think of tort suits as occasions to implement ad hoc solutions to pressing social problems do a better job if they saw themselves as instead interpreting, refining, and fashioning norms of right and wrong conduct? Can jurors be trusted in the theatrical atmosphere of a trial to temper tort rules with common sense? If judges are in fact interpreting social norms when they "make" tort law, are courts in this sense democratic institutions even if the judges are not democratically elected? We do not offer answers to these questions; our point is simply that a wrongs-based theory of tort law demands that these issues be examined and offers a rich theoretical framework for doing so.

A brief look at the development of actionable civil wrongs in American law yields some surprising observations. Under a conventional view of our legal system, it is legislatures that should announce new legal duties and new legal wrongs. Likewise, the fact that interpersonal wrongs are likely to

303. RESTATEMENT (SECOND) OF TORTS § 18 (1965).

304. *Id.* § 46.

305. *Id.* §§ 519–520.

306. *See* MAYO MORAN, RETHINKING THE REASONABLE PERSON: AN EGALITARIAN RECONSTRUCTION OF THE OBJECTIVE STANDARD 58 (2003) ("[T]he reasonable person standard's 'failure' to capture only blameworthy behaviour is actually posited as its distinctive strength. . . . [I]n part this is because this 'failure' to attend to blame is seen as crucial to maintaining the core of the objective standard. In this way, the reasonable person's rigidity in the face of the intellectual shortcomings of the defendant is taken to exemplify its distinctively egalitarian conception of fault.").

307. Civil Rights Act of 1964, Pub. L. No. 88-352, tit. 7, sec. 703, 78 Stat. 241, 255–57 (1964) (codified as amended at 42 U.S.C. § 2000e-2 (2006)); *see also* Meritor Sav. Bank v. Vinson, 477 U.S. 57 (1986) (holding that sexual harassment is a form of sex discrimination actionable under Title VII).

308. *See* Veeder v. Kennedy, 589 N.W.2d 610, 614 nn.3–4 (S.D. 1999) (surveying the status of a cause of action for alienation of affection, which the majority of states have judicially or statutorily abolished).

implicate community values arguably points toward the propriety of state-law rather than federal-law innovation in this area. How ironic it is, then, that the most dramatic and notable developments in the law of civilly actionable wrongs—products-liability law, sexual-harassment law, securities fraud, constitutional torts, and international-human-rights wrongs—have been brought about principally not just by courts, but by federal courts.[309] We offer this observation not by way of criticism, but merely to suggest that our legal system has hardly given up on its commitment to courts as fora for the articulation of legal wrongs and the provision of civil recourse.

Relatedly, there is an important set of questions worth noting as to the respective scope of judicial and legislative authority. It is quite clear that legislatures enjoy the authority to fashion statutory torts—relational wrongs that give rise to private rights of action. This is what statutes like Title VII are all about. They can also define different kinds of wrongs that call for different kinds of enforcement. For example, state consumer-fraud statutes have explicitly adopted a private-attorney-general enforcement model, and they have done so in ways that grant standing to sue to persons who would not be authorized to sue under common law principles.[310] What about the converse question? Do courts enjoy the authority to empower persons other than victims of relational wrongs to obtain remedies for those wrongs? At a minimum, it seems clear that courts are operating at the core of their common law authority when they are articulating relational wrongs and providing remedies to victims of those wrongs. Whether, in the absence of an explicit or implicit statutory grant, they also retain a penumbral authority to deputize private citizens to act as private attorneys general is a question that deserves careful consideration.

VIII. Conclusion

The law of Torts is *not* accident law, nor even accident law plus assault and battery. It is what it purports to be: a law of wrongs. Torts are legal wrongs for which courts provide victims a right of civil recourse—a right to sue for a remedy. There is nothing new or even surprising about these statements; hornbook authors have said it all along. What is newer and more surprising is that it actually means something to say these things. Torts is as basic a subject in our legal system as it purports to be. What stands next to Contracts, Property, and Criminal Law is not accident-law-plus. It is the law of private wrongs. By recognizing torts as wrongs, civil-recourse theory

309. *See, e.g.*, Bivens v. Six Unknown Named Agents of Fed. Bureau of Narcotics, 403 U.S. 388 (1971) (constitutional torts); G.A. Thompson & Co. v. Partridge, 636 F.2d 945 (5th Cir. 1981) (securities fraud); Filartiga v. Pena-Irala, 630 F.2d 876 (2d Cir. 1980) (international human rights); Williams v. Saxbe, 413 F. Supp. 654 (D.D.C. 1976) (sexual harassment); Escola v. Coca-Cola Bottling Co., 24 Cal. 2d 453, 562 (Cal. 1944) (Traynor, J., concurring) (products liability).

310. *See* John C.P. Goldberg, Anthony J. Sebok & Benjamin C. Zipursky, *The Place of Reliance in Fraud*, 48 ARIZ. L. REV. 1001, 1015–18 (2006) (outlining the development of a private right of action for consumer fraud).

permits legal scholars to make sense of and develop further a vast body of concepts and principles central to a general understanding of American law.

Afraid since Holmes's time of the sanctimonious sound of "wrongs," and confronted with modern accident epidemics, scholars have convinced themselves that the subject of Torts is really about accidentally caused *losses*, not wrongs, and that the central task of tort law is to reallocate such losses in the most justifiable manner. Included among them are economists like Calabresi and Posner, corrective-justice theorists like Coleman, and mainstream doctrinal scholars like Prosser and the Reporters for the forthcoming Restatement (Third) of Torts. Without wrongs at the center, however, all of these theories are doomed to fail. Numerous, deeply rooted features of the structure of Anglo-American tort law, as we have shown, render loss-based theories incapable of capturing this body of law. In contrast, a civil-recourse theory that predicates rights of action on wrongs, not losses, comfortably shows how tort law hangs together.

In retrospect, our "retaking" of tort law has required only two simple steps: (1) crafting a conception of legal wrongs and (2) taking seriously the idea of a victim's right to recourse against a wrongdoer. In tort, wrongs are violations of legal norms not to mistreat others in various ways. They are legal wrongs, not moral wrongs. Because the legal norms that set out wrongs are always wrongs as to a particular person or classes of persons, those legal norms go hand in hand with a set of potential victims who will be entitled, in principle, to recourse against their wrongdoers. And the principle of civil recourse is itself familiar, not esoteric, embedded as it is in the *ubi jus* maxim: Where there's a right there's a remedy. Wrongs and recourse run as deep in American law as any of its other elements, and they lie at the core of Torts.

Embattled CEOs

Marcel Kahan and Edward Rock*

In this Article, we argue that chief executive officers (CEOs) of publicly held corporations in the United States are losing power to their boards of directors and to their shareholders. This loss of power is recent (say, since 2000) and gradual, but nevertheless represents a significant move away from the imperial CEO who was surrounded by a hand-picked board and lethargic shareholders. After discussing the concept of power and its dimensions, we document the causes and symptoms of the decline in CEO power in several areas: share ownership composition and shareholder activism; governance rules and the board response to shareholder activism; regulatory changes related to shareholder voting; changes in the board of directors; and executive compensation. We argue that this decline in CEO power represents a long-term trend, rather than a temporary response to economic and political conditions. The decline in CEO power has several important implications, including implications with respect to the possibility of a regulatory backlash against certain newly empowered shareholder groups, future development in Delaware's corporate law, the type of persons who will serve on corporate boards in the future, the type of shareholder initiatives that will be introduced and the corporate response to them, the convergence of corporate laws across countries, the source of resistance to acquisitions and the legal regulation of target defenses, the desirability of legal reforms expanding shareholder voting rights, and the relationship between CEOs and private equity firms.

---

* We would like to thank Bill Allen, John Armour, Oren Bar-Gill, Margaret Blair, Jennifer Hill, Michael Halberstam, Reinier Kraakman, Roy Katzovicz, Michael Levine, Jonathan Lipson, Eric Posner, Randall Thomas, David Webber, and Scott Wornow; participants in the Harvard Law School–Sloan Foundation Conference on Corporate Governance, the University of Pennsylvania Institute for Law and Economics Roundtable, the Seattle University Law School Conference on Adolph Berle, and in workshops at Columbia Law School, Fordham Law School, New York University School of Law, the University of Toronto Law School, and Vanderbilt Law School for comments; and Daniela Fischel, Ryan Golden, and Phillip Won for research assistance. Marcel Kahan would also like to thank the Milton and Miriam Handler Foundation for financial support. Data left uncited throughout this Article are on file with the authors.

I. Introduction

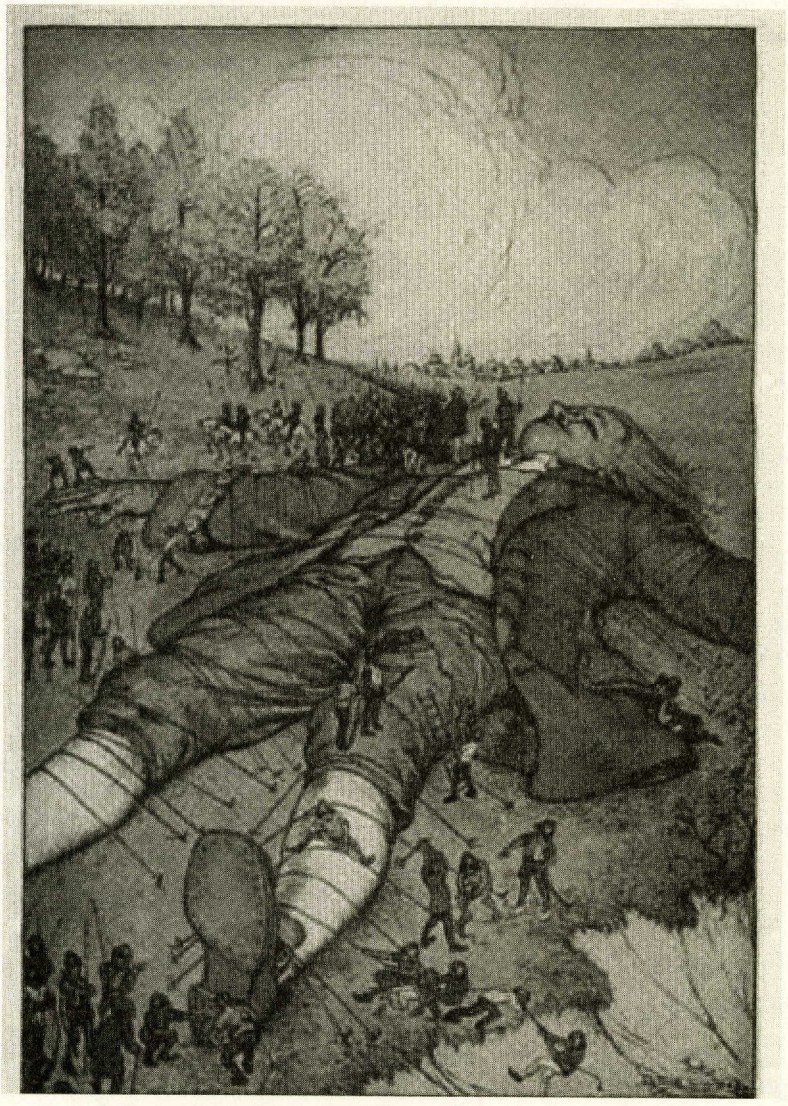
In this Article, we present a straightforward thesis. The CEOs of publicly held corporations in the United States are losing power. They are losing power to boards of directors that increasingly consist of both nominally and substantively independent directors. And, perhaps more so, they are losing power to shareholders. This loss of power is recent (say, since 2000) and gradual, but nevertheless represents a significant move away from the imperial CEO who was surrounded by a hand-picked board and lethargic shareholders.

Most significantly, we think that the recent loss of power is not some cyclical change, and the developments we discuss largely predate the "Great Recession" that started in 2008. Rather, we argue that the decline of CEO power is caused by some underlying changes in the economic and regulatory landscape that will persist, and predict that these changes will result in a further decline of CEO power, at least in the intermediate term. What we may be witnessing is the emergence of a new era of corporate governance for the early part of the twenty-first century, where power over the U.S. corporate enterprise is more evenly distributed between various participants—inside managers, outside directors, and shareholders—rather than concentrated in the hands of the CEO.

This is, to our knowledge, the first academic article to document and analyze the loss of CEO power.[1] Beyond identifying this power decline as an important element of an evolving corporate-governance structure, this Article makes a variety of contributions. First, relying on the philosophical and sociological analyses of power, we use a more sophisticated and self-conscious conceptual scheme that clarifies the various aspects of power that are relevant. Second, because CEO power is both conceptually complex and difficult to observe from outside the firm, we assemble data from a wide variety of sources that provide both direct and indirect evidence of the decline. Finally, we examine the implications of a decline of CEO power in publicly traded U.S. corporations for corporate governance and control.

1. There have been several articles in the trade press on the "decline of the imperial CEO." *See, e.g.*, David Leonhardt with Andrew Ross Sorkin, *Reining in the Imperial C.E.O.: Handshakes Are Becoming a Bit Less Golden*, N.Y. TIMES, Sept. 15, 2002, § 3, at 1 (chronicling attempts by certain boards to restrain CEOs' salaries). The focus of these articles, however, has largely been on CEO firings and compensation. *Id.* While, as we discuss below, both of these are elements of the decline of CEO power, they are only a piece of a much larger picture.

The intuition that drives this Article is nicely captured in Milo Winter's famous 1912 illustration of Gulliver tied down by the Lilliputians.[2]



"I found my arms and legs were strongly fastened on each side to the ground."

Each Lilliputian is far smaller and weaker than Gulliver, and their ropes, individually, are mere threads. Collectively, and in the aggregate, however, the threads bind Gulliver.

---

2. JONATHAN SWIFT, GULLIVER'S TRAVELS 4 (Milo Winter illustrator, Rand McNally & Co. 1912) (1726), *available at* http://www.jaffebros.com/lee/gulliver/winter/p1.jpeg.

The Article is organized as follows. In Part II, we discuss what we mean by "CEO power." In Parts III–VII, we analyze the causes for and symptoms of the loss of CEO power. We consider shareholder composition and activism (Part III), governance rules and the board response to shareholder activism (Part IV), regulatory changes related to shareholder voting (Part V), changes in the board of directors (Part VI), and executive compensation (Part VII). In Part VIII, we tie our analysis of the causes for and symptoms of the loss of CEO power to the aspects and dimensions of power examined in Part II. In Part IX, we discuss the implications of our analysis.

Before embarking on our analysis, it is worth recalling just how much power CEOs had before all the changes that we document here. Myles Mace's surveys of directors and officers during the 1960s provide a wonderful window into what it used to be like.[3] Consider the following quotes from top executives on the role of the CEO and directors:

> To put it bluntly, whether a board has any function or not, it must truly reflect the nature of the chief executive officer of the company more than anything else. If he wants to use the board, he will use them. And if he doesn't want to use the board, he will run over them pretty roughshod. Basically, the board can be made just about as useful as the president wishes it to be.
>
> The president of a company, or the chairman of the board, or whoever runs this operation, really determines the contribution the board makes. If all he wants to do is to get up in front of them and sort of go through some motions, see that fees get distributed, give them a bit of lunch—then that's the kind of performance you will get, because the chief executive officer controls the affair. If, on the other hand, the chief executive officer seeks out where in the management areas various board members might be able to make more of a contribution than in others, and then structures his board so that emphasis is placed on such questions rather than on the rote alternative, then the chief executive is making a direct impact on the contribution the board makes. This, I suppose, is a matter of style.
>
> . . . The old man [the president] has exactly the kind of a board he wants. They all live here in the city, and they just don't do a damn thing as directors. The old man thinks it is a great board, and from his point of view he is probably right. From my point of view they are a big glob of nothing. Not that there aren't some extremely able outsiders on the board—there are. But as board members, they know who is in control and they will never cross the old man.
>
> . . . .

3. MYLES L. MACE, DIRECTORS: MYTH AND REALITY (1971).

> . . . What any new board member finds out very quickly in our company is that it is very difficult to do anything except go along with the recommendations of the president. Because directors who don't go along with them tend to find themselves asked to leave.[4]

These interviews led Mace to conclude the following:

> Presidents of these [widely held] companies have assumed and do exercise the de facto powers of control of the companies for which they are responsible. To them the stockholders constitute what is in effect an anonymous mass of paper faces. Thus, presidents in these situations determine what directors do or do not do.
>
> . . . .
>
> Most presidents, it was found, choose to exercise their powers of control in a moderate and acceptable manner with regard to their relationships with boards of directors. They communicate, though, explicitly or implicitly, that they, as presidents, control the enterprises they head, and this is generally understood and accepted by the directors. Many of them are presidents of companies themselves, and they thoroughly understand the existence and location of powers of control.[5]

II. Power and the CEO

Power is a complex concept that has generated, and continues to generate, a huge literature. Drawing on that literature, in speaking of CEO power, we are interested in three related aspects of power: decision making, second-guessing, and scope.[6]

Decision making refers to the ability of the CEO to decide key issues facing the firm either on her own or by getting the pro forma approval by other decision makers in the firm, such as shareholders, the board of directors, or lower-level managers.[7] A greater ability to decide means more

---

4. *Id.* at 78–79. Mace also comments: "This point of view was confirmed many times during our study." *Id.* at 79.

5. *Id.* at 84.

6. There is a huge literature on "power" in philosophy, sociology, and political science. Important contributions from which we base some of our assertions include ROBERT DAHL, WHO GOVERNS? (2d ed. 2005); STEVEN LUKES, POWER: A RADICAL VIEW (2d ed. 2005); PETER MORRISS, POWER: A PHILOSOPHICAL ANALYSIS (2d ed. 2002); and DENNIS WRONG, POWER: ITS FORMS, BASES, AND USES (Transaction ed., Transaction Publishers 1995) (1979). Because our interest is essentially practical and comparative—whether CEOs are less powerful than they used to be?—we do not need to come up with a comprehensive analysis of power. Rather, it is enough to identify a set of features that approximate the recognized scope of the term power in corporate law and governance. We assert without proving that these aspects capture the essential elements of "CEO power," including the various aspects of "power to" and "power over" as well as the related notions of "autonomy."

7. *Cf.* LUKES, *supra* note 6, at 16–19 (describing a "one-dimensional" view of political power that is best understood as the ability to make decisions affecting others).

CEO power.[8] Decision-making ability, in turn, has several facets. First, CEO power includes the CEO's ability to control whether an issue is even presented to other potential decision makers, i.e., the power of the CEO to control the agenda of these other decision-making bodies.[9] Second, it includes the ability of the CEO to determine the outcome of an issue that is presented to these other bodies, i.e., the power to determine the decision outcome.[10] Third, it includes the ability of the CEO to act if the issue is not presented to these other bodies, i.e., the power to act independently.[11] CEOs have the greatest power if they either (i) have both agenda control and the power to act independently or (ii) have the power to determine the decision outcome.

Second-guessing refers to the ability of other actors to second-guess, and penalize, the CEO for a decision.[12] A lesser ability to second-guess by other actors means more CEO power. Second-guessing, unlike decision making, thus relates not to the ability to make a decision to start with, but to the consequences if other actors, at the time or with the benefit of hindsight, disagree with a decision that has already been made.

Finally, scope relates to the type of decisions that a CEO has the power to make.[13] Scope, in turn, has three dimensions: extension, comprehensiveness, and intensity.[14] Extension relates to the scale of the firm: Given the CEO's power within a firm, a CEO of a larger firm is more powerful than a CEO of a smaller firm.[15] Comprehensiveness relates to the type of decisions over which a CEO has power.[16] For example, can the CEO

---

8. *See* NELSON W. POLSBY, COMMUNITY POWER AND POLITICAL THEORY 3–4 (1980) (discussing the concept of power and suggesting that the amount of power a decision maker possesses is directly related to his ability to make important decisions that affect others and change future events).

9. *Cf. id.* at 25 (noting a more developed theory of political power incorporating agenda setting as an aspect of the ability to make decisions).

10. *See* DAHL, *supra* note 6, at 66 (arguing that a decision maker's power can be measured by "the frequency with which he successfully initiates an important policy over the opposition of others, or vetoes policies initiated by others").

11. *See* WRONG, *supra* note 6, at 36–41 (exploring the concept of "authority" and explaining how powerful leaders are able to impose their decisions and judgments on others without being questioned or tested).

12. In his groundbreaking work analyzing the political power structures in New Haven, Connecticut, Robert Dahl used the frequency of second-guessing veto power used by other parties as a rough measure of their influence over the decision maker. DAHL, *supra* note 6, at 66; *see also id.* at 163–65 (arguing that decision makers who must ultimately answer to other actors—such as politicians who must win future elections to remain in office—are greatly influenced by the preferences of those other actors when deciding what policies to adopt or reject).

13. *See* LUKES, *supra* note 6, at 22 (highlighting the importance scope plays in understanding the amount of power invested in the decision maker).

14. WRONG, *supra* note 6, at 14–16 (citing Bertrand de Jouvenel, *Authority: The Efficient Imperative*, *in* THE NATURE OF POLITICS 84, 85 (Dennis Hale & Marc Landy eds., 1992)).

15. *Id.* at 14–15.

16. *Id.* at 14–16.

set the price of a product sold by the firm, the price of a division that the company wants to put up for sale, or the price at which the whole firm is to be sold to a third party? The more comprehensive the type of decisions, the more powerful is the CEO. Finally, intensity relates to how far the CEO can push others without loss of compliance.[17] A CEO may, for example, have the power to set the price for a product, a division, or the whole firm as long as the price is within the range of reasonable prices but may not be able to get away with setting a ridiculously low price.

Each of these aspects of power captures a different facet of what we take people to mean when they talk of "CEO power." As we will show in the discussion below, the CEO's power has changed over some of these dimensions more than over others. In particular, we will argue that CEOs have lost decision-making power, in terms of agenda control, outcome manipulation, and the power to make decisions independently, albeit in different degrees and to different competing decision-making bodies. CEOs have also become more subject to second-guessing by both shareholders and board members. In terms of scope, CEOs have suffered a decline of power along the dimensions of comprehensiveness and intensity, though not extension.

In addition to its conceptual complexity, power is also a complex social phenomenon that emerges at the intersection of law, norms, and personal qualities such as charisma.[18] The accumulation and exercise of power often occur below the surface, giving rise to an "iceberg" problem: what we observe is likely to be only a small part of what is taking place.[19]

Moreover, in examining changes in power, sorting out cause and effect is likewise extremely difficult both conceptually and analytically. For example, is increased shareholder power evidence of a decline in CEO power, a cause of it, or an effect of it? How about the emergence of more independent boards? The answer, in these and other situations, is often, but not always, "all of the above."

These complications make our analysis somewhat conjectural. Maybe we are completely wrong that CEO power has declined. Maybe CEOs are every bit as powerful as they once were. Maybe they were never very powerful. Or maybe it is impossible to come up with a metric for measuring CEO power or, even if one can design a metric, to collect the data to determine whether CEO power has declined. On the other hand, we think that there are lots of reasons to think that one can intelligibly discuss CEO power,

17. *Id.*

18. For a comprehensive survey of the complexity of political power and decision making in one American city, see DAHL, *supra* note 6. Dahl himself calls the subject "among the most complex phenomena we struggle to understand." *Id.* at xi.

19. *See id.* at 89 (observing that indirect influences on decision making may be very great but difficult to see compared to direct influences).

that it has declined, and that this decline has important implications. That is the case we make in this Article.

III. Changes in Shareholder Composition and Activism

The profile and behavior of shareholders has been fundamentally transformed over the last decade. In this Part, we discuss several changes that have led to a reduction in CEO power: the continued increase in shareholdings by institutional investors and the rise of mutual funds as the most significant type of institutional investor; the emergence of hedge funds as significant shareholder activists; the change by mutual funds and public pension funds to a more confrontational mode of activism; and the increased prominence and power of proxy advisory firms.

A. The Never-Ending Rise of Institutional Investors

Commentators have long noted the change in the ownership structure of shares of publicly traded corporations. In an article published in 1990, for example, Bernie Black presented data showing that the percentage of institutional ownership in New York Stock Exchange (NYSE) companies, which tend to be the largest publicly held companies, had increased from 45.2% in 1980 to 54.4% in 1988.[20] Around the same period, several other articles noted the increase in institutional ownership and the commensurate decline in individual ownership of shares and argued that the concentrated ownership by institutions would be the dawn of a new era of shareholder power.[21] As we have remarked elsewhere, the hopes of these commentators have been largely unfulfilled—until recently, that is.[22]

Probably as a result of the spate of scholarship from the early 1990s, the fact that share ownership by institutions has increased has long been treated as yesterday's news. There are, however, two noteworthy developments about share ownership by institutions since 1990.

---

20. Bernard S. Black, *Shareholder Passivity Reexamined*, 89 MICH. L. REV. 520, 570 (1990). In examining the different measures of concentration of shareholding discussed in the text, it is worth remembering that different studies examine different samples and may define terms differently. The most important results are the trends within a sample.

21. *See, e.g.*, Bernard S. Black, *Agents Watching Agents: The Promise of Institutional Investor Voice*, 39 UCLA L. REV. 811, 827, 830–49 (1991) (arguing that the increase in institutional equity ownership would improve corporate performance). *But see, e.g.*, Edward B. Rock, *The Logic and (Uncertain) Significance of Institutional Shareholder Activism*, 79 GEO. L.J. 445, 452 (1991) ("[R]ecent developments, including the increased concentration of shareholdings, the emergence of new players, and the increased activism of institutional shareholders, are unlikely to bring about a fundamental change in corporate law . . . .").

22. *See* Marcel Kahan & Edward B. Rock, *Hedge Funds in Corporate Governance and Corporate Control*, 155 U. PA. L. REV. 1021, 1024–26 (2007) (exploring activism by hedge funds, which exceeds that by traditional institutional investors, including recent examples and potential problems).

The first is that the increase in share ownership by institutions and the decline in share ownership by individuals have continued since 1990 at virtually the same rate. As shown in Table 1, according to flow of funds accounts compiled by the Federal Reserve, the percentage of stock of publicly held companies held by all institutions (pension funds, mutual funds, banks, insurance companies, and brokers/dealers) increased from 19% in 1970 to 37% in 1990, an 18 percentage-point increase. Between 1990 and 2008, that percentage increased by a further 13 percentage points to 50%.[23] Ownership by households and nonprofits[24]—a category that includes large individual blockholders as well as retail investors—has decreased from 78% in 1970 to 56% in 1990—a 22 percentage-point decrease—and then decreased by another 20 percentage points to 36% in 2008.[25] If the changes in the ownership structure between 1970 and 1990 were notable and important, the further changes between 1990 and 2008 are presumably also important.

Table 1: Percent Ownership of Equities by Types of Investor

	1965	1970	1975	1980	1985	1990	1995	2000	2002	2004	2006	2008
Households and Nonprofits	84	78	70	68	54	56	52	46	39	40	38	36
Private Pension Funds	6	8	13	16	23	17	15	11	13	12	12	11
Public Pension Funds	0	1	3	3	5	8	8	8	9	9	9	9
Mutual Funds	5	5	5	3	5	7	13	19	19	21	23	22
All Institutions	14	19	26	27	40	37	42	44	50	49	51	50

23. *See infra* tbl.1; *see also* BD. OF GOVERNORS OF THE FED. RESERVE SYS., FLOW OF FUNDS ACCOUNTS OF THE UNITED STATES: ANNUAL FLOWS AND OUTSTANDINGS 1985–1994, at 83 (2009), *available at* http://www.federalreserve.gov/releases/z1/current/annuals/a1985-1994.pdf [hereinafter ANNUAL FLOWS AND OUTSTANDINGS 1985–1994]; BD. OF GOVERNORS OF THE FED. RESERVE SYS., FLOW OF FUNDS ACCOUNTS OF THE UNITED STATES: ANNUAL FLOWS AND OUTSTANDINGS 2005–2008, at 83 (2009), *available at* http://www.federalreserve.gov/releases/z1/current/annuals/a2005-2008.pdf [hereinafter ANNUAL FLOWS AND OUTSTANDINGS 2005–2008] (both providing the level values from which the percentages were calculated).

24. Generally, values for the households and nonprofit organizations sector are residuals; in other words, such values equal known totals for all sectors less known values for other sectors. 1 BD. OF GOVERNORS OF THE FED. RESERVE SYS., GUIDE TO THE FLOW OF FUNDS ACCOUNTS 170 (1993). Besides pension funds, mutual funds, banks, insurance companies, brokers/dealers, and households and nonprofits, the Federal Reserve provides data for ownership by state and local governments and foreign residents, among others. *See, e.g.*, BD. OF GOVERNORS OF THE FED. RESERVE SYS., FLOW OF FUNDS ACCOUNTS OF THE UNITED STATES: FLOWS AND OUTSTANDINGS THIRD QUARTER 2009, at 7 (2009), *available at* http://www.federalreserve.gov/releases/z1/current/zi.pdf [hereinafter FLOWS AND OUTSTANDINGS THIRD QUARTER 2009] (listing debt growth by sector: households, business, state and local governments, federal government, domestic financial sectors, and foreign); *id.* at 18–38 (providing data by sector: households and nonprofit organizations, life insurance companies, and private pension funds).

25. *See supra* tbl.1; *see also* BD. OF GOVERNORS OF THE FED. RESERVE SYS., FLOW OF FUNDS ACCOUNTS OF THE UNITED STATES: ANNUAL FLOWS AND OUTSTANDINGS 1965–1974, at 83 (2009), *available at* http://www.federalreserve.gov/releases/z1/Current/annuals/a1965-1974.pdf [hereinafter ANNUAL FLOWS AND OUTSTANDINGS 1965–1974]; ANNUAL FLOWS AND OUTSTANDINGS 1985–1994, *supra* note 23, at 55; ANNUAL FLOWS AND OUTSTANDINGS 2005–2008, *supra* note 23, at 55 (all providing the level values from which the percentages were calculated).

Moreover, the figures in the flow of funds accounts do not account separately for stock held by hedge funds. Rather, hedge fund stock beneficially owned by individuals and nonprofits is included in the household and nonprofit category[26] even though hedge fund assets have surged over the last few years[27] (though they have declined with the recent market downturn[28]). By the end of 2008, hedge funds had an estimated $1.5 trillion in assets under management.[29] If these assets include, say, $500 billion in stock of publicly held corporations, the percentage of such stock held by households would fall to 33% and the percentage held by institutions would rise to 53%. Given these trends, the 90%–10% ratio of retail to institutional stock ownership of the 1950s could soon become 90%–10% the other way.[30] Brian Cartwright, while the general counsel at the SEC, recently referred to this as the "deretailization" of the stock market.[31]

The second development relates to the composition of the institutional holdings. Pre-1990, the most important type of institutional owner was the private pension fund. Ownership by private pension funds had increased from 1% in 1950 to 17% in 1990.[32] By contrast, ownership by mutual funds and public pension funds had increased at a slower pace, from 3% in 1950 to 7% in 1990 for mutual funds and from less than 1% in 1950 to 8% in 1990 for public pension funds. But even in 1990, commentators recognized that

---

26. *See* Rochelle L. Antoniewicz, A Comparison of the Household Sector from the Flow of Funds Accounts and the Survey of Consumer Finances 10 (Oct. 2000) (unpublished manuscript, on file at http://www.federalreserve.gov/PUBS/oss/oss2/papers/antoniewicz_paper.pdf) ("Because hedge funds are not required to file any documentation on their assets or asset values, the flow of funds cannot separate these financial intermediaries from the household sector. Therefore, hedge fund assets are contained within the FFA household sector assets."); *see also* FLOWS AND OUTSTANDINGS THIRD QUARTER 2009, *supra* note 24, at 104 n.1 (indicating that the households and nonprofit organizations sector includes domestic hedge funds).

27. *See* MARKO MASLAKOVIC, INT'L FIN. SERVS., HEDGE FUNDS: CITY BUSINESS SERIES 1 (2007), http://www.ifsl.org.uk/upload/CBS_Hedge_Funds_2007.pdf (noting the hedge fund industry's recent "remarkable growth").

28. *See* INT'L FIN. SERVS. LONDON, HEDGE FUNDS 2009, at 1, (2009) available at http://www.ifsl.org.uk/upload/CBS_Hedge%20Funds%202009(1).pdf (reporting that hedge fund assets plummeted approximately 30% in 2008 and may fall a further 20% in 2009).

29. *Id.*

30. Brian G. Cartwright, Gen. Counsel, SEC, Speech by SEC Staff: The Future of Securities Regulation (Oct. 24, 2007) (transcript available at http://sec.gov/news/speech/2007/spch102407bgc. htm).

31. *Id.*

32. *See supra* tbl.1; *see also* BD. OF GOVERNORS OF THE FED. RESERVE SYS., FLOW OF FUNDS ACCOUNTS OF THE UNITED STATES: ANNUAL FLOWS AND OUTSTANDINGS 1945–1954, at 83 (2009), *available at* http://www.federalreserve.gov/releases/z1/Current/annuals/a1945-1954.pdf [hereinafter ANNUAL FLOWS AND OUTSTANDINGS 1945–1954]; ANNUAL FLOWS AND OUTSTANDINGS 1985–1994, *supra* note 23, at 68 (both providing the level values from which the percentages were calculated).

private pension funds were among the least likely institutions to take an activist approach.[33]

Since 1990, the picture has changed drastically. Holdings by private pension funds have declined from 17% in 1990 to 11% in 2008;[34] and given the disadvantages of defined benefit plans,[35] they are poised for further declines. Mutual funds, by contrast, have taken off, tripling their percentage holdings from 7% to 22%.[36] And mutual funds, as discussed below, have recently become increasingly engaged in governance activism.[37]

B. Hedge Funds: The New Player

As we have shown in an earlier article, activist hedge funds have emerged as critical new players in both corporate governance and corporate control.[38] Hedge funds have created headaches for CEOs and corporate boards by pushing for changes in management and changes in business strategy, including opposing acquisitions favored by management both as shareholders of the acquirer and as shareholders of the target, and by making unsolicited bids.[39] The list of companies that have been subjected to campaigns by hedge funds and other activist investors includes McDonald's,[40] Time Warner,[41] H.J. Heinz Company,[42] Wendy's,[43] Massey Energy,[44] KT&G,[45] infoUSA,[46] GenCorp,[47] Sovereign Bancorp,[48] Deutsche Börse,[49]

---

33. *See* Black, *supra* note 20, at 596–98 ("[W]e can't expect corporate pension managers to become visibly active in the best of circumstances . . . .").

34. *See* ANNUAL FLOWS AND OUTSTANDINGS 1985–1994, *supra* note 23, at 83; ANNUAL FLOWS AND OUTSTANDINGS 2005–2008, *supra* note 23, at 83 (both providing the level values from which the percentages were calculated).

35. *See* Alvin Lurie, *How a Lawsuit Almost Strangled Pensions*, POINTOFLAW.COM, Nov. 12, 2006, http://www.pointoflaw.com/columns/archives/003183.php ("For many years the design of pension plans has been shifting away from the 'defined benefit' format that was once typical. Employers came to dislike such plans because they can impose devastating new funding liabilities in certain situations, as when interest rates sink while the stock market declines. Employees do not find such plans as suitable as they once did because they no longer expect to follow the model of lifetime one-workplace employment for which the plans were originally designed.").

36. *See* ANNUAL FLOWS AND OUTSTANDINGS 1985–1994, *supra* note 23, at 83; ANNUAL FLOWS AND OUTSTANDINGS 2005–2008, *supra* note 23, at 83 (both providing the level values from which the percentages were calculated).

37. *See infra* notes 85–109 and accompanying text (highlighting increased buyout opposition and influence on corporate structure by mutual funds, sometimes in cooperation with traditionally more activist hedge funds).

38. *See* Kahan & Rock, *supra* note 22, at 1029–42 (discussing methods by which hedge funds have pressured for change in corporate governance, blocked acquisitions, and bought or bid for portfolio companies).

39. *See id.* (highlighting a number of examples of hedge fund activism).

40. *Id.* at 1024.

41. *Id.*

42. *Id.*

43. *Id.* at 1031.

44. *Id.* at 1024.

45. *Id.*

Novartis,[50] Sears Holdings,[51] VNU,[52] Delphi,[53] Calpine,[54] Kerr-McGee,[55] Blockbuster,[56] Hollinger International,[57] Target,[58] Kraft,[59] Home Depot,[60] WCI,[61] Motorola,[62] Biogen,[63] Comcast,[64] H&R Block,[65] Tiffany & Co.,[66] Alcoa,[67] BEA Systems,[68] CSX,[69] Circuit City,[70] Zale Corporation,[71] the *New*

---

46. *Id.*

47. *Id.* at 1025.

48. William W. Bratton, *Hedge Funds and Governance Targets*, 95 GEO. L.J. 1375, 1392 (2007).

49. Kahan & Rock, *supra* note 22, at 1025.

50. *Id.*

51. *Id.*

52. *Id.*

53. *Id.*

54. *Id.*

55. *Id.* at 1031.

56. *Id.*

57. *Id.* at 1033.

58. *See* Jared A. Favole & Mike Barris, *With 9.6% Stake in Target, Ackman Fires Value Salvo*, WALL ST. J., July 17, 2007, at C3 (reporting that a 9.6% stakeholder in Target planned to pressure the retailer to increase value).

59. *See* Brad Dorfman, *Kraft in Deal with Peltz's Trian*, REUTERS, Nov. 7, 2007, http://www.reuters.com/article/idUSN0757299320071108 (describing a deal struck between Kraft and an activist shareholder regarding the composition of the board).

60. *See* Ann Zimmerman & Mary Ellen Lloyd, *Home Depot Draws Proxy Threat from Investor*, WALL ST. J., Dec. 19, 2006, at A10 (indicating a potential conflict stemming from stakeholder Relational Development's disapproval of Home Depot's growth strategy).

61. *See* Jonathan Liss, *Icahn, WCI Management Trade Barbs at Hint of Hostile Bid; Shares Jump*, SEEKING ALPHA, Mar. 14, 2007, http://seekingalpha.com/article/29487-icahn-wci-management-trade-barbs-at-hint-of-hostile-bid-shares-jump (reporting that a 14.6% stakeholder claimed the CEO was "not qualified" to lead the company).

62. *See* Sara Silver, *Motorola Reaches Truce with Icahn*, WALL ST. J., Apr. 8, 2008, at B3 (describing a deal by which 6.4% shareholder Icahn dropped his proxy battle against the Motorola board of directors in exchange for the ability to nominate two board positions).

63. *See* Keith J. Winstein, *Biogen Considers a Sale; Pressure Comes from Icahn*, WALL ST. J., Oct. 13–14, 2007, at A3 (describing 4% shareholder Icahn's $23 billion bid to purchase Biogen).

64. *See* Merissa Marr & Dionne Searcey, *Comcast Holder Seeks CEO's Dismissal: Chieftan Targets Dual-Class Voting and Executive Pay*, WALL ST. J., Jan. 18, 2008, at A8 (describing a 2% stockholder's demands for restructuring after Comcast stock fell 40% in one year).

65. *See* Kevin Kingsbury, *H&R Block Holders Vote to Install Breeden Picks*, WALL ST. J., Sept. 7, 2007, at C3 (observing that a substantial majority of the lagging company's shareholders voted to replace three board members).

66. *See Tiffany Will Hear Ideas of Peltz Funds*, N.Y. TIMES, Feb. 2, 2007, at C4 (signaling Tiffany & Co.'s willingness to cooperate with a majority-shareholder investment firm).

67. *See Alcoa: JANA Partners LLC Calls on Alcoa's Board to Drop Alcan's Bid and Pursue Strategic Alternatives Including a Sale of the Company*, REUTERS, May 8, 2007, http://www.reuters.com/article/idUSIN20070508142830AA20070508 (highlighting a hedge fund's activist attempts to influence corporate decisions).

68. *See* Pui-Wing Tam & Vauhini Vara, *Activists Test Silicon Valley: Icahn Is Latest Investor Trying to Unlock Value in Technology Industry*, WALL ST. J., Oct. 20–21, 2007, at B1 (describing hedge fund managers' increasing willingness to influence undervalued technology firms).

69. *See* Michael J. de la Merced, *Hedge Funds Propose CSX Directors, Starting Proxy Battle*, N.Y. TIMES, Dec. 20, 2007, at C2 (describing hedge funds' proposals of five alternate directors as

York Times,[72] and Sprint Nextel.[73] According to Wachtell Lipton partner Patricia Vlahakis, hedge funds conducted 137 activist campaigns just in the fourth quarter of 2007.[74] In many of these instances, hedge funds have been able to win outright or at least to wrest substantial concessions from the management of the companies they target.[75]

This new activism by hedge funds has become a prime irritant for CEOs. Martin Lipton, the renowned advisor to corporate boards, recently listed "attacks by activist hedge funds" as a key issue for directors.[76] Alan Murray from the *Wall Street Journal* calls hedge funds the "new leader" on the "list of bogeymen haunting the corporate boardroom,"[77] and his colleague Jesse Eisinger notes that these days hedge funds are the "shareholder activists with the most clout."[78]

What is particularly noteworthy is the degree of activist bang generated by relatively little buck. Although hedge funds manage substantial amounts of investor money,[79] the large majority of hedge funds do not pursue activist strategies. Rather, according to a recent estimate by J.P. Morgan, only 5% of

---

yet one more example of a hostile maneuver); Heidi N. Moore, *En Garde, Activist Hedge Funds*, WALL ST. J. DEAL J., Mar. 24, 2008, http://blogs.wsj.com/deals/2008/03/24/can-investment-banks-take-a-stand-against-activist-hedge-funds (describing a claim over swap agreements filed by CSX against two hedge funds).

70. *See* Gary McWilliams, *Activists Circle Circuit City*, WALL ST. J., Mar. 24, 2008, at C2 (forecasting changes in management due to activist hedge funds' increased control of Circuit City).

71. *See* Karey Wutkowski, *Breeden Builds Up Stake in Jeweler Zale*, REUTERS, Jan. 7, 2008, http://www.reuters.com/article/idUSN0740372120080108 (noting that activist-investor Breeden increased his holdings in Zale Corporation to 15.85%).

72. *See* Sarah Ellison, *New York Times Holder Protest Grows*, WALL ST. J., Apr. 25, 2007, at B10 (noting that a growing number of shareholders, led by money manager Morgan Stanley, withheld their votes from directors up for reelection as a sign of their dissatisfaction).

73. *See Sprint Nextel Announces Appointment of Ralph Whitworth to Its Board of Directors*, REUTERS, Feb. 12, 2008, http://www.reuters.com/article/pressRelease/idUS145208+12-Feb-2008+BW20080212 (noting the appointment of a large stakeholder to the Sprint Nextel Board as an investment advisor).

74. *Hedge Fund Activism, Possible Recession Will Play Roles in Upcoming Proxy Season*, CORP. L. DAILY, Feb. 1, 2008 (on file with authors).

75. *See supra* notes 45, 48, 49, 53, 56–57, 59 and accompanying text.

76. Memorandum from Martin Lipton, Wachtell, Lipton, Rosen & Katz, to Clients 1 (Dec. 17, 2007) (on file with Texas Law Review).

77. Alan Murray, *Hedge Funds Are New Sheriffs of Boardroom*, WALL ST. J., Dec. 14, 2005, at A2.

78. *See* Jesse Eisinger, *Memo to Activists: Mind CEO Pay*, WALL ST. J., Jan. 11, 2006, at C1 ("The shareholder activists with the most clout these days are hedge-fund managers . . . .").

79. *See* ALEXANDER INEICHEN & KURT SILBERSTEIN, AIMA'S ROADMAP TO HEDGE FUNDS 10 (2008), http://www.aima.org/en/knowledge_centre/education/aimas-roadmap-to-hedge-funds.cfm (estimating hedge fund assets to be between $1.8 and $4 trillion); *see also supra* notes 26–28 and accompanying text.

hedge fund assets are available for shareholder activism.[80] Given these figures, the list of companies targeted by hedge funds is indeed impressive.[81]

C. *Activism by Traditional Institutions*

Traditional institutional investors—specifically public pension funds and mutual funds—have long engaged in low-pressure (and low-cost) "soft" forms of activism, such as voting in favor of corporate-governance shareholder resolutions (and occasionally introducing them) or lobbying boards behind-the-scenes to improve their corporate governance.[82] As discussed below, the corporate-governance issues "du jour" to which this activism is directed are constantly evolving, and the level of support for shareholder resolutions has increased.[83] Overall, however, there has been no major qualitative change in the nature of the soft activism by traditional institutions.[84] On different fronts, however, qualitative changes seem to be occurring.

1. The Awakening of Mutual Funds.—Mutual funds are increasingly engaging in the hard-core activism that has been the hallmark of hedge funds. For example, mutual funds have shown an increased willingness to oppose acquisition of their portfolio companies by private equity firms or large family owners.[85]

In 2007, the most recent year with significant buyout activity, mutual funds successfully opposed a number of buyout transactions approved by the board of directors. For example, mutual-fund giant Fidelity, the largest

---

80. Kahan & Rock, *supra* note 22, at 1046.

81. Hedge fund activism is continuing even after the recent financial crisis. *See* David Walker, *Activist Hedge Funds Passively Rebound*, WALL ST. J. (EUR.), Sept. 29, 2009, at 19 (chronicling activist hedge funds' gain of 29.23% in 2009 after dropping 30.81% in 2008); *see also* Alistair Barr, *Pershing Square to Fight Landry's Buyout*, MARKET WATCH, Nov. 13, 2009, http://www.marketwatch.com/story/pershing-square-to-fight-landrys-buyout-2009-11-13-174800 (describing a major shareholder's successful opposition of an attempted buyout by Landry's founder).

82. *See* Kahan & Rock, *supra* note 22, at 1042–44 (describing the lower impact forms of activism traditionally favored by pension funds and mutual funds).

83. *See infra* notes 156–61 and accompanying text.

84. *See* PAUL LANGLEY, THE EVERYDAY LIFE OF GLOBAL FINANCE 122 (2008) (observing that even after a corporation becomes the target of an activist resolution, it is still common for the institutional investors to voice their concerns in private meetings with executives).

85. *See* Scott Barancik, *OSI Buyout Down but Not Out*, ST. PETERSBURG TIMES, May 9, 2007, at 1D (commenting on the trend of hedge fund managers who are also becoming more vocal critics of unfavorable deals); Tom Lauricella, *Mutual Funds Get Mad*, WALL ST. J., Oct. 2, 2007, at R1 (profiling recent instances of mutual-fund managers "rabble-rousing" and taking action against companies in which they have a stake); Tom Lauricella, *Oppenheimer Revolt Shows Mutual Funds' New Mood*, WALL ST. J., Apr. 11, 2007, at C1 [hereinafter Lauricella, *Opphenheimer*] (reporting that mutual funds are "borrowing a tactic from hedge funds . . . to publicly battle with companies they own"); John Laide, *Investor Activism Against Mergers on the Rise*, SHARKREPELLENT, Mar. 7, 2007, https://www.sharkrepellent.net/pub/rs_20070308.html (tracking major instances of activist shareholders opposing mergers in 2006).

shareholder of Clear Channel Communications, threatened to vote against an acquisition of the company by Bain Capital Partners and Thomas H. Lee Partners, two private equity firms, for $37.60 per share and thereby forced the buyers to raise the price to $39.20.[86] T. Rowe Price, an 8.1% stockholder of Laureate Education, led the opposition to the proposed acquisition of the company for $60.50; the offer was sweetened to $62.[87] Mutual fund Lord Abbett & Co., the second-largest shareholder (6.95%) of OSI Restaurant Partners, complained about the $40-per-share buyout price as having the "markings of a 'panic sale.'"[88] Shareholders approved the deal only after Bain Capital increased the price to $41.15.[89] Investment manager Pzena Investment Management, the second-largest shareholder of Lear, blocked Carl Icahn's offer for the company for $36[90] and also for the sweetened $37.25.[91] And the public pension fund CalPERS opposed a $44 buyout of Biomet Inc. by a group of private equity firms, forcing a price increase to $46.[92] In all of these cases (except for the Clear Channel acquisition), the offer that the institutional investors considered too low—and that was subsequently raised—had the blessing of the respective company's board of directors and its CEO.[93]

The new hard-core activism by traditional money managers is not confined to the buyout area. The money-management arm of the investment bank Morgan Stanley has urged the *New York Times* to dismantle its dual share structure, which assures the founding Sulzberger family of continued control of the company.[94] A campaign led by Morgan Stanley—and supported by mutual funds T. Rowe Price and Legg Mason[95]—for shareholders to withhold their votes from nominees to the board of directors resulted in a 42% withhold vote, which amounted to a majority of the votes not controlled

86. Zachery Kouwe, *Back in Play: Shareholder Outcry Revives Clear Channel LBO*, N.Y. POST, May 8, 2007, at 47; Sarah McBride & Dennis K. Berman, *Clear Channel Is Asked to Reconsider Offer*, WALL ST. J., May 7, 2007, at B12.

87. Jonathan Vuocolo, *Laureate Accepts Sweetened Buyout Bid*, WALL ST. J., June 5, 2007, at C3.

88. Barancik, *supra* note 85.

89. Richard Gibson, *OSI Holders Clear Bid to Go Private*, WALL ST. J., June 6, 2007, at C3.

90. Andrew Farrell, *Icahn's Lear Bid Not Enough, Pzena Says*, FORBES, July 9, 2007, http://www.forbes.com/2007/07/09/lear-icahn-update-markets-equity-cx_af_0709markets20.html.

91. *Lear Holders Reject Icahn's Buyout Bid*, N.Y. TIMES, July 17, 2007, at C4.

92. Jon Kamp, *New Buyout Offer for Biomet Is Now Sweeter and Tender*, WALL ST. J., June 8, 2007, at C3.

93. McBride & Berman, *supra* note 86 (Clear Channel); Susan Carey & Jonathan Vuocolo, *Biomet Agrees to Be Acquired for $10.9 Billion*, WALL ST. J., Dec. 19, 2006, at C4 (Biomet); Michael J. de la Merced, *Parent of Outback Steakhouse Is Sold in $3.2 Billion Deal*, N.Y. TIMES, Nov. 7, 2006, at C2 (OSI Restaurant Partners); Josee Valcourt, *Lear Execs in Hot Seat After Deal Dies*, DETROIT NEWS, July 17, 2007, at 1A (Lear); Vuocolo, *supra* note 87 (Laureate Education).

94. *See* Sarah Ellison, *Paper Chase: How a Money Manager Battled New York Times*, WALL ST. J., Mar. 21, 2007, at A1 (covering a Morgan Stanley portfolio manager's campaign to reform the dual-class structure of the company, among other perceived problems).

95. Ellison, *supra* note 72.

by the Sulzberger family.[96] Morgan Stanley also hired a governance expert respected in activist circles, which fueled speculation that Morgan Stanley may itself be planning to engage in more activism.[97]

A related noteworthy—and novel—phenomenon is the cooperation between mutual funds and hedge funds in pressuring management.[98] For example, in March 2007, Oppenheimer Funds teamed up with several hedge funds to stage a coup and install new top executives at Take-Two Interactive Software Inc., the struggling maker of the popular videogame series Grand Theft Auto.[99] According to the *Wall Street Journal*, this was "the first time in [Oppenheimer's] 46-year history to take such a step."[100] Other recent instances of cooperation between mutual funds and hedge funds include joint efforts to block the acquisition of the London Stock Exchange by Deutsche Börse,[101] of Chiron by Novartis,[102] of MONY by AXA,[103] of Lear by Carl Icahn,[104] and of IMS Health by VNU;[105] joint bids for Beverly Enterprises;[106]

---

96. Landon Thomas Jr., *Shareholders of Times Co. Hold Out 42% of Board Vote*, N.Y. TIMES, Apr. 25, 2007, at C5.

97. *See* Kaja Whitehouse, *Morgan Stanley Buffs Activist Profile*, WALL ST. J., Dec. 13, 2006, at C15 (publicizing speculation that "the decision to hire [the expert], combined with the efforts at the New York Times, suggests the Morgan Stanley unit may be growing more interested in shareholder activism").

98. *See* Christopher Young & Qin Tuminelli, *A New World Order in M&A and Proxy Fights*, *in* INSTITUTIONAL S'HOLDER SERVS., POSTSEASON REPORT 25, 25 (2006), http://www.riskmetrics.com/system/files/private/2006PostSeasonReportFINAL.pdf (relating that asset managers recently have taken the "first tentative steps toward activism").

99. Lauricella, *Oppenheimer*, *supra* note 85.

100. *Id.*

101. *See* David Reilly, *Deutsche Boerse Drama Ends*, WALL ST. J., May 24, 2005, at C14 (reporting on the joint opposition of Atticus Capital (a hedge fund) and Fidelity Investments (a mutual-fund investor) to the takeover of the London Stock Exchange).

102. *See* David P. Hamilton, *Novartis Raises Chiron Bid, Virtually Sealing Deal*, WALL ST. J., Apr. 4, 2006, at A2 (recording that CAM North America and Legg Mason (both mutual funds) and ValueAct Capital (a hedge fund) were investors of Chiron that opposed the takeover by Novartis).

103. *See* Theo Francis, *MONY Holder, Resisting AXA, Suggests a New Chairman, CEO*, WALL ST. J., Jan. 23, 2004, at C4 (noting that MONY's four largest shareholders, composed of mutual funds and hedge funds, opposed the acquisition by AXA, though they claimed not to be working together).

104. *See* Jeff Bennett & Mike Ramsey, *Lear Accepts Carl Icahn's $2.8 Billion Cash Offer*, BLOOMBERG, Feb. 9, 2007, http://www.bloomsberg.com/apps/news?pid=20601087&refer=home&sid=alWGZxoz9GxU (noting opposition to Icahn's offer by Pzena (a hedge fund) and Brandes (a mutual fund)).

105. *See* Jason Singer, *For VNU, a Shareholder Revolt May Lead to Its Sale or Breakup*, WALL ST. J., Oct. 25, 2005, at A3 (reporting that a group of VNU shareholders including Fidelity Investments (a mutual fund) and Knight Vinke (a hedge fund) banded together to tell board members to abandon the friendly acquisition).

106. *See* Kahan & Rock, *supra* note 22, at 1045 (recounting the occasion in 2005 when Franklin Mutual Advisors (a mutual fund) and various hedge funds joined forces to bid jointly on Beverly Enterprises).

and a joint proxy fight over Time Warner.[107] Often, as in the case of Deutsche Börse, it is activist hedge funds that take the lead and mutual funds that follow.[108] But increasingly, as in the case of Oppenheimer, mutual funds are in the forefront.[109]

2. Pension Fund Support of Activist Hedge Funds.—Although pension funds predominantly do not themselves engage in hard-core activism, they often have affiliations with activist hedge funds that do. There are two models for this division of responsibility. The rare but most transparent approach is illustrated by Hermes, the British Telecom-owned fund manager that "manages the assets of the BT Scheme and the Post Office Staff Superannuation Scheme, two of the largest four pension funds in the U.K."[110] In 1998, Hermes established an independent fund, the Hermes U.K. Focus Fund,[111] which has successfully pursued activist strategies.[112] When an activist hedge fund sits on top of, or beside, an index fund, it can be thought of as providing the activist corporate-governance strike force for the associated index fund.[113]

The more common model involves investments by institutional investors in activist hedge funds.[114] Functionally, this is quite similar to the Hermes model: institutions who invest in independent activist funds that target underperforming companies can make money on their direct investments

---

107. *See* Young & Tuminelli, *supra* note 98, at 27 (recalling a 2006 proxy fight where Carl Icahn led a gang of mutual funds and hedge funds to ultimately pressure Time Warner into making valuable concessions).

108. In the case of the failed takeover of the London Stock Exchange by Deutsche Börse, a London-based hedge fund manager was an early shareholder to rattle sabers—mutual funds like Fidelity Investments only followed the activist hedge funds' war cries weeks later. *See* David Reilly, *Deutsche Boerse Faces Mounting Opposition to Its Bid for LSE*, WALL ST. J., Feb. 28, 2005, at C4 (reporting that Fidelity Investments and Merrill Lynch investment managers jumped onto the activist bandwagon in late February); Jason Singer et al., *Fund Chief Fights Deutsche Boerse on Buyout Plan*, WALL ST. J., Jan. 21, 2005, at C4 (profiling the crusade of Christopher Hohn of Children's Investment Fund Management to kill the ill-fated purchase since at least January of 2005).

109. *See* Don Jeffrey & Michael White, *Take-Two Dissidents Win Control, Install New Chief*, BLOOMBERG, Mar. 29, 2007, http://bloomberg.com/apps/news?pid=20601087&src=mwm&sid=agdRA1p8Epzc (last updated Mar. 29, 2007) (reporting that Oppenheimer led other investors, including hedge funds, to force a change of CEO at Take-Two).

110. Marco Becht et al., *Returns to Shareholder Activism: Evidence from a Clinical Study of the Hermes U.K. Focus Fund* 13 (Eur. Corporate Governance Inst., Fin. Working Paper No. 138/2006, 2008), *available at* http://ssrn.com/abstract=934712.

111. *Id.* at 14.

112. *See id.* at 41 (concluding that the Hermes U.K. Focus Fund's activism has created "substantial shareholder gains").

113. *See* Kahan & Rock, *supra* note 22, at 1043 (observing that institutional investors often engage in "passive activism" by taking heed of and supporting the proposals of co-investors).

114. *See* Marcel Kahan & Edward Rock, *Hedge Fund Activism: The Case for Non-intervention*, ADMIN. & REG. L. NEWS, Winter 2008, at 6 (stating that many public pension funds have invested in hedge funds).

and may also increase the value of their portfolio overall.[115] On the other hand, indirect investment provides the institutions with a great deal more insulation from criticism since it is the pension fund itself that selects the target for activism and devises the activist strategy. It is easier and cleaner for the Harvard or Yale endowments or CalPERS or NYCERS to invest in an activist hedge fund than to take the responsibility of starting and operating one themselves on the Hermes model.[116]

D. The Role of Proxy Advisory Firms

Proxy advisory firms have arisen in parallel with the increased share ownership by institutional investors.[117] Such firms make recommendations to their clients—which include most institutional investors[118]—on how to vote their shares in the election of directors, shareholder resolutions, merger proposals, or any other matter on which shareholders vote, as well as provide services that simplify the casting of votes.[119] Commentators have described ISS (the largest advisory firm and now a division of RiskMetrics)[120] as an entity that exercises "tremendous clout,"[121] wielding "extraordinary" influence,[122] being "belligerent,"[123] and to which "powerful CEOs come on bended knees."[124] Claims about its power range from swaying 19% of the

---

115. *See* Randall S. Thomas, *The Evolving Role of Institutional Investors in Corporate Governance and Corporate Litigation*, 61 VAND. L. REV. 299, 312 (2008) (remarking that institutional investors have frequently supported hedge funds in their activism and noting that hedge funds generate value by being good stock pickers and by intervening in undervalued firms on behalf of shareholders).

116. Jim Manzi, *Is Harvard Just a Tax Free Hedge Fund?*, AM. SCENE, May 15, 2008, http://www.theamericanscene.com/2008/05/12/is-harvard-just-a-tax-free-hedge-fund.

117. *See* Tamara C. Belinfanti, *The Proxy Advisory and Corporate Governance Industry: The Case for Increased Oversight and Control*, 14 STAN. J.L. BUS. & FIN. 384, 393 (2009) (highlighting the sharp increase in the percentage of equity securities held by institutional investors and emphasizing that these investors are likely to receive voting advice from proxy advisors).

118. *See id.* at 385–86 (explaining that proxy advisors provide research and voting advice and citing a 2006 *New York Times* article reporting that the advice of the largest proxy advisor—Institutional Shareholder Services (ISS)—affects the decisions of professional investors controlling half the value of the world's common stock).

119. *See* THEODORE ROOSEVELT MALLOCH & SCOTT T. MASSEY, RENEWING AMERICAN CULTURE: THE PURSUIT OF HAPPINESS 170 (2006) (describing proxy advisory firms and noting that they "give advice to institutional shareholders on how to vote their stock").

120. David S. Hilzenrath, *Investor Advisor ISS Is Sold to RiskMetrics*, WASH. POST, Nov. 2, 2006, at D1.

121. Dennis K. Berman & Joann S. Lublin, *Advisor ISS Puts Itself on Sale, Could Fetch Up to $500 Million*, WALL ST. J., Sept. 6, 2006, at C4.

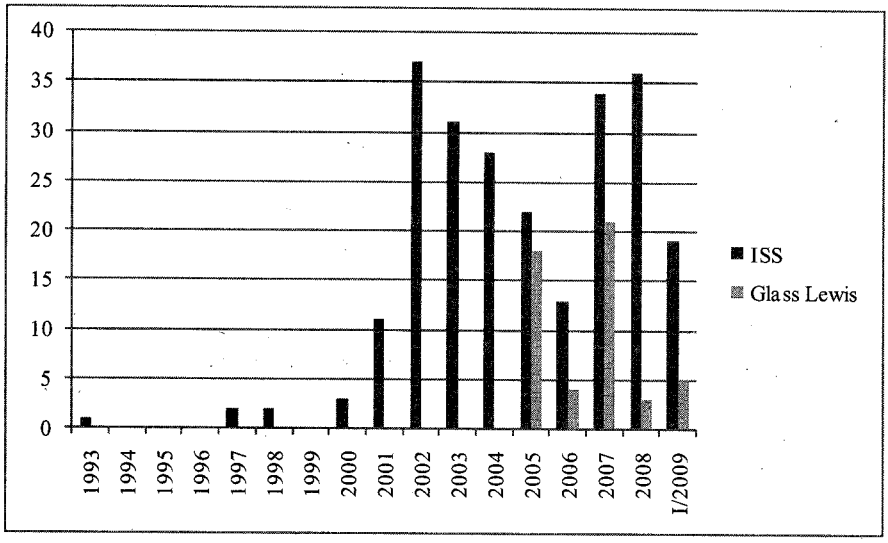
122. *See* Robert D. Hershey, *A Little Industry with a Lot of Sway on Proxy Votes*, N.Y. TIMES, Jan. 18, 2006, § 3, at 6 (describing proxy advisors generally as wielding extraordinary influence and identifying ISS as the most prominent advisory firm).

123. John Goff, *Who's the Boss?*, CFO MAG., Sept. 1, 2004, http://www.cfo.com/article.cfm/3127506?f=singlepage.

124. Leo E. Strine, Jr., *The Delaware Way*, 30 DEL. J. CORP. L. 673, 688 (2005).

votes,[125] to 21%,[126] to 30%,[127] to affecting the vote of $25 trillion in assets,[128] to getting "whatever [it] wants."[129] Martin Lipton blames "influential proxy advisory firms," together with hedge funds and other activist shareholders, for undermining the board-centric model of governance.[130]

Figure 1: Advisor References per Year[131]



As a rough metric of the increased power of proxy advisors, or at least of the public perception of such power, we collected references to ISS (or RiskMetrics) as well as to Glass Lewis (the second-largest proxy advisory firm). As reported in Figure 1, the number of times that these firms are referred to as powerful, influential, and the like has grown substantially since 2000.

While many of the more extreme claims about the power of these firms are likely to be exaggerated,[132] proxy advisors are important in at least two

125. Jie Cai et al., *Electing Directors*, 64 J. FIN. 2389, 2404 (2009).

126. Jennifer E. Bethel & Stuart L. Gillan, *The Impact of the Institutional and Regulatory Environment on Shareholder Voting*, FIN. MGMT., Winter 2002, at 29, 30.

127. *Is ISS Too Powerful? And Whose Interests Does It Serve?*, Posting of William J. Holstein to Corner Office, http://blogs.bnet.com/ceo/?p=1100 (Feb. 7, 2008, 8:03 EST).

128. *See* Hershey, *supra* note 122 (describing ISS's influence on the governance decisions of professional investors who control $25 trillion in assets).

129. *Is ISS Too Powerful? And Whose Interests Does It Serve?*, *supra* note 127.

130. Martin Lipton, *Some Thoughts for Boards of Directors in 2008*, BRIEFLY, Jan. 2008, at 1.

131. The data reflected in this figure were gathered by conducting a search on Westlaw for articles mentioning ISS, RiskMetrics, or Glass Lewis within ten words of shareholder and either powerful, clout, or influential. The data for 2009 include results through June 10.

132. Stephen Choi, Jill Fisch & Marcel Kahan, *The Power of Proxy Advisors: Myth or Reality?*, 59 EMORY L.J. (forthcoming 2010).

ways. First, they may be new and independent power centers that, to some significant degree, influence the votes of clients.[133] Second, they may function as central coordinating and information agents who help create a unified front of institutional investors, and thereby increase collective institutional shareholder influence.[134]

IV. Changes in Governance Rules and Boards' Responses to Activism

In recent years, the governance structure of large publicly held corporations has been transformed through a combination of regulatory changes and shareholder activism.[135] These changes are both a cause and a reflection of a decline in CEO power.[136] In this Part, we document and analyze trends with respect to staggered boards, the voting rules in director elections, and the adoption and implementation of shareholder proposals.

A. The Decline of Staggered Boards

Modern corporate law scholarship regards staggered boards as one of the most potent and controversial anti-takeover devices.[137] In companies with "effective" staggered boards, it takes two consecutive annual shareholder meetings to replace a majority of a board of directors against the opposition of incumbents.[138] While poison pills that are not coupled with staggered boards are nowadays viewed as relatively harmless,[139] several commentators have argued that staggered boards, coupled with the (virtually) universally available poison pill, serve to illegitimately entrench managers

133. *See* James F. Cotter et al., *ISS Recommendations and Mutual Fund Voting on Proxy Proposals*, 50 VILL. L. REV. (forthcoming 2010) (manuscript at 107, on file with Texas Law Review) (finding that mutual-fund votes correspond more closely to ISS recommendations than to management recommendations).

134. *See* Thuy-Nga T. Vo, *Rating Management Behavior and Ethics: A Proposal to Upgrade the Corporate Governance Rating Criteria*, 34 J. CORP. L. 1, 8 (2008) ("The increasing concentration of stock ownership in the hands of institutional investors, and the interest of these institutional investors in the governance of public companies, have also fueled the need for information about corporate governance practices.").

135. *See id.* at 11–13 (describing how public companies are now required by the Sarbanes–Oxley Act to have audit committees composed of independent directors and how the ratings systems used by major advisory services reflect generally accepted views of strong governance).

136. *See id.* at 13 (arguing that companies are increasingly assigning separate individuals to the board-chair and CEO positions despite the absence of a legal duty to do so in order to enhance the board's role as an independent monitor of management's performance).

137. *See, e.g.*, Lucian Arye Bebchuk et al., *The Powerful Antitakeover Force of Staggered Boards: Theory, Evidence, and Policy*, 54 STAN. L. REV. 887, 890 (2002) (commenting that a staggered board "offers a more powerful antitakeover defense than has previously been recognized" and suggesting that the staggered board's power as a takeover prevention tool may warrant changes in takeover regulation and in general takeover doctrine).

138. *Id.* at 912.

139. *Id.* at 899.

and that courts should find some way to render them ineffective.[140] The policy battlefront for takeover defenses, in other words, has shifted to staggered boards.

For existing companies, conventional wisdom had it that shareholders and boards are in a stalemate. Boards of companies without staggered boards may want to adopt staggered boards, but they do not propose a charter amendment because they know that shareholders will not approve it.[141] Shareholders in companies with staggered boards want to get rid of them but cannot because the board refuses to approve the requisite charter amendment.[142]

The conventional wisdom is wrong. The tide on staggered boards has turned and, at least for the largest companies, the day is not far off when staggered boards will be the rare exception. In Table 2, we present data on staggered boards in the S&P 100 companies. S&P 100 companies are among the largest and the most established companies in the U.S., representing, in aggregate, almost 45% of the market capitalization.[143]

Table 2: Staggered Boards in S&P 100 Companies

	Companies with Staggered Boards	New Adoptions	Eliminations	Eliminations as % of Companies with Staggered Boards
2003	44			
2004	41	0	3	6.80%
2005	34	0	7	17.10%
2006	25	0	8	23.50%
2007	21	0	5	24.00%
2008	16	0	1	6.30%
2009	15			

140. *See* Bernard Black & Reinier Kraakman, *Delaware's Takeover Law: The Uncertain Search for Hidden Value*, 96 NW. U. L. REV. 521, 561 (2002) (suggesting that courts should review staggered board terms because "[n]either the finance literature nor the norms of corporate law support vesting such unbalanced power in the hands of the board"); *see also* Bebchuk et al., *supra* note 137, at 949 (indicating that shareholders should not be permitted to adopt an anti-takeover device, such as staggered boards, that does not allow for a one-time up-or-down referendum on acquisition offers).

141. Bebchuk et al., *supra* note 137, at 900.

142. This was largely true until 2003. *See* Jennifer Levitz, *Getting the Message*, WALL ST. J., Oct. 9, 2006, at R6 (showing that virtually all shareholder resolutions that received majority support were ignored prior to 2003).

143. STANDARD & POORS, S&P 100, at 1 (2008), www2.standardandpoors.com/spf/pdf/index/SP_100_Factsheet.pdf.

As Table 2 shows, the incidence of staggered boards has declined from 44% to 16% between 2003 and 2009.[144] Put differently, over the six-year period, two-thirds of the companies that had staggered boards have dismantled them.

To be sure, the decline of staggered boards among the largest and most established companies does not necessarily mean that staggered boards are universally in decline. Arguably the managers of the largest companies are least in need of insulation against takeovers, and thus most willing to agree to destagger. Indeed, staggered boards are alive and well in companies at the time of their IPO. In a sample of twenty-six companies that went public in the first part of 2007, we found that twenty had a staggered-board provision in their charter.

That said, the largest and most established companies act as trendsetters for what is considered good corporate governance.[145] The directors of these companies sit on boards of smaller companies, and their managers are members of influential groups like the Business Roundtable.[146] With most of these companies dismantling their staggered boards over the last six years, it will become increasingly difficult for other companies to resist shareholder pressure.

In fact, smaller companies have started to go down the same path of dismantling their staggered boards that the S&P 100 companies have almost completed. Thus, according to *SharkRepellent*, the incidence of staggered boards among the (still large) S&P 500 companies declined from 57% in 2003 to 36% in 2007; among midsize S&P 400 companies, it declined from 67% in 2003 to 58% in 2007; and among small S&P 600 companies, it declined from 61% in 2003 to 55% in 2007.[147] Thus, we already see that other companies have started, and we predict that they will continue to, follow the lead of the S&P 100 companies.

---

144. Of the sixteen companies that still had staggered boards in place for 2008, one had a "non-effective" staggered board (which is not regarded as a forceful anti-takeover mechanism), and the boards of two others had, since 2003, proposed charter amendments to destagger that failed to get the requisite (supermajority) shareholder approval.

145. *See* Jared A. Favole, *Big Firms Increasingly Declassify Boards*, WALL ST. J., Jan. 10, 2007, at B2 (noting that in 2006 S&P 500 companies were the first group, ahead of small and midcap companies, to surpass the 50% mark for having declassified boards).

146. *See* BusinessRoundtable.org, About Us, http://www.businessroundtable.org/about (detailing the influence CEOs from leading U.S. companies have on the association and noting that member companies "comprise nearly a third of the total value of the U.S. stock markets and pay more than 60 percent of all corporate income taxes paid to the federal government").

147. *See* SharkRepellent.net, S&P 1500 Classified Boards at Year End (2009) (on file with Texas Law Review) (reporting the number of classified boards in S&P 400, 500, and 600 companies from 1998 to 2009).

B. The Meteoric Rise of Majority Voting for Directors

Perhaps the most astonishing change in the corporate-governance environment is the meteoric rise of majority voting for directors. The traditional voting standard for director elections was plurality voting.[148] Under plurality voting, the directors who receive the most votes are elected. This means, in effect, that if the number of nominees is equal to the number of vacancies—as is the case in the overwhelming majority of director elections—every nominee is assured election since it takes only one vote to be elected.[149]

Until recently, the directors of most corporations were elected under a plurality-voting regime. Of S&P 100 companies, only ten deviated from plurality voting in 2003.[150] By 2009, that number had increased to ninety (see Table 3). Moreover, of the ten remaining companies, one had not yet filed its 2009 proxy statement, four were no longer publicly traded, and four others had some form of cumulative or dual-class voting regime in place, which complicates majority voting for directors.[151] Only a single company definitely retained a regular plurality-voting regime. As Table 3 shows, most of the change from plurality to majority voting took place in the two-year span from 2005 to 2007, where the number of S&P 100 companies with majority voting increased from nine to eighty-one. Thus, within just two years, we have moved from a regime in which majority voting was the rare exception to a regime in which it has been adopted by virtually all of the largest companies. Though the rise of majority voting among the broader set of S&P 500 companies has been somewhat slower,[152] experienced observers

---

148. *See* Annalisa Barrett & Beth Young, *Majority of Votes Withheld: Shareholders Say "No," Boards Say "Yes,"* 16 CORP. GOVERNANCE ADVISOR 6, 6 (2008) (discussing the shift from plurality to majority voting as it relates to withheld votes); CLAUDIA H. ALLEN, NEAL, GERBER & EISENBERG, LLP, STUDY OF MAJORITY VOTING IN DIRECTOR ELECTIONS, at ii (2007), http://www.ngelaw.com/files/upload/majoritystudy111207.pdf ("Until recently, virtually all directors of U.S. public companies were elected under a 'plurality' vote standard.").

149. *See* Allen, *supra* note 148, at ii ("A nominee in an election to be decided by a plurality could theoretically be elected with as little as one vote, thereby ensuring that, in an uncontested election, nominees slated by a board will be elected and that board seats will not be left vacant.").

150. Two as a result of state law, five due to charter or bylaw provisions, and three for unknown reasons.

151. Majority voting is not well defined for cumulative voting.

152. *See* ALLEN, *supra* note 148, at i (highlighting the increase in S&P 500 majority voting from 16% to 66% in the period from February 2006 to November 2007). Companies that have adopted majority voting differ in whether they have done so through a bylaw amendment, which usually specifies that a director who receives more "withhold" or "against" votes than "for" votes is not elected, or through corporate-governance guidelines requiring a director to tender her resignation if she receives more "withhold" or "against" votes than "for" votes. *Id.* at ii–iii, ix. Delaware law was recently changed to clarify that a resignation conditional on not receiving a specified vote can provide that it is irrevocable. DEL. CODE ANN. tit. 8, § 141(b) (2009). The distinction between these two variants, however, is not large. Even if the director is not elected, the remaining board members could, if they wanted to, fill the resulting vacancy with the very director who failed to receive the requisite shareholder vote. Directors, of course, will be reluctant to do so,

like Martin Lipton opined that "it is clear today that majority voting will become universal."[153]

Table 3: Number of S&P 100 Companies Applying Majority Voting

	2003	2004	2005	2006	2007	2008	2009
Majority Voting	10	9	9	51	81	85	90

Two further comments are in order to put this shift in perspective. First, it is important to highlight that boards simply caved in to demands for majority voting. Unlike the shift from staggered boards to annual election, the more dramatic shift from plurality to majority voting was not preceded by a long and tortured shareholder campaign; it happened over a very short time span and was more complete.[154] Second, the shift to majority voting makes the shift from staggered boards all the more important. To the extent that majority voting provides a tool for shareholders to show their disapproval for specific directors, rather than the board or management as a whole, annual voting means that shareholders have the opportunity to do so, *for each director*, on a yearly basis. Thus, while staggered boards have hitherto been viewed largely as an anti-takeover device, they now are also important as a mechanism to insulate board members from shareholder "withhold" campaigns. And the demise of staggered boards documented in the previous sections means that the ability to exert pressure via withhold campaigns is increasing.

C. More—and More Successful—Shareholder Proposals

Another piece of evidence suggesting that the landscape is changing relates to precatory shareholder resolutions. In precatory resolutions, shareholders request the board of directors take a certain action—such as redeem a pill or propose a charter amendment—without mandating the action. Virtually all of these resolutions are introduced under Rule 14a-8 of the Securities Exchange Act, which permits shareholders, at little cost, to force the company to include a resolution in its own materials.[155] Precatory resolutions thus represent a low-cost and (since they are not binding) relatively low-pressure form of activism.

---

but they will be equally reluctant to reject the resignation of a director who received more votes "against" than "for." In any case, most recent moves to majority voting are via bylaw amendments, and many companies that had initially adopted corporate-governance guidelines have subsequently adopted a bylaw. ALLEN, *supra* note 148, at ii, ix, fig.1.

153. Lipton, *supra* note 130, at 4.

154. *See* Vincent Falcone, Note, *Majority Voting in Director Elections: A Simple, Direct, and Swift Solution?*, 2007 COLUM. BUS. L. REV. 844, 853–55 (2007) (describing the widespread success of the majority-voting movement).

155. 17 C.F.R. § 240.14a-8 (2008).

It was long thought that precatory shareholder resolutions did not have much of an effect.[156] This used to be true—but no longer is. First, an increasing number of shareholder resolutions are adopted by shareholders.[157] Since 2001, Georgeson, a major proxy solicitor, has prepared an annual *Corporate Governance Review* showing the voting results on corporate-governance-related shareholder resolutions filed with S&P 1500 companies.[158] These proposals concern issues like majority voting, declassifying the board, executive compensation, or the right to call a shareholder meeting.[159] Table 4 shows, for each year, the number of proposals receiving majority shareholder support and whether the board in the year after passage implemented the proposal (i.e., did what the shareholders asked it to do), ignored the proposal (i.e., refused to do what the shareholders asked it to do), or did neither (e.g., because the company was acquired or because the proposal asked the board to refrain from taking an action which the board ordinarily would not have taken anyway within that time frame).

Table 4: Shareholder Proposals (S&P 1500)[160]

	2001	2002	2003	2004	2005	2006	2007	2008
Adopted Proposals	25	51	85	79	55	73	80	86
Implemented Proposals	3	9	21	28	34	46	39	43
Ignored Proposals	20	31	44	24	9	10	28	33
Implementation Percentage	12	18	25	35	62	63	49	50
Ignored Percentage	80	61	52	30	16	14	35	38

Since 2001, the number of implemented proposals has been rising steadily, from three in 2001 to forty-three proposals in 2008. This is due partly to an increase in the number of proposals receiving majority support but, to an even greater degree to an increase in the percentage of implemented proposals, from 12% in 2001 to 50% in 2008. Correspondingly, the percentage of ignored proposals has declined. Thus, in 2001, only twenty-five proposals were adopted and 80% of those were simply ignored by the

156. Roberta Romano, *Less Is More: Making Institutional Investor Activism a Valuable Mechanism of Corporate Governance*, 18 YALE J. ON REG. 174, 177 (2001).

157. We considered a resolution adopted if it received more votes in favor than the combined votes against and abstentions (including broker no-votes). This appears to be the standard used by most companies. *See* DEL. CODE ANN. tit. 8, § 216(2) (2001 & Supp. 2008) ("In all matters other than the election of directors, the affirmative vote of the majority of shares present in person or represented by proxy at the meeting and entitled to vote on the subject matter shall be the act of the stockholders.").

158. Georgeson, Research, http://www.georgesonshareholder.com/emea/resources_research.php. Prior to 2001, Georgeson had also prepared such a report, but analyzed only corporate-governance proposals made by institutional investors. *Id.*

159. *Id.*

160. The data reflected in this table was gathered from Georgeson's Annual Corporate Reviews. *Id.* For each adopted resolution, we conducted research to see whether it was implemented or ignored over the next year.

board of directors. By 2008, eighty-six proposals were adopted and half of them were implemented.[161]

V. Regulatory Changes—Voting and Solicitation of Proxies

In this Part, we discuss several regulatory initiatives, some that have passed and others that are pending. As we will argue, these changes have made, and have the potential to make, shareholder activism less costly and more effective. We discuss four regulatory initiatives: the 1992 amendments to the proxy rules, the reduction in the scope of discretionary broker voting, the recently enacted notice and access rules, and the proposals for proxy access.

A. The 1992 Amendments: An Honorable Mention

Though they predate the time period that we focus on in this Article, the 1992 amendments[162] to the proxy rules merit an honorable mention. In many respects, the 1992 amendments enabled the increased levels of activism[163] that started only about a decade after the amendments were adopted.[164] To understand the import of these amendments and some later reforms, it is important to give a short overview of the federal regulation of proxy solicitations.

The federal proxy rules prohibit *any person* from engaging in a solicitation unless either that person has filed a proxy statement with the SEC and sent it to each shareholder who is being solicited or an exception applies.[165] The definition of "solicitation" in the proxy rules, moreover, is extraordinarily broad and includes virtually any comment about the company, management, or any proposal to be voted on.[166] Preparing a proxy statement is a somewhat tedious and costly task, and printing and mailing it to each solicited shareholder further adds to the expense. Thus, unless one was willing to spend a fairly substantial amount of money, one would only want to engage in a solicitation if an exception to the proxy statement filing requirement applied.

Prior to the 1992 amendments, these exceptions were highly limited.[167] The natural effect of these rules was thus to stifle communication and coor-

---

161. For another study on shareholder resolutions arriving at similar conclusions, see Randall S. Thomas & James F. Cotter, *Shareholder Proposals in the New Millennium: Shareholder Support, Board Response, and Market Reaction*, 13 J. CORP. FIN. 368, 369 (2007).

162. 17 C.F.R. §§ 240, 249 (1992).

163. Catherine M. Daily et al., *Corporate Governance: Decades of Dialogue and Data*, 28 ACAD. MGMT. REV. 371, 376 (2003).

164. *See supra* Part III.

165. 17 C.F.R. § 240.14a-3 (2009).

166. *See id.* § 240.14a-1(*l*) (defining solicitation to include any communication reasonably calculated to result in the procurement or withholding of a proxy).

167. Daily et al., *supra* note 163, at 376.

dination among shareholders. The solicitations that did occur were part of full-fledged proxy contests, conducted by large shareholders who sought to obtain control over the company.[168]

The 1992 amendments added an additional important exception to the requirement to prepare and file a proxy statement. Under Rule 14a-2(b)(1), most persons who do not either seek the power to act as proxy or furnish a form of proxy need not file a proxy statement.[169] When the solicitation is oral, no filings of any sort are required.[170]

This exception is useful as long as the solicited shareholders can vote on the form of proxy distributed *by the company*. This is possible in solicitations in favor of a shareholder resolution introduced under Rule 14a-8, solicitations in opposition to a management proposal, and solicitations to withhold authority to vote for certain directors.[171]

Rule 14a-2(b)(1) was mostly meant to encourage involvement by institutional investors.[172] However, until recently, it does not appear that institutional investors—or anyone else, for that matter—made much use of the exceptions in the rules. Active campaigns in favor of a shareholder proposal or in opposition to a management proposal, or large-scale efforts to withhold votes in director elections were rare until 2004.[173]

More recently, however, the exceptions in Rule 14a-2(b)(1) have become much more important. As discussed above, institutions now commonly oppose proposed mergers endorsed by the board of directors.[174] They also increasingly engage in campaigns in support of shareholder proposals.[175] Hedge funds can also avail themselves of this exemption to coordinate their activities as long as they stop short of forming a 13(d) "group."[176] Finally, the 2004 campaign to withhold votes from Disney CEO Michael Eisner—a campaign that contributed to Eisner's resignation a year

168. *See* Lucian Ayre Bebchuk & Marcel Kahan, *A Framework for Analyzing Legal Policy Towards Proxy Contests*, 78 CAL. L. REV. 1071, 1075 (1990) (explaining that by 1990 proxy contests were becoming the takeover method of choice for large shareholders).

169. 17 C.F.R. § 240.14a-2(b)(1).

170. *Id.* § 240.14a-6(g)(2).

171. *Id.* § 240.14a-4.

172. *See* John C. Coffee, Jr., *The SEC and the Institutional Investor: A Half-Time Report*, 15 CARDOZO L. REV. 837, 840 n.17 (1994) (describing how 14a-2(b)(1)'s safe harbor was designed to encourage broad communication among shareholders, particularly institutional investors).

173. *See* Lisa M. Fairfax, *The Future of Shareholder Democracy*, 84 IND. L.J. 1259, 1290 (2009) (noting that since 2004 there has been a dramatic increase in majority vote shareholder proposals to be included on corporations' proxy statements). Prior to 2004, Georgeson did not keep track of "other activist events" where dissidents did not distribute a separate proxy card, indicating that such solicitations were rare.

174. *See supra* subpart III(C).

175. *See supra* notes 155–72 and accompanying text.

176. *See* CSX Corp. v. Children's Inv. Fund Mgmt., 562 F. Supp. 2d 511, 552 (S.D.N.Y. 2008) (discussing when hedge funds are deemed to have formed a group for § 13(d) purposes).

later—showed shareholders the power of just voting no.[177] Since 2004, large-scale moves to withhold votes for directors have become increasingly common.[178] Without the 1992 amendment, such campaigns would be much more difficult.

Some of the other governance changes discussed above will create even more opportunities for making use of Rule 14a-2(b)(1). Specifically, the moves to annual elections of the entire board and majority voting[179] mean that the opportunities for, and the incentives to engage in, withhold campaigns increase. With annual elections, the opportunity to withhold a vote for a specific director arises every year, rather than once every three years under staggered boards.[180] And with majority voting, the result of a sufficient withhold vote is that the director is not elected or is required to offer to resign, rather than mere embarrassment under plurality voting.[181] Moreover, the ability of activists to threaten to engage in a withhold campaign is greatly enhanced by the fact that such a campaign would fall under the Rule 14a-2(b)(1) exceptions. This is especially true for activists that are viewed as cost sensitive, such as mutual funds or public pension funds. Thus, whatever the use of the Rule 14a-2(b)(1) is today, we think it is likely that the use will increase significantly in the years ahead.

B. The End of Brokers' Discretionary Voting in Director Elections

Under long-standing practice, brokers may vote shares held in their accounts according to their discretion when they do not receive specific instructions from the beneficial owners of these shares and when the matter is designated as "routine" by the NYSE.[182] These discretionary broker votes have been a reliable and significant source of pro-management votes.[183] Many individual shareholders, who tend to hold their shares in brokerage accounts, do not bother to provide voting instructions, and brokers have

---

177. Kara Scannell, *'Broker Votes': Opponents May Win One*, WALL ST. J., June 13, 2007, at C1. The term "just vote no" goes back to Joe Grundfest, who had proposed such a strategy in a 1993 article. Joseph A. Grundfest, *Just Vote No: A Minimalist Strategy for Dealing with Barbarians Inside the Gates*, 45 STAN. L. REV. 857, 927 (1993).

178. *See* Fairfax, *supra* note 173, at 1289 (stating that the 2004 director elections for Disney and Federated Department Stores, Inc. represent the two elections with the most withheld votes "in recent history").

179. *See supra* notes 137–63 and accompanying text.

180. *See* Bebchuk et al., *supra* note 137, at 893 ("In a company with a staggered board, directors are grouped into classes (typically three), with each class elected at successive annual meetings. . . . With three classes, directors in each class would be elected to three-year terms.").

181. *See* Fairfax, *supra* note 173, at 1288–89 (explaining the effectiveness of a withhold-the-vote campaign in both majority and plurality systems).

182. NYSE, Inc., Listed Company Manual § 452 (2003).

183. *See* Marcel Kahan & Edward Rock, *The Hanging Chads of Corporate Voting*, 96 GEO. L.J. 1227, 1250 (2008) (explaining how a failure to vote has a unique impact when mediated through a broker).

tended to use their discretion to vote shares in accordance with the board's recommendations.[184]

Historically, the NYSE had regarded uncontested director elections—that is, elections where there is only one slate of nominees—as routine, even when some shareholders waged an active campaign to convince other shareholders to "withhold" their votes from certain nominees. For example, in 2004, 43% of the shares voted to withhold support from Disney's CEO Michael Eisner—a far from routine occurrence.[185] Brokers, however, were permitted to cast the votes of uninstructed shares, and, according to some, if broker votes had been ignored, Eisner would not have received majority support.[186]

The practical import of discretionary broker votes in director elections has recently declined. First, as discussed above, the percentage of shares held by individual investors—the type of investor most likely to hold their shares in brokerage accounts and not to return a ballot[187]—is steadily decreasing.[188] Second, some brokers have moved from voting uninstructed shares in accordance with the board recommendation to voting them in the same proportion as those shares in their accounts for which they received voting instructions.[189] On the other hand, voting is more important than ever. With the rise of withhold-vote campaigns and, as discussed above, majority voting for directors, an increasing number of director elections will likely become truly nonroutine.

Either way, discretionary broker votes in director elections are now gone. In October 2006, the NYSE proposed to amend Rule 452 governing broker votes to redefine all director elections as nonroutine.[190] The proposed change required SEC approval to become effective.[191] After not taking any action for over two years, the SEC last February solicited comments on the

184. *See* Scannell, *supra* note 177 ("Brokers generally vote for management, partly, they say, because if clients wanted them to oppose management they would let them know.").

185. *Id.*

186. *Id.*

187. *See SEC Hears Testimony on Broker Votes*, Posting of Ted Allen to RiskMetrics Group, http://blog.riskmetrics.com/gov/2007/05/sec-hears-testimony-on-broker-votessubmitted-by-ted-allen-director-of-publications.html (May 24, 2007, 10:58 EST) ("While most institutions now vote their shares or give voting instructions, only 30 to 40 percent of retail investors bother to vote their shares.").

188. *See supra* subpart III(A).

189. *See* Scannell, *supra* note 177 (reporting that Goldman Sachs, Merrill Lynch, and Morgan Stanley voted proportionally in the 2007 season and that Charles Schwab has done so since 2005).

190. *See* PROXY WORKING GROUP TO THE N.Y. STOCK EXCH., REPORT AND RECOMMENDATIONS OF THE PROXY WORKING GROUP TO THE NEW YORK STOCK EXCHANGE 3–4 (2006), http://www.nyse.com/pdfs/PWG_REPORT.pdf ("[T]he election of directors can no longer be considered a 'routine' event in the life of a corporation.").

191. 15 U.S.C. § 78s(b)(1) (2006) ("No proposed rule change shall take effect unless approved by the Commission or otherwise permitted in accordance with the provisions of this subsection.").

proposed changes.[192] On July 1, 2009, the SEC finally approved the amendments, effectively ending discretionary broker voting in director elections for all shareholder meetings held on or after January 1, 2010.[193]

An end of discretionary broker voting will obviously make it easier for withhold campaigns to succeed. On nonroutine shareholder proposals, it is estimated that broker nonvotes amount on average to 19% of the votes cast at an annual meeting.[194] With discretionary broker voting, a board could until recently count on brokers voting most of these shares in favor of its nominees. That would imply that, to get a majority-withhold vote, activists would have had to get over 60% of the instructed shares—a pretty steep task.[195] Without broker nonvotes, this task is much easier.

C. *Notice and Access*

In 2007, the SEC enacted new rules governing the electronic delivery of proxy materials.[196] Under the new rules, instead of furnishing the whole set of proxy materials, companies may mail shareholders a short notice providing some basic information about the issues to be voted on at the annual meeting and refer them to a website where the proxy statement and other solicitation materials are available.[197] The "notice and access" option is meant to reduce the cost of printing and mailing solicitation materials.[198] Importantly, the notice and access option is also made available to shareholders engaged in a solicitation opposed by the company.[199]

---

192. Notice of Proposed Rule Change to Amend NYSE Rule 452 and Eliminate Broker Discretionary Voting for the Election of Directors, Exchange Act Release No. 59,464, 95 SEC Docket 744 (Feb. 26, 2009).

193. *See* Melissa Aguilar, *SEC Approves NYSE Broker Vote Ban in Director Elections*, COMPLIANCE WK., July 1, 2009, http://complianceweek.com/blog/aguilar/2009/07/01/sec-approves-nyse-broker-vote-ban-in-director-elections/ (reporting the SEC's approval of the proposed rule, its effective date, and its likely effects).

194. *See SEC Hears Testimony on Broker Votes*, *supra* note 187 (attributing this figure to Broadridge Financial).

195. The exact percentage of instructed shares is a function of both the percentage of broker no votes and the overall percentage voting. With a broker no-vote percentage of 19% and 100% voting, activists need to get 62% of the instructed shares to get a majority. With a broker no-vote percentage of 19% and 80% voting, activists need to get 66% of the instructed shares to get a majority.

196. For background and a similar point regarding notice and access as compared with issuer proxy access, see Jeffrey N. Gordon, *Proxy Contests in an Era of Increasing Shareholder Power: Forget Issuer Proxy Access and Focus on E-Proxy*, 61 VAND. L. REV. 475, 487 (2008).

197. *See* 17 C.F.R. § 240.14a-16 (2009) (dictating the requirements for making proxy statements available on the Internet). Moreover, as of 2008, large accelerated filers and, as of 2009, everyone else, are required to post materials on a Web site.

198. Internet Availability of Proxy Materials, Exchange Act Release No. 52,926, 86 SEC Docket 2145 (Dec. 8, 2005).

199. *See* Fairfax, *supra* note 173, at 1285–86 (observing that e-proxy rules give individual shareholders more control since shareholders are able to "control both the content of the proxy statement and the solicitation process").

The notice and access option is especially important when shareholders cannot rely on the Rule 14a-2(b)(1) exemption discussed above. Under that exemption, shareholders can already engage in an effective proxy campaign without furnishing a proxy statement in support of a shareholder proposal, in opposition to a board proposal, or to withhold votes for board nominees to the board of directors.[200]

Rule 14a-2(b), however, has limitations. Most importantly, the exemption does not cover contests to elect a dissident slate to the board of directors.[201] Right now, such contests tend to be full-blown campaigns—such as Trian's 2006 campaign to elect some of its nominees to the board of Heinz[202]—often conducted by a hedge fund or in the context of a takeover bid. For full-fledged campaigns, the printing and mailing savings from notice and access is likely immaterial.[203] Thus, we do not expect that notice and access will lead to a significant increase in the number of such contests.

Rather, the most important impact of the notice and access rule may be to spur a new type of lower key, lower cost contest, probably for a "short slate" minority representation on the board. The campaigns in these lower key election contests are likely to resemble the campaigns currently waged in support of shareholder resolutions, in opposition to board proposals, or to withhold votes in favor of directors, which are now conducted in reliance on the Rule 14a-2(b)(1) exemption.[204] Unlike full-fledged campaigns, these campaigns are often run by cost-conscious traditional institutional investors.[205] The existence of such campaigns shows that there is some de-

---

200. *See supra* notes 169–80 and accompanying text.

201. *See* 17 C.F.R. § 240.14a-2(b)(1) (stating that the exemption does not apply to nominees for the board of directors and persons acting on their behalf). Moreover, it is practically impossible to get a dissident elected without distributing one's own proxy cards. *See* Stephen Taub, *Dissidents Win Proxy Fight, Without Proxy*, COMPLIANCE WK., Apr. 10, 2007, http://www.complianceweek.com/article/3253?printable=1 (noting how rare it is for dissidents to win proxy fights without proxy materials, especially at large companies). Other campaigns that cannot be effectively conducted under Rule 14a-2(b) relate to campaigns by shareholders who must file a Schedule 13D (mostly 5% shareholders with an activist agenda), campaigns in favor of shareholder resolutions that the company excluded from its proxy statement under Rule 14a-8, or campaigns where for strategic reasons the proponents want to distribute their own proxy forms. 17 C.F.R. §§ 240.14a-2(b)(1), 240.14a-8. Note that campaigns related to mergers, which are not covered by the 14a-2(b) exemption are also not subject to the notice and access rule. *Id.* § 240.14a-16(m).

202. *See* Andrew R. Sorkin, *Enough Anger to Make Ketchup Boil: Raider in a Bruising Fight with Heinz*, N.Y. TIMES, July 27, 2006, at C1 (describing Trian's strategy in its fight for control of Heinz's board and noting that Heinz's shares had risen 20% since Trian publicly announced its campaign).

203. Trian estimated that its total expenses in conducting the proxy contest would be $7 million. THE TRIAN GROUP, PROXY STATEMENT 27 (2006). In our estimate, the printing and mailing costs of the proxy statement do not amount to a significant portion of these expenses.

204. *See* GEORGESON, ANNUAL CORPORATE GOVERNANCE REVIEW 48 (2007), http://www.georgesonshareholder.com/emea/research/4a)%20US%20Corp%20Gov%202007.pdf [hereinafter 2007 ANNUAL CORPORATE GOVERNANCE REVIEW] (listing thirteen other activist campaigns).

205. *See "Just Vote No" Campaigns in Uncontested Director Elections—Renewed Vitality for the 2010 Proxy Season*, Client Memorandum from Wilkie, Farr, and Gallagher 1 (Sept. 24, 2009)

mand by investors for activism that goes beyond making (or voting for) a mere shareholder proposal under Rule 14a-8 but does not go as far as a full-fledged contest. For similar campaigns to elect directors, the costs savings resulting from not having to print and mail proxy statements may well be significant.[206] Notice and access opens up a wider range of issues to these intermediate-intensity campaigns.

D. The Roller-Coaster Ride of Proxy Access

Proxy access refers to a requirement that a company include director candidates nominated by shareholders in the company's proxy statement.[207] In 2003, the SEC approved a complex proposal for rules on proxy access and solicited comments on the proposal.[208] This proposal followed earlier considerations of the issue—in 1942 and 1977—which did not result in regulatory action.[209]

Under the 2003 proposal, proxy access was subject to a number of limitations. Any shareholder proxy access was conditioned on the prior occurrence of a "triggering event"—specifically a 35% or more "withhold" vote for a director or a majority vote electing to subject the company to proxy access.[210] Such proxy access would be limited to two years after such a trigger had occurred.[211] And during these two years, only shareholders who held at least 5% of a company's stock continuously for two years could obtain proxy access and could nominate only a minority slate.[212]

The 2003 proposal was initially supported by three of the five commissioners: the Republican Chairman Donaldson and the two Democratic commissioners.[213] But the proposal elicited strong negative reactions from managerial interests, including the Business Roundtable (an association of CEOs of leading U.S. companies) and the Chamber of

---

(on file with authors) (hypothesizing that the low-cost "just vote no" campaigns are likely to be widely used by institutional investors, possibly even as an alternative to "short slate" proxy contests).

206. *See* Corporate Governance: A Seismic Shift in the Mechanics of Electing Directors, Client Memorandum from David A. Katz & Laura A. McIntosh, Wachtell, Lipton, Rosen & Katz n.20 (July 27, 2006) (noting that, in the authors' experience, mailing costs are a substantial part of the dissidents' cost in a proxy fight).

207. DIV. OF CORP. FIN., SEC, STAFF REPORT: REVIEW OF THE PROXY PROCESS REGARDING THE NOMINATION AND ELECTION OF DIRECTORS 1–2 (2003).

208. *Id.* at 1, 5.

209. *Id.* at 2–3, 5.

210. Security Holder Director Nominations, 68 Fed. Reg. 60,790 (proposed Oct. 23, 2003) (to be codified at 17 C.F.R. pts. 240, 249, 274).

211. *Id.* at 60,794.

212. *Id.*

213. *See* Jonathan Peterson, *SEC Offers Conflicting Shareholder Proposals*, L.A. TIMES, July 26, 2007, at 3 (stating that Chairman Donaldson and the SEC's Democratic commissioners supported the SEC's 2003 proposal).

Commerce,[214] and Donaldson's support waned.[215] When Donaldson resigned as chairman in 2005,[216] the practical effect was that the proposal, which had been lingering in limbo for some time, was considered dead.[217]

A hard battle had been fought between proponents and opponents of greater shareholder rights, and the Business Roundtable had won—or so it seemed. Curiously, however, majority voting for directors—which started spreading at about the time of Donaldson's resignation and is now in place in most large companies[218]—gives shareholders many of the same powers and in a more useful form.[219] Most importantly, majority voting (like proxy access) gives shareholders the power to "diselect" a director from the board without having to file a proxy statement with the SEC.[220] Furthermore, while the proxy-access proposal was subject to limitations, majority voting is not so constrained.

To be sure, majority voting differs from proxy access in that shareholders cannot pick the director to replace the one they diselect. But this may be a net advantage. First and foremost, shareholders will have a much easier time agreeing on diselecting a director than agreeing both on

---

214. *See* Bill Baue, *Opening Up Pandora's Box: SEC Proxy Roundtable Questions Role of Non-binding Resolutions*, SOCIALFUNDS, May 15, 2007, http://www.socialfunds.com/news/article.cgi/2293.html ("The SEC allowed the rule it proposed in October 2003, allowing shareowners proxy access to nominate directors in certain circumstance, to die on the vine due to opposition by the Business Roundtable and the US Chamber of Commerce, which threatened a lawsuit.").

215. *See* Stephen Labaton, *S.E.C. Rebuffs Investors on Board Votes*, N.Y. TIMES, Feb. 8, 2005, at C2 (explaining that Donaldson "no longer supported the proposal put forward by the commission in 2003").

216. Stephen Labaton, *Donaldson Announces Resignation as S.E.C. Chairman*, N.Y. TIMES, June 1, 2005, *available at* http://www.nytimes.com/2005/06/01/business/01wiresec.html?ex=1275278400&en=d89d9d8be5440394&ei=5090&partner=rssuserland&emc=rs.

217. *Id.*

218. *See supra* notes 148–63 and accompanying text.

219. For a similar opinion, see Rachel McTague, *Grundfest: Proxy Access Unnecessary in View of Issuers' Shift to Majority Voting*, CORP. L. DAILY, Nov. 19, 2007, http://corplawcenter.bna.com/ pic2/clb.nsf/id/BNAP-78ZV9X?OpenDocument.

220. *See* Louis M. Thompson, Jr., *Shareholder Democracy to March On in '07*, COMPLIANCE WK., Nov. 21, 2006, http://www.complianceweek.com/article/2935/shareholder-democracy-to-march-on-in-07 (comparing the power to remove directors under the "majority-voting concept" with proxy access for director nominations). With the change in NYSE rules eliminating discretionary broker votes in director elections, another distinction between proxy access and majority voting will be eliminated. With proxy access, a vote of directors where shareholders nominated a competing slate would not have been viewed as routine, and brokers would not have discretionary voting power. *See* David A. Cifrino et al., *SEC Eliminates Broker Discretionary Voting in Director Elections, Proposes Changes to Disclosure & Other Requirements Regarding Corporate Governance & Compensation*, MCDERMOTT NEWLS., July 6, 2009, http://www.mwe.com/index.cfm/fuseaction/publications.nldetail/object_id/a89585fc-a483-4fed-9358-7cccf7b00616.cfm (explaining that, in the current system, "elections are already considered 'non-routine' matters on which discretionary voting is not allowed"). With the new NYSE rules, brokers do not have discretionary voting power in regular elections either. *See id.* (explaining that the change "eliminates 'discretionary voting' for all director elections").

rejecting the board's nominee *and* on replacing her with a specific person. As a result, majority voting gives shareholders a much more useful tool than proxy access. Moreover, it is often the ability to remove an offensive director (and the ability to threaten such a removal), rather than the ability to pick a replacement, that shareholders are really after. This is all the more so because any shareholder nominees would have to have broad appeal to maximize their chances of getting elected and would thus likely be drawn from the same pool of candidates as regular directors.

And even if what shareholders really want is to elect someone of their choice to the board, they have won half the battle. The proxy-access proposal would have spared shareholders who wanted to conduct a proxy contest the costs of preparing, printing, and mailing a proxy statement. Notice and access similarly saves shareholders printing and mailing costs (albeit not preparation costs) and does so without any of the limitations that were part of the proxy-access proposal.

In any case, there is an aftermath. Like the Sorcerer's Apprentice, the SEC could not control the forces it set in motion. In 2005, the American Federation of State, County and Municipal Employees (AFSCME) submitted its own homemade proposal for proxy access under Rule 14a-8 to American International Group (AIG).[221] The SEC's Division of Corporate Finance issued a no-action letter permitting AIG to omit the proposal under Rule 14a-8(i)(8).[222] In a stinging opinion issued in September 2006, the Court of Appeals for the Second Circuit rejected the SEC's reasoning as inconsistent with the SEC's own prior interpretations of its rules and held that the proposal could not be excluded.[223] The SEC immediately announced that it would consider amending Rule 14a-8.[224] After several delays,[225] in July 2007, the SEC finally approved two alternative proposals for public comments, each by a 3–2 vote with the new chairman Cox once siding with the two other Republican commissioners and once with the two Democrats.[226] The first proposal would have put on a firmer regulatory

221. AFSCME v. AIG, 462 F.3d 121, 123–24 (2d Cir. 2006).

222. *Id.* at 124.

223. *See id.* at 129 (upbraiding the SEC for failing to acknowledge its changed position regarding the excludability of proxy-access bylaw proposals and for failing to offer a reasoned basis for the change).

224. Press Release, SEC, Commission Calendars Proposed Amendment to Rule 14a-8 Governing Director Nominations by Shareholders (Sept. 7, 2006), *available at* http://www.sec.gov/news/press/2006/2006-150.htm) (last modified Sept. 7, 2006.

225. Since October 2006, the SEC has delayed scheduled consideration of proxy access at least twice. *See Atkins Says SEC Roundtable Likely on Proxy Access Issue; Time Not Yet Set*, 39 Sec. Reg. & L. Rep. (BNA) No. 11, at 379 (Mar. 12, 2007). In the meantime, there has been no groundswell of shareholder proposals resembling AFSCME's in the 2007 season—a fact quite consistent with our view that majority voting (and, to a lesser extent, notice and access) has made the fate of the proxy-access rule largely irrelevant.

226. Stephen Labaton, *A Public Airing for Proposals on Shareholders*, N.Y. TIMES, July 26, 2007, at C3.

footing the SEC's position rejected by the Second Circuit that shareholder proposals on proxy access can be excluded under Rule 14a-8(i)(8); the second resembled the 2003 proposal for proxy access.[227] In November 2007, the SEC adopted the first proposal by a party-line 3–1 vote.[228]

But with the 2008 election of President Obama, tables—and the party makeup of the SEC—turned again. Now with a majority of Democrats, the SEC on June 10, 2009, voted 3–2 (again along party lines) to approve for public comment another proxy-access proposal (with somewhat fewer limitations than the 2003 proposal) and for good measure a proposal to reverse the November 2007 amendments to Rule 14a-8(i)(8).[229] Because of the number of comments it received, the SEC has indicated that it will not act until spring of 2010.[230]

VI. Changes in the Board of Directors

Up to now, our analysis has focused on developments related to CEOs losing power to shareholders. In this Part, we discuss several developments related to CEOs losing power to corporate boards. Specifically, we will examine the 2003 amendments to the NYSE and NASDAQ listing standards, changes in nominal director independence, and changes in substantive director independence.

A. Listing Standards

In 2003, in the wake of the Enron scandal and the Sarbanes–Oxley Act of 2002, the NYSE and the NASDAQ Stock Market adopted new governance rules for listed companies.[231] Both sets of rules now require

---

227. *See* Nicholas Rummell, *SEC Splits Proxy Access Votes as Cox Says 'Yea' to Two Proposals*, FIN. WK., July 25, 2007, http://www.financialweek.com/apps/pbcs.dll/article?AID=/20070725/REG/70725013/1036 (describing how the first proposal would allow shareholders to put forth proxy proposals calling for bylaw changes to allow shareholder-approved directors during corporate elections and how the second would restate the SEC's position prior to the Second Circuit's invalidation).

228. *The SEC Denies Proxy Access*, Posting of L. Reed Walton to RiskMetrics Group, http://blog.riskmetrics.com/gov/2007/11/the-sec-denies-proxy-accesssubmitted-by-l-reed-walton-publications.html (Nov. 30, 2007, 10:33 EST).

229. Georgeson Inc. & Latham & Watkins LLP, *Proxy Access Proposed Rules Published by SEC*, CORP. GOVERNANCE COMMENT., June 15, 2009, http://www.georgeson.com/usa/download/corp_gov_commentary.html.

230. Mary Schapiro, Chairman, SEC, Address to the Practising Law Institute's 41st Annual Institute on Securities Regulation (Nov. 4, 2009) (on file at http://sec.gov/news/speech/2009/spch110409mls.htm).

231. Goodwin Procter, *SEC Approves Final NYSE and Nasdaq Corporate Governance Standards*, PUB. CO. ADVISORY, Nov. 11, 2003, at 1, http://www.goodwinprocter.com/~/media/Files/Publications/Newsletters/Public Company Advisory/2003/SEC_Approves_Final_NYSE_and_Nasdaq_Corporate_Governance_Standards.ashx.

boards of most companies[232] to consist of a majority of "independent" directors (albeit with somewhat varying definitions of "independence"[233]), to establish an audit committee consisting entirely of independent directors,[234] and to conduct regular separate meetings ("executive sessions") of the independent directors.[235] The NYSE rules further require that each board have nominating/corporate-governance and compensation committees consisting entirely of independent directors.[236] The NASDAQ rules do not require boards to establish such committees, but if a company does establish a nominating or a compensation committee, it must consist entirely of independent directors.[237] Both sets of rules became effective in January 2004 for some companies, and later for others.[238] And the Sarbanes–Oxley Act itself requires that each listed company have an audit committee consisting entirely of independent directors.[239]

B. Nominal Board Independence

As Jeff Gordon has recently shown, the nominal independence of board members has increased dramatically since the 1950s.[240] Gordon estimates that the percentage of inside directors has steadily decreased from 50% in 1950 to around 10% in 2005 and that the percentage of independent directors has correspondingly increased from around 20% to around 80%.[241]

What is less clear, however, is whether there has been a significant change in board makeup over the last ten years and, if so, whether any change is attributable to the changed listing requirements. Korn/Ferry, which conducts annual reviews of proxy statements of Fortune 1000 companies, reports that the average number of insiders on boards remained steady at two between 1997 and 2007, while the average number of outsiders has declined

---

232. Boards of certain controlled companies are exempt. NYSE, Inc., Listed Company Manual § 303A.0 (2003) ("A listed company of which more than 50% of the voting power is held by an individual, a group or another company need not comply.").

233. NASDAQ, Inc., Rule 5605(b)(1); NYSE, Inc., Listed Company Manual § 303A.01–.02.

234. NASDAQ, Inc., Rule 5605(c)(2)(A); NYSE, Inc., Listed Company Manual § 303A.07(b).

235. NASDAQ, Inc., Rule 5605(b)(2); NYSE, Inc., Listed Company Manual § 303A.03.

236. NYSE, Inc., Listed Company Manual § 303A.04–.05.

237. NASDAQ, Inc., Rule 5605(d)–(e). The NASDAQ rules do not address the composition of any separate corporate-governance committee.

238. Goodwin Procter, *SEC Expected to Approve Final Nasdaq Corporate Governance Standards*, PUB. CO. ADVISORY, Oct. 21, 2003, http://www.goodwinprocter.com/~/media/Files/Publications/Newsletters/Public Company Advisory/2003/SEC_Expected_to_Approve_Final_Nasdaq_Corporate_Governance_Standards.ashx.

239. 15 U.S.C. § 78j-1 (2006).

240. Jeffrey N. Gordon, *The Rise of Independent Directors in the United States, 1950–2005: Of Shareholder Value and Stock Market Prices*, 59 STAN. L. REV. 1465, 1468 (2007).

241. *See id.* at 1473–75 (presenting a methodology and graphical data demonstrating the trend from insider to independent directors).

from nine to eight.[242] Another source of data is the Investor Responsibility Research Center (IRRC) (now part of RiskMetrics), which categorizes each director as an employee of the company, a linked director (a former employee, family member of an employee, or a director who provides, or whose employer provides, services to the company, or is a significant customer), or an independent director.[243] We collected information of these categorizations for the years 2000 and 2007 for companies in the S&P 500 Index, for the Midcap (S&P 400) Index, and for the SmallCap (S&P 600) Index. The IRRC data shows a decline of average total board size for S&P 500 companies (but not for companies in the other indices), as well as a decline in the number of employee directors from about 2.1 to 1.5. Depending on the index, the average percentage of employee directors declined from 18% to 24% in 2000 to 14% and 18% in 2007. Linked directors experience a steeper decline, from around 1.3 to 1.6 in 2000 to 0.6 in 2007, while the number of directors categorized as independent increased. For all companies combined, the percentage of linked directors declined from 14.5% to 6.4% over this seven-year period.

Table 5: Nominal Director Independence

	Employee	Independent	Linked	Total
	2007	2007	2007	2007
SP 500	1.5	8.5	0.6	10.6
SP 400	1.5	7.2	0.6	9.3
SP 600	1.5	6.2	0.6	8.3
	2000	2000	2000	2000
SP 500	2.1	7.9	1.6	11.5
SP 400	2.0	5.9	1.4	9.3
SP 600	2.0	5.0	1.3	8.4

Compared to the longer period investigated by Gordon, it thus appears that the move away from employee directors largely preceded 2000. However, post-2000, there was a significant drop in linked directors.

There appears to have been no change in the makeup of the key committees. According to Korn/Ferry data, in each year between 1997 and 2006, neither audit, nor compensation, nor nominating, nor corporate-governance committees have had (on average) *any* insider directors (the average number of outside directors on each committee varied between three

242. KORN/FERRY INT'L, 32ND ANNUAL BOARD OF DIRECTORS STUDY 36 (2007) [hereinafter 32ND ANNUAL BOARD OF DIRECTORS STUDY], *available at* http://www.kornferry.com/Library/ViewGallery.asp?CID=1573&LanguageID=1&RegionID=23.

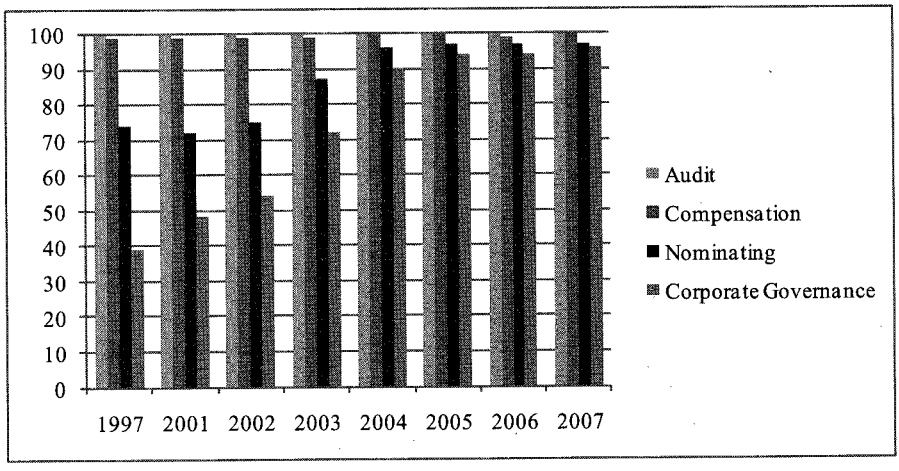
243. IRRC Legacy Directors Database, Overview of IRRC Directors in WRDS, http://wrds-web.wharton.upenn.edu/wrds/support/Data/_001Manuals%20and%20Overviews/_115RiskMetrics/Overview%20of%20IRRC%20Legacy%20Directors%20Database%20on%20WRDS.cfm.

and four).[244] Moreover, according to IRRC data, even in 2000, 60% of directors in S&P 600 companies, 64% of directors in S&P 400 companies, and 69% of directors in S&P 500 companies were independent.[245] Thus, it is likely that most companies fulfilled the requirements of the new 2004 listing standards for committee and board composition several years prior to their adoption.

C. *Substantive Board Independence I: What Do Boards Do?*

More important than nominal board independence, however, is whether boards are substantively independent: whether directors act as independent decision makers, rather than as yes-men for the CEO. One way to get a handle on whether boards have become more substantively independent of the CEO is to examine what boards spend their time on. Specifically, boards that spend relatively more time monitoring the CEO are likely to be more substantively independent, and the CEOs of companies with such boards are likely to be less powerful.

Figure 2: Percentage of Companies with Board Committees: Audit, Compensation, Nominating, and Corporate Governance



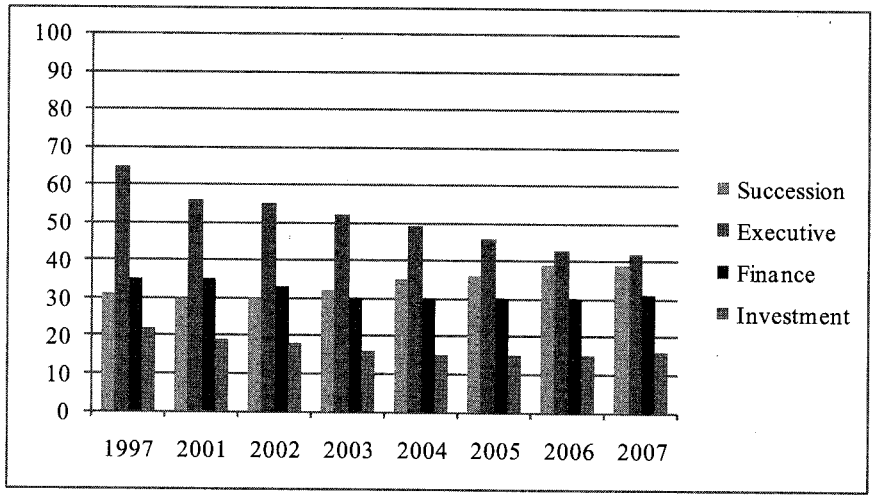
There are several useful metrics for determining what boards spend their time on. One important measure is whether a board has established a committee devoted to certain tasks and how frequently that committee meets. Virtually all larger companies have had audit and compensation committees

244. KORN/FERRY INT'L, 33RD ANNUAL BOARD OF DIRECTORS STUDY 13 tbl.D (2006) [hereinafter 33RD ANNUAL BOARD OF DIRECTORS STUDY], *available at* http://www.kornferry.com/Publication/3322.

245. IRRC Legacy Directors Database, *supra* note 243.

for a significant period of time.[246] But the number of companies with nominating and corporate-governance committees has increased significantly. According to Korn/Ferry, the percentage of companies with nominating committees hovered in the low- to mid-seventies until 2002, increased to 87% in 2003, and further increased to over 95% from 2004 on.[247] The percentage of companies with corporate-governance committees (which are not regulated by NASDAQ standards) gradually increased from 39% in 1997 to 48% in 2001, and then increased at a more rapid rate to 96% in 2007.[248] The changed NYSE and NASDAQ listing requirements presumably account for at least a portion of this increase. Many companies, however, had added these committees before they were required to do so.[249] The trend in corporate-governance committees, not required by Sarbanes–Oxley or NASDAQ listing standards, showing an increase even in the pre-Sarbanes–Oxley period, suggests that a significant portion of the increase may be unrelated to the changed standards.

Figure 3: Percentage of Companies with Board Committees: Succession, Executive, Finance, and Investment



246. Of companies participating in the Korn/Ferry survey, 100% had audit and 99% had compensation committees by 1995. 33RD ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 244, at 12 tbl.C.

247. *Id.*

248. KORN/FERRY INT'L, 34TH ANNUAL BOARD OF DIRECTORS STUDY 18 tbl.C (2007) [hereinafter 34TH ANNUAL BOARD OF DIRECTORS STUDY], *available at* http://www.kornferryinstitute.com/leadership/board_of_directors/publication/1225/34th_Annual_Board_of_Directors_Study.

249. *See* James S. Linck et al., *The Effects and Unintended Consequences of the Sarbanes–Oxley Act on the Supply and Demand for Directors*, 22 REV. FIN. STUD. 3287, 3288 (2009), *available at* http://rfs.oxfordjournals.org/cgi/reprint/22/8/3287.pdf ("[C]hanges in boards and directors have been occurring for some time."); *see also id.* at 3292 (enumerating the recently mandated "major governance provisions").

Another interesting trend can be observed by looking at some other committees. The three committees included in the Korn/Ferry data that relate to "management"—the executive committee, the finance committee, and the investment committee—experienced a steady decline.[250] By contrast, the one committee charged with monitoring functions that is not affected by the changed listing standards—the succession committee—experienced a steady (if slow) increase from 31% in 1995 to 39% in 2007.[251]

A further indicator of whether these committees serve as window dressing or whether they perform important functions is the number of times they meet. As Table 6 indicates, the number of meetings of committees with monitoring functions—the audit, compensation, nominating, corporate-governance, and succession committees—has generally increased.[252] With the exception of the audit committees, this increase does not seem to be due to an increased burden placed on these committees by the Sarbanes–Oxley Act. Rather, the number of meetings increased at approximately the same rate in the pre-Sarbanes–Oxley period (1997–2001) as in the post-Sarbanes–Oxley period (2002–2006). By contrast, the number of meetings of the committees with management functions—executive, finance, and investment—has largely remained steady.[253]

Table 6: Committee Meetings per Year

	1997	2001	2007
Audit	3	4	9
Compensation	4	5	6
Nominating	2	3	4
Corporate Governance	3	3	4
Succession	5	5	6
Executive	4	4	4
Finance	5	4	5
Investment	4	4	5

250. 34TH ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 248, at 18 tbl.C.

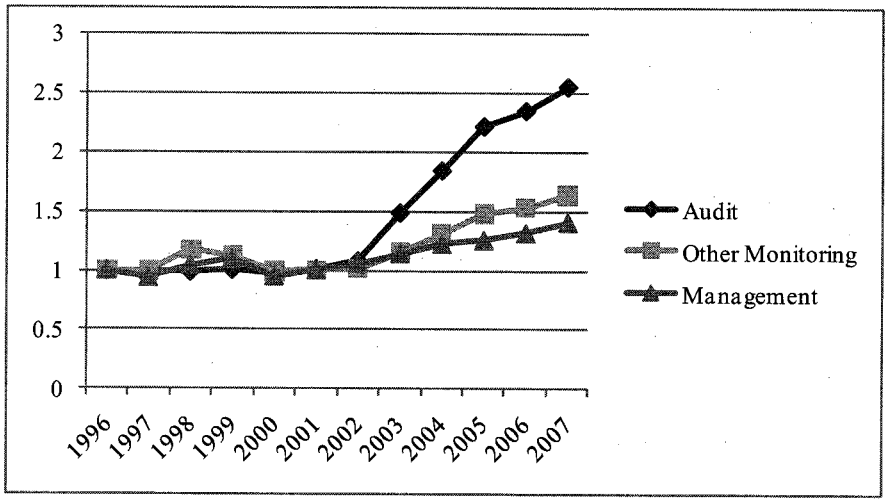
251. The only other committees included in the Korn/Ferry data but not included in Figure 2 or 3 are the "Corporate Responsibility" committee, which experienced a slight decline, and the "Director Compensation" committee, which experienced a major increase. *Id.*

252. KORN/FERRY INT'L, 30TH ANNUAL BOARD OF DIRECTORS STUDY 13 tbl.F (2003) [hereinafter 30TH ANNUAL BOARD OF DIRECTORS STUDY], *available at* http://www.kornferry institute.com/about_us/thought_leadership_library/publication/1492/30th_Annual_Board_of_Directors_Study; 34TH ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 248, at 19 tbl.E.

253. 30TH ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 252, at 13 tbl.F; 34TH ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 248, at 19 tbl.E.

We also examined the compensation received by board members for serving on various committees. As a measure of compensation, we used the retainer received by the committee chair because cash compensation levels for that measure were available for each committee in most years.[254] Between 1996 and 2001, compensation for committee service adjusted for inflation barely budged. Average compensation (adjusted for inflation) changed by less than 1% per year for all committees combined, all committees but the audit committee, the four other monitoring committees, and the three management committees.[255] But between 2001 and 2007, the picture is starkly different. Compensation for the chair of the audit committee increased on average by 17% a year over the six years, and by 11% a year since 2004.[256] Compensation for the four other monitoring committees also increased by an average total of 64% over the six-year period. In contrast, compensation for the three management committees increased by only 40%.[257]

Figure 4: Committee Compensation Trends[258]



254. Because information for 2001 was not available for some committees, we interpolated the figures for 2000 and 2002.

255. 30TH ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 252, at 15 tbl.1; 34TH ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 248, at 21 tbl.H.

256. 34TH ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 248, at 21 tbl.H.

257. *Id.* at tbls.H–I. Another fact of perhaps symbolic significance: between 1996 and 2002, the highest average retainer (usually by a large margin) was paid to the chair of the executive committee, traditionally the CEO. By 2005, the average retainer of the chair of the executive committee was less than the retainer for the chair of each of the five committees with monitoring tasks. *Id.* at tbl.I.

258. The figure is based on cash retainer for committee chair. All amounts are adjusted for inflation. Other monitoring committees are compensation, nominating, corporate governance, and succession. Management committees are executive, finance, and investment.

These changes indicate a shift in what the board is doing. Rather than help the corporate insider with managing the business of the corporation, boards are now increasingly engaged in monitoring management and planning for management changes.

Some other data in the Korn/Ferry survey provide additional evidence that outside directors work harder. Survey responses indicate that between 1997 and 2007 the number of hours worked per month increased from thirteen to sixteen.[259] In 2004 and 2005, when Korn/Ferry instead asked whether the board had more meetings than in the prior year, 29% and 34%, respectively, responded yes.[260]

D. Substantive Board Independence II: Changed Board Dynamics

Over the last few years, boardroom dynamics have changed, with outside directors emerging as a power center independent of CEOs. Until recently, outside directors never met without the CEO present and received most of their information from management. This insider control of the information flow, both to and among outside directors, has diminished. Nowadays, it is not unusual for directors to meet with significant shareholders and even with employees.[261] In some instances, they even hire outside consultants to review business plans presented by management.[262] In addition, since 2004, outside directors are required by stock-exchange rules to meet in "executive sessions" outside the presence of the CEO.[263] According to reports, "directors who fear a company is heading off course can use executive session meetings to reinforce each others' concerns and settle on a plan of action"—including, on occasion, a plan to fire the CEO.[264]

Responses in the Korn/Ferry survey confirm this change in board dynamics. The percentage of boards with a formal process for evaluating CEOs increased from the high sixties in 1997 and 2001 to around 92% in 2007.[265] The percentage of boards with a lead outside director (if the CEO is

---

259. *See* 30TH ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 252, at 24; 32ND ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 242, at 53; 33RD ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 244, at 23; 34TH ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 248, at 34.

260. 32ND ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 242, at 53.

261. Kaja Whitehouse, *Move Over CEO: Here Come the Directors*, WALL ST. J., Oct. 9, 2006, at R1; *see also* Lipton, *supra* note 130, at 7 (discussing increased demand by public pension funds and other activists to meet with independent directors).

262. Whitehouse, *supra* note 261.

263. NASDAQ, Inc., Rule 5605(b)(2); NYSE, Inc., Listed Company Manual § 303A.03 (2003).

264. George Anders, *Private Time*, WALL ST. J., Oct. 9, 2006, at R4.

265. 30TH ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 252, at 22; 32ND ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 242, at 22, 63; 33RD ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 244, at 9; 34TH ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 248, at 12 chart D.

also the chairman) increased from around 30% in 2002 to 84% in 2007.[266] And according to an annual survey conducted by the Business Roundtable, 90% of companies had an independent chairman, lead director, or presiding director in 2007 (up from 83% in 2005 and 71% in 2004).[267]

The latter increase could be attributable to the requirement that independent directors meet in executive sessions.[268] Though there is no requirement for a lead director, a board may find it useful to appoint a lead director to run these meetings. But we think more is going on. For one, the percentage of respondents who said that companies *should have* a lead outside director increased from 55% in 2001 to 84% in 2006.[269] Second, the percentage of companies with a lead director *out of those that conduct executive sessions* increased from 34% in 1997 to 78% in 2001 to 85% in 2007. This indicates that the increase in lead directors is not merely a pragmatic adjustment to the requirement to hold executive sessions but also reflects a change in the board attitude that a greater dispersion of power—away from the CEO and towards the independent directors—is desirable.

Finally, the long-standard U.S. practice of having the CEO also serve as Chairman of the Board seems to be eroding.[270] According to the Business Roundtable Survey, the percentage of companies that had split the CEO and Chairman positions increased from 4% in 2004 to 13% in 2007.[271] And our own review of S&P 100 companies indicates that the percentage of companies with split positions increased from 18% in 2003 to 26% in 2006.

Perhaps the most telling indicator that boardroom dynamics have changed is the annual list of "Key Issues for Directors" prepared by Martin Lipton. In 2007, the number one item on the list was "Anticipating attacks by activist hedge funds."[272] For 2008, attacks by activist hedge funds had dropped to number seven (of nine) and a new entry headed the list: "Maintaining collegiality and the culture of common enterprise with the CEO and senior management."[273]

E. Substantive Board Independence III: CEO Turnover

As another indicator of the greater substantive independence of the board of directors, CEO tenure is declining. According to a 2007 report pre-

---

266. 32ND ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 242, at 54; 34TH ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 248, at 7.

267. BUS. ROUNDTABLE, BUSINESS ROUNDTABLE CORPORATE GOVERNANCE SURVEY KEY FINDINGS (2007).

268. *See supra* note 235 and accompanying text.

269. 30TH ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 252, at 20; 33RD ANNUAL BOARD OF DIRECTORS STUDY, *supra* note 244, at 24.

270. Lipton, *supra* note 130, at 17.

271. BUS. ROUNDTABLE, *supra* note 267.

272. Memorandum from Martin Lipton, Wachtell, Lipton, Rosen & Katz, to Clients 1 (Dec. 6, 2006) (on file with Texas Law Review).

273. Memorandum from Martin Lipton, *supra* note 76, at 1.

pared by Booz Allen Hamilton, directors are "becoming more critical . . . and are far more likely to insist that CEOs deliver acceptable shareholder returns."[274] Importantly, Booz Allen finds that boards are increasingly prepared to replace CEOs *in anticipation of disappointing future performance*, rather than in response to poor past performance.[275] For 2006, total turnover (which includes turnover due to retirement, dismissal, and acquisition) was 14.3%.[276] Among the other specific findings, Booz Allen reports that, between 1995 and 2006, annual turnover of CEOs had increased by 59% and performance-related turnover by 318%.[277] Correspondingly, the fraction of CEOs who were forced from office increased from one out of eight to nearly one out of three.[278]

A study by Steve Kaplan and Bernadette Minton arrives at similar conclusions.[279] Kaplan and Minton find a total turnover rate, including both external (takeover related) and internal (nontakeover related) turnover, of 17.4% and an internal turnover rate of 12.6% for 1998–2005, which corresponds to an average CEO tenure period of as low as six years.[280] This tenure, the authors say, is substantially shorter than the ones reported in previous work for the 1970s, 1980s, and 1990s.[281] They conclude that boards respond more broadly to poor performance than they have in the past and monitor more frequently and aggressively.[282]

In our analysis of S&P 100 companies, we found that of the ninety-six companies that had not been acquired between March 2006 and March 2008, nineteen had a turnover in CEOs. This corresponds to a somewhat higher internal takeover rate than reported in the study above.[283] Of these nineteen changes, one can be classified as a promotion (Goldman Sachs's CEO became Secretary of the Treasury), and nine (based on press reports) as

274. Chuck Lucier et al., *The Era of the Inclusive Leader*, STRATEGY + BUS., Summer 2007, at 2.

275. *Id.*

276. *Id.* at 3.

277. *Id.*

278. *Id.*

279. Steven N. Kaplan & Bernadette A. Minton, How Has CEO Turnover Changed? 1 (Aug. 2008) (unpublished manuscript, on file at http://faculty.chicagobooth.edu/steven.kaplan/research/km.pdf).

280. *Id.* at 2.

281. *Id.* at 4.

282. *See id.* at 4 (explaining boards' broader and more immediate responses to poor market performance as well as poor industry performance).

283. Kaplan & Minton, *supra* note 279, at 1. Our analysis results in a 19.8% internal takeover rate for this period, as opposed to the 12.6% rate found from 1998–2005. Our study's rate covers two years, meaning that the rate per year is about half—less than Kaplan and Minton's reported annualized rate.

involuntary.[284] The remainder were claimed to be retirements. Using the academic convention of treating a "retirement" of a CEO who is sixty or older as voluntary and a "retirement" of a CEO under sixty as forced (unless the reported reason is health), a total of twelve changes can be classified as involuntary. Thus, our sample yields a somewhat higher estimate for involuntary turnover, both absolutely and as a fraction of total turnover, than the Booz Allen study.[285]

Increased CEO turnover is not only a symptom of increased substantive independence.[286] It is also a cause for further independence. As CEO turnover increases, the tenure of outside directors relative to the CEO increases. As Table 7 below shows, an increasing percentage of S&P 500 companies have significant fractions of outside directors whose tenure on the board precedes the CEO.

Table 7: CEO and Outside Director Tenure in S&P 500 Firms

	2000	2001	2002	2003	2004	2005	2006	2007
At least 25% of outsiders have tenure greater than CEO's (% of cos.)	47	50	50	52	53	58	58	54
At least 50% of outsiders have tenure greater than CEO's (% of cos.)	23	26	23	24	25	30	30	28
At least 75% of outsiders have tenure greater than CEO's (% of cos.)	5	6	5	4	5	8	8	8
Mean CEO tenure (years)	10.2	9.5	9.7	9.1	9.1	9	8	8.3
Mean Outside Director tenure (years)	8	8.1	7.9	7.8	7.8	7.9	7.8	7.7

Such relative tenure contributes to substantive independence in two respects. First, outside directors who became board members before the CEO should in no way feel that they owe their board seat to the CEO. Second, such outside board members will often have been involved in the selection of the CEO.[287] Thus, rather than viewing themselves as having been hired by the CEO, they are more likely to view themselves as having hired the CEO, and are thus in a stronger position to contradict the CEO or even fire him.

284. *See* Michael Mandel, *Mr. Risk Goes to Washington*, BUS. WEEK, June 12, 2006, http://www.businessweek.com/magazine/content/06_24/b3988001.htm (describing Henry Paulson's resignation as CEO of Goldman Sachs Group in order to serve as Secretary of the Treasury).

285. *See* Lucier et al., *supra* note 274, at 3 (stating that nearly one in three CEOs left involuntarily in 2006).

286. It is possible that the increased turnover is exclusively due to other factors, such as more attractive severance packages for CEOs.

287. *See generally* Kenneth A. Borokhovich et al., *Outside Directors and CEO Selection*, 31 J. FIN. & QUANTITATIVE ANALYSIS 337 (discussing the role of outsiders in the selection of a new CEO).

VII. Executive Compensation: The Final Frontier

As another metric of changes in CEO power, we examine executive compensation. We discuss three developments: the disclosure rules adopted in 2006, the recent initiatives to give shareholders a "say on pay," and recent trends in CEO compensation.

A. Enhanced Disclosure

In July 2006, the SEC adopted new and enhanced disclosure requirements for executive compensation.[288] The new rules expand the previous disclosure regime in several ways. First, a proxy statement must contain a new section, "Compensation Discussion and Analysis," with a narrative discussion of objectives, design of compensation program, and how the company determines the amount of various compensation elements.[289] Second, more information is required for stock options and retirement benefits, including the fair value of these options on the date of grant.[290] Third, an enhanced summary-compensation table must provide a dollar value for each compensation item as well as elements for total compensation.[291] The last requirement, in particular, makes it harder to camouflage compensation by shifting it into categories that need not be quantified.[292] Prior to the 2006 reforms, the reported figure for "total compensation" did not include the value of stock awards, option grants, and retirement benefits.[293] And although some information on these items was disclosed elsewhere in the proxy statement, it was hard to decipher their dollar value.

Thus, for example, in 2005, GE reported that its CEO Jeffrey Immelt received "total compensation" of $3.4 million, that he was also granted 430,000 performance stock units (PSUs), and that he held PSUs and restricted stock with a value of $45.7 million as of December 31, 2005; but it did not disclose either the fair value of the PSUs granted in 2005 nor the total compensation including these PSUs for 2005.[294] In 2006 (postreform), GE disclosed that Immelt received total compensation of $17.9 million, a figure

---

288. Executive Compensation and Related Person Disclosure, Exchange Act Release No. 8,732A, 71 Fed. Reg. 53,158 (Sept. 8, 2006).

289. 17 C.F.R. § 229.402(b) (2009).

290. *Id.* § 229.402(a)(6)(iv), (d)(2)(ii).

291. *Id.* § 229.402(c)(2).

292. *See* Lucian Arye Bebchuk et al., *Managerial Power and Rent Extraction in the Design of Executive Compensation* 39–42 (Nat'l Bureau of Econ. Research, Working Paper No. 9068, 2002) (arguing that companies are trying to camouflage compensation paid to executives as options).

293. Kathryn Yeaton, *The SEC's New Rules on Executive Compensation*, CPA J., July 2007, at 26, 29.

294. Gen. Elec. Corp., Notice of 2006 Annual Meeting and Proxy Statement (Schedule 14A), at 28–29, 34–35 (Mar. 3, 2006).

that includes stock and option awards valued at $8 million and an increase in pension value of $1 million.[295]

B. "Say on Pay"

One of the latest shareholder-rights initiatives goes by the poetic label "say on pay."[296] Say on pay requires a company to give its shareholders a nonbinding, advisory vote on the compensation of its executives.[297] This could have serious ramifications. The combination of traditional institutional investors with various performance or governance gripes, union-affiliated pension funds that may be willing to campaign for a "say NO on pay" vote, and populist sentiments against executives and their high salaries—together with the disclosure requirements that make it harder to camouflage executive compensation[298]—means that CEOs could find their packages disapproved by shareholders. Moreover, given the recent trend of boards to heed shareholders requests,[299] even an advisory vote could be a significant threat to CEO pocketbooks. Unsurprisingly, management lawyers like Martin Lipton recommend that such votes be "strongly resisted."[300]

In 2007, the SEC ruled that shareholder proposals requesting boards to adopt say on pay are not excludable under Rule 14a-8.[301] According to ISS, the number of such proposals has skyrocketed from 0 in 2005, to 7 in 2006, 41 in the first half of 2007, and at least 67 in 2008[302]—the single most

---

295. Gen. Elec. Corp., Notice of 2007 Annual Meeting and Proxy Statement (Schedule 14A), at 21 (Feb. 27, 2007).

296. *See* Malini Manickavasagam, *Shareholder Proposals: Annual Meeting Voting Compels More Accountability*, 11 BNA CORP. GOV. REP. 30 (2008) (listing "say on pay" as one of top three issues on corporate ballots for 2008).

297. Sandeep Gopalan, *Say on Pay and the SEC Disclosure Rules: Expressive Law and CEO Compensation*, 35 PEPP. L. REV. 207, 220–21 (2008).

298. *See* 7 C.F.R. § 229.402(c)(2) (2009) (requiring corporations to report dollar amounts in a number of categories); Bebchuk et al., *supra* note 292, at 39–42 (discussing the use of options as a means of camouflaging executive compensation).

299. *See* Whitehouse, *supra* note 261 (observing that if shareholders make enough of an impact, directors will respond).

300. Lipton, *supra* note 130, at 8.

301. *See Proposals on Policy for 'Advisory' Votes Regarding Executive Pay Not Excludable*, 39 Sec. Reg. & L. Rep. (BNA) No. 9, at 370 (Mar. 5, 2007) (reviewing three separate no-action responses in which the staff of the Division of Corporate Finance advised AT&T Inc., Qwest Communications International Inc., and Clear Channel Communications Inc. respectively that they may not exclude proposals that the board adopt policies allowing shareholders to cast "advisory" votes on executive compensation from proxy materials for upcoming shareholder meetings).

302. *See Companies Ignore 'Say on Pay' Votes*, DIRECTORSHIP, July 23, 2008, http://www.directorship.com/companies-ignore-say-on-pay-votes (reporting seventy-six proposals so far in 2008); RISKMETRICS GROUP, POSTSEASON REPORT 8 (2007), http://www.riskmetrics.com/system/files/private/2007PostSeasonReportFINAL.pdf [hereinafter 2007 POSTSEASON REPORT] (comparing the forty "say on pay" proposals that were voted on between January 1, 2007, and June 30, 2007, with the seven proposals voted on in 2006); *id.* at 6 (illustrating in chart one the fact that in 2005 there were zero votes on proposals to give shareholders an advisory vote on executive compensation).

numerous category of proposals for that year.[303] Georgeson, which tracks proposals at a smaller set of companies, found 1 say on pay proposal in 2004, 39 in 2007, and 67 in 2008.[304]

Shareholder support for these proposals is high—the average proposal received 41.7% support in 2007—but not nearly as high as support for proposals to destagger the board (63.9% support) or to adopt majority voting (50.3% support).[305] Of 39 such proposals for which results were reported by Georgeson in 2007, only 4 (at Blockbuster, Ingersoll-Rand, Motorola, and Verizon) garnered a majority of the votes cast.[306] In 2008, 6 of 67 for which results were reported received a majority.[307] The fact that many proposals receive significant, but not majority, support partly explains why these proposals are so numerous. When proposals to destagger a board or to adopt majority voting are introduced, proposals shareholders regularly adopt, companies often agree to make the requested changes without a shareholder vote and thus remove the proposal from the ballot.[308] But because boards have a high chance of defeating a say on pay proposal, they have less of an incentive to adopt say on pay before a vote. At the same time, the level of support is sufficiently high for shareholders to keep introducing these proposals in order to put pressure on the board and gather momentum for an eventual passage.

The problem for boards—and CEOs, who would presumably be most affected by say on pay votes—is that for new types of shareholder proposals, the percentage of shares voted in favor and the number of proposals introduced tends to increase over time. Thus, for example, support for majority-voting proposals, also of relatively recent vintage, increased from 12% (on 12 proposals) in 2004, to 44% (on 54) in 2005, to 48% (on 84) in 2006, to 50% (on 37) in 2007.[309] For say on pay proposals, the number of proposals and the support they garner (no proposal in 2005, 40% on 7 proposals in 2006, 42% on 41 proposals in 2007, 42% on 62 proposals in 2008) seem to

303. *See* 2007 POSTSEASON REPORT, *supra* note 302, at 6 (observing that shareholders' say on pay proposals outnumbered eleven other types of proposals in 2007).

304. *See* GEORGESON, ANNUAL CORPORATE GOVERNANCE REVIEW 18 (2004), http://www.georgesonshareholder.com/usa/download/acgr/acgr2004.pdf [hereinafter 2004 ANNUAL CORPORATE GOVERNANCE REVIEW] (listing a single proposal to "approve executive compensation" in 2004); GEORGESON, ANNUAL CORPORATE GOVERNANCE REVIEW 5 (2008), http://www.georgesonshareholder.com/usa/download/acgr/acgr2008.pdf [hereinafter 2008 ANNUAL CORPORATE GOVERNANCE REVIEW] (comparing the number and success of say on pay proposals in 2007 with proposals in 2006).

305. 2007 POSTSEASON REPORT, *supra* note 302, at 6.

306. 2007 ANNUAL CORPORATE GOVERNANCE REVIEW, *supra* note 204, at 4.

307. 2008 ANNUAL CORPORATE GOVERNANCE REVIEW, *supra* note 304, at 5.

308. *See* 2007 POSTSEASON REPORT, *supra* note 302, at 3 (observing that these proposals are frequently withdrawn as companies become more willing to negotiate directly with shareholders on these issues).

309. 2004 ANNUAL CORPORATE GOVERNANCE REVIEW, *supra* note 304, at 6; 2007 POSTSEASON REPORT, *supra* note 302, at 4.

follow a similar pattern.[310] For January 1 to June 30, 2009, RiskMetrics estimates that say on pay proposals received 46% support on 71 proposals.[311]

Moreover, signs are that the board front against say on pay is starting to break. In 2008, Aflac Inc. became the first company to hold an advisory say on pay vote.[312] Verizon Communications, where a 2007 proposal received slightly more "for" than "against" votes (but less than majority support), and Blockbuster, where the proposal received majority support, decided to adopt say on pay for 2009,[313] as did Occidental Petroleum, Intel, Hewlett-Packard, MBIA, Motorola, and Ingersoll-Rand.[314]

The current economic crisis provides further impetus for say on pay. In times of declining stock prices and rising unemployment, high compensation for CEOs is an easy target, both for activist shareholders and politicians. Shareholder proposals submitted in 2009 have gathered unprecedented support, with ten of twenty-nine receiving a majority of the votes cast.[315] Under federal regulations, companies that receive federal TARP funds must hold say on pay votes if they want to pay their executives more than $500,000.[316] And under legislation proposed by Senator Schumer, all publicly traded companies would have to hold say on pay votes.[317]

C. *Actual Compensation*

Actual executive compensation may present the final frontier in the erosion of CEO dominance. Many commentators believe that CEOs, through their influence over the board, essentially set their own pay.[318] Even if one does not subscribe to the more extreme versions of the theory, which accords a minimal role to market forces in setting CEO pay, compensation is surely

---

310. 2007 POSTSEASON REPORT, *supra* note 302, at 6; RISKMETRICS GROUP, POSTSEASON REPORT 5 (2009), http://www.riskmetrics.com/system/files/private/2009_PSR_Public_final.pdf [hereinafter 2009 POSTSEASON REPORT].

311. 2009 POSTSEASON REPORT, *supra* note 310, at 5.

312. Joann S. Lublin, *Say on the Boss's Pay: Aflac CEO Amos Bets on His Track Record as Insurer Becomes First U.S. Company to Hold Vote on Executive Compensation*, WALL ST. J., Mar. 7, 2008, at B1.

313. George Anders, *'Say-on-Pay' Gets a Push, but Will Boards Listen?*, WALL ST. J., Feb. 27, 2008, at A2.

314. *Say-on-Pay Is on the Way*, SMARTPROS, Mar. 2, 2009, http://accounting.smartpros.com/x65641.xml.

315. Press Release, Am. Fed'n of State, County and Mun. Employees, Say on Pay Shareholder Proposals Garner Record Support During Tumultuous Shareholder Season (May 4, 2009), *available at* http://www.afscme.org/press/26145.cfm.

316. *See Obama Imposes Limits on Executive Pay*, MSNBC, Feb. 4, 2009, http://www.msnbc.msn.com/id/29003620 (noting that future recipients of TARP funds will be required to hold a nonbinding shareholder vote in order to pay executives more than $500,000).

317. DealBook, http://dealbook.blogs.nytimes.com/2009/05/19/schumer-seeks-shareholder-vote-on-executive-pay (May 19, 2009, 14:27 EDT).

318. *See generally* Bebchuk et al., *supra* note 292, at 2–4 (arguing that the influence executives have over boards exerts substantial pressure on compensation decisions, which in cases of great influence leads to compensation that is constrained only by fear of public outrage).

important for CEOs, and CEOs can be expected to use the levers of power they have to notch up the amount they earn. Thus, if we are right and CEOs have lost power, we may expect that the decline in power has, or will soon have, an adverse impact on their compensation.

Executive compensation, of course, may also respond to macroeconomic factors. As such, one should be careful not to over interpret short-term changes in executive compensation. Moreover, other things being equal, a loss of power would make the CEO job less attractive. Thus, to the extent that CEO compensation is determined by supply-and-demand forces, a loss of CEO power could result in higher monetary compensation.

Table 8: Executive Compensation

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Median Total Compensation	2.1	2.5	2.6	3	3.8	4.3	4.9	5.4	6.4	5.8	5.8	6.2	6.1	6.9	6.6	5.9
Average Total Compensation	3.1	3.7	4	5.6	7.4	9.5	9.5	12.3	11.1	8.5	7.8	8.6	8.8	9.2	8.8	7.3

All this being said, it appears that the rise in CEO compensation has come to a halt. In Table 8, we present data (in millions of 2007 dollars) of the total amount of executive compensation (salary, bonus, stock options valued at grant time, other incentive compensation, and other compensation) for the CEOs of S&P 500 companies. The table provides both the average and median compensation in that group. The table shows a steep rise in compensation during the 1990s until around 2001. Since then, median compensation has flattened and average compensation has declined. Thus, for the most recent five-year period of 2004 to 2008, median total compensation was $6.3 million and average total compensation was $8.5 million, respectively about the same and 25% below their 2000 levels. Moreover, these data fail to take account of the enhanced disclosure rules,[319] which may have resulted in higher levels of reported compensation from 2006 onwards (and thus make comparisons between pre-2005 compensation and post-2006 compensation more difficult). If one were to assume, for example, that the difference in reported 2005 and 2006 compensation is due entirely to changes in disclosure rules (and that reported 2007 and 2008 compensation under the pre-2005 rules would be lower by a like amount), then even median compensation in the 2004 to 2008 period dropped by around 10% from its 2000 level. Thus, for those who believe that CEO power is positively correlated with compensation, recent compensation trends are consistent with shrinking CEO power.

VIII. The Effects of These Changes on CEO Power

In the previous pages, we have analyzed a large number of changes in the relationships between CEOs, boards, and shareholders. In this Part, we

319. *See supra* notes 288–93 and accompanying text.

analyze how these changes affect CEO power, using the taxonomy developed in Part II.

A. Decision Making: Decisions and Agenda Control

Consider the single most important decision in the life of a company: whether to sell control. In a world of dispersed shareholdings—think back to the 1950s and 1960s—this was a decision in the first instance for the CEO, possibly with the advice of the board of directors.[320] A CEO who determined it was a good time to sell the company or to buy another company would reasonably expect that decision to carry the day, even if the particular form of corporate combination required board and shareholder approval.[321] Likewise, a CEO who decided it was not a good time to sell had reasonable grounds for assuming that the decision would end discussion.

In today's world of activist hedge funds, more independent directors, and assertive shareholders, that is clearly no longer true. How does it play out today? First, the changes in shareholder composition and activism mean that shareholding is far more concentrated, and concentrated in the hands of shareholders—hedge funds and more traditional institutional shareholders—who are more willing to challenge a CEO's decisions than ever before.[322] Such challenges are becoming easier to mount because the decline of staggered boards, the rise of majority voting for directors, and the ever-increasing success of shareholder proposals give shareholders far more opportunities to hold directors accountable for any excessive deference to the CEO.

Moreover, the 1992 partial deregulation of the proxy rules, combined with the end of discretionary broker voting and the adoption of notice and access, means that the costs of challenging the CEO's decision have markedly declined, and the chances of success have increased. The emergence of proxy advisors can further contribute to the success of such challenges. All this takes place against the backdrop of directors who show more substantive independence than ever before: they spend more time monitoring management; exercise greater control over auditing, governance, and compensation decisions; meet regularly in executive sessions; have more control over the information presented to them; and fire the CEO more frequently and more readily.[323]

---

320. *See generally* MACE, *supra* note 3, at 73–85 (discussing the use of the power of control by company presidents and the interplay between presidents and boards in corporate decision making).

321. *See id.* at 186 ("In most companies the allocation of capital resources, including the acquisition of other enterprises, is accomplished through a management process of analysis resulting in recommendations to the board and in requests for approval by the board . . . . Approval by boards in most companies is perfunctory, automatic, and routine.").

322. *See* Kahan & Rock, *supra* note 22, at 1029–33 (providing examples of institutional-investor shareholders challenging CEOs).

323. *See id.* at 1029–42 (discussing various cases of activist shareholders exercising control over corporate governance).

A decision by the CEO to sell or not to sell the company—as, say, PeopleSoft CEO Craig Conway learned—is therefore but the beginning of the conversation.[324] And, because all the players know that the rules of the game have changed, some conversations do not even start. In today's environment, a decision by the CEO to sell the company to a favored bidder over a competing bidder offering more would be doomed from the outset.

Equally as dramatic, the CEO has lost significant control over the agenda to the shareholders. The changes summarized above combine to eliminate the CEO's ability to keep matters off the corporate agenda. With hedge funds and more traditional institutional shareholders willing to agitate in favor of proposals on the issuer's proxy under 14a-8 or pursue matters directly in their own proxy solicitations, with the costs of such solicitations declining because of regulation and technology, and with increasing success in passing and implementing such proposals, CEO agenda control has declined. In case after case, shareholders have proved themselves capable of forcing unwanted topics onto the front burner.[325]

These examples of the loss of CEO decision-making powers are the most visible tips of an iceberg. In several other areas, our analysis suggests CEO decision-making power has also declined. These include, in particular, the areas delegated to the responsibility of wholly independent board committees: audit, compensation, and nomination. How much more of the iceberg is hidden under water is harder to tell. Presumably, board members still generally defer to the CEO when it comes to operational decisions (or else decide to fire the CEO). By the same token, we believe that CEOs involve board members more in major strategic decisions and that board members have become more willing to share any concerns over operations with their CEOs outside the boardroom. Anecdotal evidence also suggests that "friendly" hedge funds—who do not engage in adversarial activism—share their views about major business decisions with the CEO. Thus, it is likely that CEO decision-making power has declined notably with respect to some key issues and more moderately over a wider set of issues, with both large shareholders and independent directors gaining power at the expense of CEOs.

B. Second-Guessing

What has changed more than anything else is the ability and incentives for other players to second-guess the CEO's actions. Consider first a CEO who acts imperiously with regard to selling the company.[326] If this decision

---

324. Jim Kerstetter, *Finally, Oracle Nails PeopleSoft*, BUS. WK., Dec. 13, 2004, *available at* http://www.businessweek.com/technology/content/dec2004/tc20041213_8884_tc024.htm.

325. *See* Kahan & Rock, *supra* note 22, at 1029–46 (examining how a variety of shareholders, including hedge funds and mutual funds, leverage their voting powers against corporate boards).

326. CEOs who consume excessive perks may also face criticism. Personal use of corporate jets must be disclosed under 17 C.F.R. § 229.402(c)(2)(ix)(A) (2009). Under these disclosure rules,

was once considered the final word, it no longer is. In today's environment, one would expect hedge funds to buy shares in order to challenge the decision. Thus, when Yahoo's CEO Jerry Yang cold-shouldered an offer by Microsoft to acquire the company, it did not take long for Carl Icahn to commence a proxy contest and place three nominees on Yahoo's board.[327] More generally, the evidence we presented—the emergence of hedge funds, the greater power of institutional investors and their greater proclivity to activism, the regulatory changes making it easier for shareholders to challenge managements, the increased monitoring of management by outside directors, and last but not least the reduced tenure of CEOs—suggests that if a CEO makes mistakes (or perhaps just has bad luck), both shareholders and directors will voice their criticism sooner and more strongly than in the days of yore, be it informally, through a proxy challenge or other activist campaign, or through a board-induced CEO resignation.

C. The Scope of CEO Power: Extension, Comprehensiveness, and Intensity

As noted earlier, CEO power can also be divided along the dimensions of extension, comprehensiveness, and intensity. Extension essentially relates to the scale of the firm. Since firms have not gotten smaller, and since CEOs remain on top of the firm, CEO power has not declined along that dimension. In terms of comprehensiveness—the number of topics over which power is exercised—and intensity—the degree to which the holder of power can impose his or her will—CEO power has declined. As noted before, the decline is most pronounced (that is, sharpest along the dimension of intensity) in areas that require board or shareholder approval, such as decisions to sell the firm, audit matters, compensation, corporate governance, and board nominations. In other areas, we believe CEO power has declined as well, but due to the lack of transparency over how these decisions are made and whether they are second-guessed, it is harder to document the decline. Moreover, since independent directors and even activist shareholders have limited capacity to micromanage a company, it is likely that CEOs still have substantial decision-making power over most nonstrategic business matters, as long as their decisions produce acceptable results.

---

the total value of perks must be disclosed "unless the aggregate amount of such compensation is less than $10,000." *Id.* A further requirement is that the company must identify by type each individual perk and must quantify and disclose each perk that exceeds the greater of $25,000 or 10% of the total amount of individual perks. *Id.* § 229.402(c)(2)(ix), instruction 4.

327. *See* Aaron Smith, *Yahoo Puts Icahn on Board, Settling Spat*, FORTUNE, July 21, 2008, *available at* http://money.cnn.com/2008/07/21/news/companies/yahoo_icahn/index.htm (noting that Yahoo agreed to place Icahn on its board and allowed him to appoint two additional board members while at the same time thwarting his efforts to take over control of the board).

IX. Implications

Our thesis that CEO power has declined notably over the last several years has important implications for corporate law and corporate governance. In this Part, we discuss these implications. First, we argue that the changes we describe reflect a long-term trend that is likely to continue and intensify, rather than some short-term cyclical movement that will reverse itself. Second, because of the nature of these changes and their underlying causes, we do not think that the loss of CEO power will generate a political backlash. Third, the new role of the board will lead to the appointment of board members having different backgrounds and competencies than before. Fourth, we predict increased velocity in the types of shareholder initiatives introduced via Rule 14a-8. Fifth, we discuss the implications of our thesis for the debate over the extent to which the corporate law of various countries is converging. Sixth, we analyze the implications of the loss of CEO power on Delaware law and the state-competition debate. Seventh, we argue that the loss of CEO power may reduce CEOs' resistance to having their company acquired. Eighth, it weakens the case for new corporate law rules that grant shareholders greater voting rights. And ninth, we examine the relationship between CEO power and private equity. Finally, in the conclusion, we comment on whether the loss of CEO power is a positive or a negative development.

A. Fundamental Shift or Perfect Storm?

Some observers, noting some of the issues we discussed in this Article, have characterized the current state of affairs as a "perfect storm."[328] The perfect-storm metaphor evokes a temporary and accidental alignment of forces that creates a special situation or opportunity.[329] But like other storms, perfect ones ultimately pass and the situation returns to normal.

We do not think this captures what is happening. The changes we discuss are not temporary and their simultaneous occurrence is not accidental. Any changes in the regulatory environment—including the changes in proxy rules, the revised listing standards in the stock exchange rules, or the elimination of broker voting in uncontested director elections—are likely to persist. The shift in equity ownership from individuals to institutions reflects fundamental long-term change forces[330] that will continue. Companies that have agreed to destagger the board are unlikely to receive shareholder approval to reintroduce a staggered board.[331] And while most

328. ALLEN, *supra* note 148, at iv; IR MAG. GUIDE, A PERFECT PROXY STORM 2 (2007), http://www.altmangroup.com/pubs/IRMag/IR Magazine - US Guide March (2).pdf.

329. *See, e.g.*, D. Michael Fields, *Perfect Storm*, BIZED MAG., Jan./Feb. 2006, at 34, http://www.aacsb.edu/publications/Archives/JanFeb06/p34-37.pdf ("A perfect storm, by definition, is a convergence of independent events . . . .").

330. Specifically, the way retirement benefits are financed.

331. Marcel Kahan & Edward Rock, *Corporate Constitutionalism: Anti-takeover Charter Provisions as Precommitment*, 152 U. PA. L. REV. 473, 495–96 (2003).

companies that have adopted majority voting could return to plurality voting without shareholder approval,[332] we think this is both unlikely and ultimately ineffective: even under a plurality-vote regime, a director who receives a majority of withhold votes faces enormous pressures to resign.[333]

These changes, in turn, have caused some of the other changes we observe. To be successful, activist hedge funds need allies, and institutional investors with their increased holdings are likely candidates. Successful hedge fund activism has led traditional institutions first to lend their active support to hedge funds, and then to lead the charge themselves. The rise in institutional holdings has generated demand for voting advice by proxy advisors. The destaggering of boards and majority voting has increased the meaningfulness and the frequency of director elections. That directors are up for election more frequently, that they are worried about a large withhold vote, and that proxy advisors are more likely to recommend a withhold vote if the board ignored a shareholder resolution are all at least part of the reason why boards have become more responsive to shareholder resolutions. This, in turn, means that more companies will destagger, adopt majority voting, or even give shareholders a say on pay. Independent nominating and governance committees reduce the ability of CEOs to stop this. Increased holdings by institutions and fear of hedge funds increase both the demand by shareholders to meet with outside directors and the willingness of directors to do so. Directors meeting in executive session create the opportunity to discuss company developments unmonitored by the CEOs. As more boards question their CEOs, it becomes more acceptable for directors in other companies to do so. Greater director independence and greater pressure from shareholders, in turn, increase CEO turnover. Increased turnover means that, at any point in time, there will be more members of the board who have picked the CEO and fewer who were picked during the CEO's tenure. Thus, even if CEOs continue to influence the selection of board members, despite the requirement of wholly independent nominating committees, shorter tenure implies less CEO influence over board membership. And we could go on.

We are not so bold to claim that all the trends we described will continue unabated. But we think that it is much more likely that CEOs, in the intermediate term (over the next ten years or so), will lose more power than that they will regain some of the power they have lost.

---

332. Companies require shareholder approval only if majority voting is embedded in the charter, a majority voting bylaw is adopted by shareholders, or a board-adopted bylaw provides that it can be amended only by shareholders. DEL. CODE ANN. tit. 8, §§ 216, 242 (2001 & Supp. 2008).

333. *See supra* note 152.

B. Backlash

If we are correct and the changes we discuss presage a continuing decline in the power of CEOs, rather than a cyclical and self-reversing shift, there is the possibility of a political or regulatory backlash. Such a backlash, in the form of state anti-takeover statutes and Delaware's sanctioning of the poison pill, helped stop the hostile takeover wave of the 1980s, the last significant threat to managerial power.[334] These days, advocates of managerialism already argue that the increased power of shareholders and decreased board collegiality induce an excessive short-term orientation that harms U.S. competitiveness.[335]

While the possibility of backlash cannot be excluded, we believe that its likelihood is remote. Unlike in the 1980s, the threat to managers derives from multiple sources—traditional institutions, hedge funds, proxy advisors, technology, and their fellow directors—rather than from a small group of raiders. And compared to raiders of the 1980s, who were in many respects outsiders,[336] even hedge funds (and, a fortiori, institutional investors and board members) are part of (or well connected to) the establishment and have significant political power.[337] The threat to managers is more gradual and broad based than in the 1980s and thus less likely to result in a strong response. Finally, there is little reason to expect populist support for pro-management changes; organized labor, who supported anti-takeover legislation in the 1980s, is lined up against management in this round;[338] and, for the moment at least, populist anger is directed against highly compensated CEOs, rather than at shareholder activists.

---

334. *See* Kenneth W. Hollman, *Merger Mania: Human and Economic Effects*, REV. BUS., June 22, 1991, *available at* http://www.allbusiness.com/buying-exiting-businesses/mergers-acquisitions/268232-1.html ("Over 30 states (including Delaware) passed laws in the latter half of the 1980s to thwart the takeover effects of corporate raiders.").

335. *Has Shareholder Influence Gone Too Far? Or Not Far Enough?*, Posting of Heidi N. Moore to Deal Journal, http://blogs.wsj.com/deals/2008/05/23/has-shareholder-influence-gone-too-far-or-not-far-enough/ (May 23, 2008, 12:06 EST).

336. For example, many prominent investment banks and law firms refused to work for hostile bidders. *See, e.g.*, RON CHERNOW, THE HOUSE OF MORGAN 707 (2001) (acknowledging that until the late 1980s, J.P. Morgan did not do work for hostile bidders).

337. During the 2008 election cycle, hedge-fund-associated individuals and PACs made over $16 million in political contributions to federal candidates and parties. *See* Hedge Funds: Long-Term Contribution Trends, http://www.opensecrets.org/industries/indus.php?cycle=2010&ind=f2700 (showing amount of political contributions made by hedge funds). For six funds, the contributions exceeded $500,000. *See* Hedge Funds: Top Contributors to Federal Candidates and Parties, http://www.opensecrets.org/industries/contrib.php?ind=f2700&cycle=2008 (showing top hedge fund political contributors and detailing their contributions). Presidential candidate John Edwards worked with hedge fund Fortress Investment. Emily Thornton, *John Edwards Hits the Streets*, BUS. WK., Oct. 13, 2005, http://www.businessweek.com/bwdaily/dnflash/oct2005/nf20051013_3314_db016.htm. Chelsea Clinton has also worked with hedge funds. *Chelsea Clinton Joins New York Hedge Fund*, MSNBC, Nov. 3, 2006, http://www.msnbc.msn.com/id/15549672/.

338. For example, union-affiliated pension funds sponsor some of the anti-management resolutions discussed above. *See supra* notes 295–99 and accompanying text.

C. *Board Composition*

The shift of power from CEOs to outside board members also has implications for the type of persons who will serve on corporate boards. Compared to outside directors fifteen years ago, outside directors today are likely to have more power, to enjoy a less collegial relationship to the insiders, to have a greater workload, to earn greater pay, to have occasional need to become confrontational, and to deal more often with vocal and restive shareholders.[339] Accordingly, board composition will shift to persons who are good at these new tasks, who derive greater enjoyment from them, and who have the needed time and energy to devote to the job.

One category of persons who may be particularly well qualified for board service in the current environment are retired CEOs and other retired high-level executives, bankers, accountants, consultants, or investment professionals. They tend to have the time, the background, the independence, and the interest to perform the tasks set to them. We would predict that, over time, the percentage of board members from these categories will increase.

D. *Shareholder "Flavor of the Year" Initiatives*

Shareholder resolutions often come in waves, with every year or so witnessing the emergence of a new "flavor of the year" type of precatory resolution and the decline of some prior types. The last two proxy seasons (2007 and 2008), for example, saw the rise of proposals asking the board to grant shareholders the right to call a special meeting.[340] These proposals, virtually unheard of until 2006,[341] were proposed in twenty-three companies in 2008, were on average supported by 47% of the votes cast, and passed in eleven of the companies.[342] By contrast, proposals to redeem or get a shareholder vote on poison pills went from fifty in 2004 to three in 2008.[343]

In the past, management's response has largely been to duck and cover: to hope for the storm to pass before the topic gained sufficient traction to generate real pressure for change.[344] This tactic looks increasingly untenable.

---

339. MICHAEL J. STAHL & DAVID W. GRIGSBY, STRATEGIC MANAGEMENT: TOTAL QUALITY AND GLOBAL COMPETITION 8–9 (1997).

340. 2007 ANNUAL CORPORATE GOVERNANCE REVIEW, *supra* note 204, at 33; 2008 ANNUAL CORPORATE GOVERNANCE REVIEW, *supra* note 304, at 33–34.

341. GEORGESON, ANNUAL CORPORATE GOVERNANCE REVIEW 32 (2005), http://www.georgesonshareholder.com/emea/resources_research.php [hereinafter 2005 ANNUAL CORPORATE GOVERNANCE REVIEW]; GEORGESON, ANNUAL CORPORATE GOVERNANCE REVIEW 31 (2006), http://www.georgesonshareholder.com/emea/resources_research.php [hereinafter 2006 ANNUAL CORPORATE GOVERNANCE REVIEW].

342. 2008 ANNUAL CORPORATE GOVERNANCE REVIEW, *supra* note 304, at 33–34.

343. 2004 ANNUAL CORPORATE GOVERNANCE REVIEW, *supra* note 304, at 21–22; 2008 ANNUAL CORPORATE GOVERNANCE REVIEW, *supra* note 304, at 32–33.

344. *See* Neil O'Hara, *Prepare for Attack: What to Do When Hedge Funds Move In*, COMPLIANCE WK., Mar. 28, 2006, http://www.complianceweek.com/article/2404/defensive-moves-

First, with the rise of institutional investors, it takes less time for a new proposal to gain significant shareholder support.[345] Second, once a proposal has received (or is expected to receive) support, boards are increasingly willing to adopt the recommendation. Thus, in 2007, eleven of the seventeen shareholder right-to-call-a-special-meeting proposals passed, and eight were implemented by the following year.[346] In 2008, eleven passed, and five were implemented.[347] As for new types of shareholder initiatives, we generally predict proposals to gain more traction than in the past and to do so more quickly, which in turn will lead to more and more types of initiatives.

E. Convergence

There is a long running debate among corporate scholars over whether the corporate law and governance systems in different countries are converging.[348] Our evidence suggests that we may be witnessing the end of a particular exceptionalism in U.S. corporate governance: the imperial CEO. In many respects, the changes we discussed in this Article, while new from the U.S. perspective, have long been part of the corporate-governance regime in other Anglo-American countries, such as the U.K., Canada, and Australia. Thus, for example, most U.K. companies have a non-executive chairman of the board, and U.K. law gives shareholders a nonbinding say on pay.[349] Poison pills are not permitted under Australian law.[350] Under Canadian law, directors of a company with a classified board can be removed without cause,[351] making this device an ineffective takeover defense. As U.S. practice moves closer to the practice of these other countries, both with

---

when-hedge-funds-attack (noting that before the rise in activist shareholders companies used to be able to simply ignore the investors); *see also* Kirsten Grind, *WaMu Likely to Ignore Ire of Shareholders at Meeting*, PUGET SOUND BUS. J., Apr. 11, 2008, http://seattle.bizjournals.com/seattle/stories/2008/04/14/story2.html ("Boards usually don't listen to the messages sent by shareholders when they withhold votes . . . .").

345. *See supra* notes 156–68 and accompanying text.

346. 2007 ANNUAL CORPORATE GOVERNANCE REVIEW, *supra* note 204, at 33.

347. 2008 ANNUAL CORPORATE GOVERNANCE REVIEW, *supra* note 304, at 33–34.

348. *See, e.g.*, CONVERGENCE AND PERSISTENCE IN CORPORATE GOVERNANCE 4–5 (Jeffrey N. Gordon & Mark J. Roe eds., 2004) (chronicling arguments about convergence and ways in which convergence could prevail); John C. Coffee, Jr., *The Future as History: The Prospects for Global Convergence in Corporate Governance and Its Implications*, 92 NW. U. L. REV. 641, 645–48 (1999) (describing several alternative positions that have emerged in the convergence debate); Jennifer G. Hill, *The Persistent Debate About Convergence in Comparative Corporate Governance*, 27 SYDNEY L. REV. 743, 743–44 (2005) (book review) (analyzing the "convergence–divergence" corporate governance debate over the last few decades).

349. Donald Kalfen et al., *The Future of Say on Pay: Current Status and Possible Impact*, BOARDMEMBER, http://www.boardmember.com/Article_Details.aspx?id=2076.

350. Jennifer G. Hill, *The Shifting Balance of Power Between Shareholders and the Board: New Corp's Exodus to Delaware and Other Antipodean Tales* 33 (Vanderbilt Univ. Law Sch., Law & Econ. Research Paper No. 08-06, 2008), *available at* http://ssrn.com/abstract=1086477.

351. Canada Business Corporations Act, R.S.C. § 109 (1985).

regard to specific issues and with regard to the overall power of the CEO, the corporate law regimes are converging.

F. Delaware Law

There is long standing debate in corporate law as to the tilt of Delaware law. Delaware, of course, is the jurisdiction in which most publicly traded companies are incorporated and is thus generally acknowledged as the most "attractive" corporate law jurisdiction.[352] But attractive to whom? According to "race to the bottom" commentators, the incorporation decision is largely made by managers, and Delaware succeeds in attracting corporations because it has pro-management rules.[353] According to "race to the top" commentators, the incorporation decision is driven by market forces, and Delaware succeeds in attracting corporations because it has rules that maximize the value of the corporation.[354] According to a third set, both management and shareholders have power over the incorporation decision, and Delaware succeeds because it generally has both a better corporate law than most other states and, in areas where shareholder and manager interests conflict, adopts rules that are acceptable to both sides.[355]

Some of the developments we have described pose substantial challenges for race to the bottom and race to the top commentators. Race to the bottom commentators, who believe that managerial power is predominant, may have a hard time explaining why so many companies have destaggered their boards, adopted majority voting, and implemented precatory shareholder resolutions—all developments that, on their face, reflect shareholders exercising power over governance decisions at the expense of the board. Race to the top commentators would have a somewhat easier time explaining the same developments but would still need to explain why it was optimal for a majority of large companies to have staggered boards and plu-

352. *See* Robert Daines, *The Incorporation Choices of IPO Firms*, 77 N.Y.U. L. REV. 1559, 1563 (2002) ("Delaware has a nearly 70% share of IPO firms and 95% share of firms incorporating outside their home state. Delaware's share is growing over time.").

353. *E.g.*, Lucian A. Bebchuk, *Federalism and the Corporation: The Desirable Limits on State Competition in Corporate Law*, 105 HARV. L. REV. 1435, 1438 (1992); *see also* William L. Cary, *Federalism and Corporate Law: Reflections upon Delaware*, 83 YALE L.J. 663, 666 (1974) (arguing for federal rules as a solution to the "race for the bottom" situation created by Delaware and other states' management "enabling" acts).

354. *See, e.g.*, Roberta Romano, *Empowering Investors: A Market Approach to Securities Regulation*, 107 YALE L.J. 2359, 2383 (1998) (asserting that investors benefit from competition and changes in corporate domicile to states such as Delaware); Ralph K. Winter, Jr., *State Law, Shareholder Protection, and the Theory of the Corporation*, 6 J. LEGAL STUD. 251, 257–58 (1977) (arguing that the fact that other states have had to change their laws in response to Delaware indicates that investors do not believe the race to the bottom theory and instead believe that they do better under Delaware law).

355. *E.g.*, Michael Barzuza, *Price Considerations in the Market for Corporate Law*, 26 CARDOZO L. REV. 126, 135–36 (2004); Marcel Kahan & Ehud Kamar, *The Myth of State Competition in Corporate Law*, 55 STAN. L. REV. 679, 739–40 (2002).

rality voting several years ago, and yet it is now optimal for most of these companies to destagger their boards and adopt majority voting.

As to the third set of commentators, the loss of CEO power suggests that the optimal compromise between shareholder and manager interests that Delaware strives to adopt has moved towards the shareholder side. As shareholder power over incorporation decisions increases and management power shrinks, the Delaware law that, at the margin, appeals to the greatest set of relevant decision makers has become more shareholder friendly.

Indeed, several recent changes in Delaware law are consistent with this prediction. Thus, in 2006, Delaware adopted legislation that a shareholder-adopted bylaw mandating majority voting for directors cannot (unlike most other bylaws) be repealed by the board of directors.[356] And in 2009, Delaware adopted legislation expressly permitting the adoption of bylaws to provide for proxy access or to require reimbursement for expenses incurred in soliciting proxies for director elections in opposition to the board's nominees.[357]

G. Changing Dynamics of Resistance to Acquisition

Being CEO of a public company has become less fun. You get to call fewer shots, you are being second-guessed by boards and shareholders, your compensation has plateaued, and your job security has decreased. All of this will make CEOs more willing to let their company be acquired and cash in on appreciated stock options or severance payments. CEO resistance to acquisitions should thus decline.

On the other hand, boards and shareholders may now be the ones to offer roadblocks. Boards may get more involved in the negotiations of acquisition terms and may reject offers that the CEO may want to accept. And we have already witnessed instances of shareholders trying to renegotiate a deal struck by management.[358] Board and shareholder resistance to acquisitions—which, until recently, was negligible—has thus increased.

We are working on documenting the divergence in financial incentives to engage in control transactions between the CEO and independent directors. While today's CEOs have strong monetary incentives to support a

---

356. DEL. CODE ANN. tit. 8, § 216 (2009).

357. *Id.* § 112. There have also been some recent judicial decisions favoring shareholders. In *CA, Inc. v. AFSCME Employees Pension Plan*, 953 A.2d 227 (Del. 2008), the court held that governance rules—specifically, rules on reimbursement of proxy expenses—are proper subjects of bylaws and need not be included in the corporate charter, as long as they do not require violations of Delaware law (including fiduciary duty law). *Id.* at 240. And in *Gantler v. Stephens*, 965 A.2d 695 (Del. 2009), the court made it more difficult for boards to obtain shareholder ratification of breaches of fiduciary duties by holding that "the scope of the shareholder ratification doctrine must be limited to . . . circumstances where a fully informed shareholder vote approves director action that does *not* legally require shareholder approval in order to become legally effective." *Id.* at 713.

358. *See supra* notes 320–27 and accompanying text.

change in control, especially as they approach the end of their tenure,[359] for outside directors such control changes are a losing proposition. While outside directors are able to sell any shares they have for a premium, these gains are dwarfed by the loss of the very substantial director fees.[360] Moreover, such directors are typically not able to replace their lost directorship with another of comparable status.

These changes, in turn, may impact Delaware law on acquisitions. First, Delaware law on hostile takeovers and, in particular, the "just say no" defense, will become less important. By the same token, Delaware law on the ability of boards to "lock up" deals in the absence of a bidding contest could become more important.[361] More profoundly, Delaware law rests to some extent on the premise that shareholders want to sell the company at a premium but that management may want to block the sale and stay independent.[362] To the extent that this premise is no longer correct, Delaware law will have to adapt the substantive standard by which it evaluates transactions.

H. The Need for Greater Shareholder Voting Rights

In a series of articles published in 2005 and 2006, Lucian Bebchuk argued that shareholder voting rights should be expanded to include the power, without board approval, to change the company's governance structure (including the power to change the charter and to reincorporate into a different state) and to make certain specific business decisions (such as the power to instruct the board to auction the company to the highest bidder).[363] The premise of Bebchuk's argument is that, even though shareholders elect the board of directors (and thus indirectly already control all of these decisions), directors do not heed shareholder wishes. Predictably, other

---

359. *See* John C. Coates, IV & Reinier Kraakman, *CEO Tenure, Performance and Turnover in S&P 500 Companies* 5 (European Corporate Governance Inst., ECGI-Finance Working Paper No. 191/2007, 2010), *available at* http://ssrn.com/abstract=925532 ("CEOs on the cusp of retirement or discharge might opt to sell their companies instead, in order to trigger option plans and liquidate equity holdings.").

360. Charles M. Elson, *Corporate Law Symposium: The Duty of Care, Compensation, and Stock Ownership*, 63 U. CIN. L. REV. 649, 694 (1995).

361. In practice, Delaware law addresses lock-ups only in the context of competing bids. *See* Hastings-Murtagh v. Tex. Air Corp., 649 F. Supp. 479, 484 (S.D. Fla. 1986) (determining that Delaware law allows lock-up provisions "where there is a live auction with competing bidders").

362. *See* Unocal Corp. v. Mesa Petroleum Co., 493 A.2d 946, 955 (Del. 1985) (setting the standard used by courts to determine whether a board's defensive measures to a perceived threat of hostile takeover, which interfere with the exercise of shareholder voting, are reasonable).

363. *See* Lucian A. Bebchuk, *Letting Shareholders Set the Rules*, 119 HARV. L. REV. 1784, 1787–95 (2006) (responding to criticism of his recommendation to give shareholders control over "rules-of-the-game" decisions); Lucian A. Bebchuk, *The Case for Increasing Shareholder Power*, 118 HARV. L. REV. 833, 850–75 (2005) (arguing for giving shareholders the power to initiate and adopt "rules-of-the-game" decisions such as amending corporate charters and reincorporation in another jurisdiction).

commentators have ridden to the defense of the current system where the board retains greater control.[364]

The evidence we present in this Article suggests that, whatever the merits of Bebchuk's proposal may have been when it was conceived, the need for (and desirability of) any reform suggested by Bebchuk has declined. Bebchuk and his detractors fundamentally differ with respect to one major issue: when shareholders and the board of directors disagree—e.g., over whether the company should be auctioned off—who is more likely to be right? Both sides to the debate, however, would presumably agree that boards are more likely to heed shareholder wishes if they believe that what shareholders want is good for the company. That is, the merits of what shareholders want and the likelihood of boards following a nonbinding shareholder vote are correlated.

In the ideal corporate-governance world, boards would retain just that modicum of power that permits them to block, at the margin, more bad ideas than good ideas. In the real world, of course, board power cannot be fine-tuned in that manner. As a formal matter, a board can either block certain types of decisions or it cannot. Bebchuk, in effect, argues that we would be better off if boards could not block governance changes and certain business decisions. His detractors argue that we are better off if they can.

As we have shown in this Article, however, even though the formal powers of the board have not changed, boards have become much more receptive to shareholders. Thus, boards voluntarily, albeit selectively, implement more shareholder-proposed governance changes.[365] This obviously reduces the need for removing board veto power over governance changes, as advocated by Bebchuk.

But if, as is likely, from among all the governance changes desired by a majority of shareholders, the governance changes implemented by boards are better than those rejected by boards, it may also mean that the time for Bebchuk's proposal has passed. Even if we would be better off with letting shareholders set the rules than with giving the veto right to boards when boards regularly ignored what shareholders want, we may be better off with board veto when boards implement a significant portion of nonbinding proposals passed by shareholders.

---

364. Stephen M. Bainbridge, *Director Primacy and Shareholder Disempowerment*, 119 HARV. L. REV. 1735, 1736–44 (2006) (arguing that the current regime offers substantial efficiency benefits); Leo E. Strine, Jr., *Toward a True Corporate Republic: A Traditionalist Response to Bebchuk's Solution for Improving Corporate America*, 119 HARV. L. REV. 1759, 1769–75 (2006) (arguing that the capital markets have not indicated a need for substantial change in corporate governance).

365. *See supra* notes 158–72 and accompanying text.

I. Private Equity to the Rescue? The Trade-Off Between Power and Wealth

As remarkable as the growth in hedge funds may have been, it is not unparalleled. The funds raised by private equity firms in the U.S. have experienced a growth rate—23%, annually compounded, from 1999 to 2006—as high as the hedge fund assets under management (19% during that period).[366] The total dollar volume of private equity M&A transactions in 2006 was $900 billion, a magnitude comparable to the total hedge fund assets ($1.427 trillion), especially considering that a significant portion of hedge fund assets are not invested in equity securities.[367]

Until the recent credit crunch, private equity funds played an increasingly large role in M&A.[368] As a percentage of total M&A dollar volume, private equity M&A had grown from less than 5% in 1999 to more than 25% in 2006.[369] In 2006, there were 151 going-private transactions sponsored by private equity funds, up from only sixty-seven in 2000.[370]

Private equity funds, like hedge funds, are significant new players. But unlike hedge funds, private equity is considered management friendly.[371] Private equity funds rarely if ever engage in hostile transactions.[372] Instead, they offer CEOs a safe harbor in a storm, by expanding CEOs' options and opportunities, and, when taking companies private, by offering the possibility of great wealth.

However, even if CEOs, on the whole, view private equity funds favorably, these funds contribute to the decline in CEO power. As an institution, private equity weakens CEOs by increasing the likelihood of a change of control, closely monitoring their investments in public companies, and tightly controlling portfolio companies (setting and monitoring goals, and firing underperforming CEOs). While having one's company acquired by a private equity firm may make a CEO rich, his power is reduced. When a company is acquired, the CEO either leaves the firm or stays on to manage

---

366. The Blackstone Group L.P., Amendment Number 9 to Form S-1, at 148, 151 (June 21, 2007).

367. *Id.* at 149, 151.

368. *See* Chris Snow, *Impact of Credit Crisis on Private Equity Markets*, 28 REV. BANKING & FIN. L. 71, 79 (2009) (emphasizing the future opportunities in the private equity industry due to significant uncommitted capital from the recent "boom" despite the fact that "[b]ecause of recent financial instability and frozen credit markets, traditional private equity buyout activity has essentially stopped").

369. The Blackstone Group, *supra* note 366, at 149.

370. *Id.*

371. *See* Brian Cheffins & John Armour, *The Eclipse of Private Equity*, 33 DEL. J. CORP. L. 1, 13 (2008) (contending that when a private equity fund buys a company "management can become very rich" and avoid the "adverse publicity associated with generous executive pay in public companies").

372. *See id.* at 12 (explaining that when private equity firms carry out buyouts they usually "opt to negotiate a 'friendly' deal with the target").

the firm, which is now a portfolio company in a private equity fund. In the former case, the CEO gives up any power that comes with the job. In the latter case, the CEO now has a boss—the management of the private equity firm—that has the ability and the incentives to monitor him and to fire him if they are dissatisfied. Whatever financial rewards the CEO may obtain in his new position, one thing is clear: the power of a CEO of a company owned by a private equity fund is much less than the power of a CEO of a comparable company that is publicly traded.

X. Conclusion: Searching for the Sweet Spot

The story we tell above is a story of declining CEO power over the last several years, a decline that has occurred across almost all of the relevant dimensions and that we believe will last and continue. Is this a good thing?

One of the great virtues of the corporate form is centralized management.[373] Much of corporate law can be interpreted as establishing and protecting that centralized management because of the benefits that it provides to the participants in the firm. At the same time, the centralization of management in the hands of paid managers creates agency costs for the shareholder–manager, the prevention of which forms such an important part of corporate law.

There is, for a given firm operating in a specific environment, a point at which the net benefits of delegation are maximized. The difficulty is that it is very tough to know whether we are at that point.

In this Article, we argue that the balance of power between CEOs, boards, and shareholders has shifted notably in the last decade away from CEOs towards outside directors and shareholders. If, as we expect, that shift will continue in the same direction, CEOs are left ever more embattled, at least in comparison to their predecessors a generation ago. But we cannot claim, and do not know, whether the balance has shifted too far, or whether under current conditions the CEO is not powerful enough. On the other hand, those arguing to strip the CEO of even more power also cannot show that the CEO of today is too powerful.

As we search for the sweet spot, it is worth keeping in mind that for every story about a domineering CEO who should have been replaced long ago, there is an Andrew Grove or a Jack Welch who used the power of the position to make billions of dollars for their shareholders.

373. *See* Edward B. Rock & Michael L. Wachter, *Islands of Conscious Power: Law, Norms, and the Self-governing Corporation*, 149 U. PA. L. REV. 1619, 1698 (2001) (anchoring centralized management within the theory of the firm).

* * *

Book Review

Metaphysical Truth vs. Workable Tort Law: Adverse Ambitions?

CAUSATION AND RESPONSIBILITY: AN ESSAY IN LAW, MORALS, AND METAPHYSICS. By Michael S. Moore. Oxford, United Kingdom: Oxford University Press, 2009. 605 pages. $130.00.

David W. Robertson*

In helping law students learn to write legal briefs, law review articles, and seminar papers addressing legal issues, I start with the same basic message: "Your overarching aim must be producing a reader-friendly product. Your reader will strive to understand and evaluate your work with her own goals in mind. You need to give careful, imaginative, and—above all—empathetic thought to what those goals are likely to be and to try your best to help your reader meet at least some of them. Do not make your reader work harder than can be avoided.[1] And please remember at all times, nobody wants to see a picture of the inside of your head."

Professor Michael Moore would likely regard my advice as hogwash. His treatise, *Causation and Responsibility*,[2] is an abstruse, ambitious, and intermittently quite difficult presentation of Moore's intuitions about the nature of causation and the implications of those intuitions for tort law, criminal law, and moral philosophy. For me, the arduous journey into the admirable mind of Michael Moore has been intellectually beneficial and for the most part enjoyable. But my preferred academic and professional focus is the day-to-day operation of law, particularly tort law, in the actual realms of legislation and litigation; my priority for choosing work-time reading is

* W. Page Keeton Chair in Tort Law and University Distinguished Teaching Professor, University of Texas. I am grateful to David Anderson, Mitch Berman, and Dan Dobbs for insightful comments.

1. In the late 1970s, when I was about 40 years old and my friend and mentor Leon Green was approaching 90, Green asked me—actually he ordered me (he regularly did this)—to edit a draft of something he was working on. Normally he rejected virtually all of my suggestions, and this time was no exception. When I said that one section of the piece seemed to lead Green's readers into unnecessarily strenuous effort (and proposed a way to fix it) Green banged his fist on his desk and said, "Let the readers do some work! It's good for them." Green's pronouncements were almost always full of wisdom, but remember that he was about 90 when he made this one. For a brief introduction to Green, see David W. Robertson, *Green, Abner Leon*, *in* THE YALE BIOGRAPHICAL DICTIONARY OF AMERICAN LAW 231–32 (Roger K. Newman ed., 2009).

2. MICHAEL S. MOORE, CAUSATION AND RESPONSIBILITY: AN ESSAY IN LAW, MORALS, AND METAPHYSICS (2009).

zeal to improve my grasp of and facility with teaching, implementing, manipulating, and critiquing the law in actual operation; and so my bottom-line criterion for evaluating a book that professes to address tort law[3] is simply: "Does this help?" This Review presents my reasons for answering no.[4]

I. Lexicon

Moore evidently hopes for tort-lawyer readers, but the core of his intended audience seems to be a smallish group of specialists in moral philosophy.[5] Often speaking primarily to those specialists,[6] Moore uses terms of recondite or uncertain meaning without defining them. When I initially engaged with the book, I soon realized that I could not analyze or evaluate it until I had settled on some probable definitions. Dictionaries were not very helpful. There were a number of these daunting terms,[7] but only two are immediately relevant: (1) When used as an adjective modifying morality or ethics, *deontological* can merely mean binding, but I think for Moore it means something like absolutist up to a point;[8] (2) *metaphysical truth* means *the* truth—God's own truth.

II. Moore's Concept of Causation

The book begins by setting forth four postulates[9] that the reader is asked to accept on faith[10]:

(a) "[C]ausation [is] a natural relation lying at the heart of scientific explanation."[11] What this means is that causation is not merely a mental or

---

3. *See, e.g.*, *id.* at xiii ("This book is mostly written to correct . . . errors of legal theory."). Indeed, the book's subtitle states that this is a book on causation in the law.

4. Sometimes works of philosophy do help. See, for example, the excellent essay on corrective justice, NEIL MACCORMICK, *The Obligation of Reparation*, *in* LEGAL RIGHT AND SOCIAL DEMOCRACY: ESSAYS IN LEGAL AND POLITICAL PHILOSOPHY 212 (1982).

5. *See, e.g.*, MOORE, *supra* note 2, at xv–xvi (thanking, in the book's preface, dozens of commentators, including a singled-out core group); *id.* at 50 (referencing works by Greg Kafka and Philippa Foot when discussing the doctrine of double effect); *id.* at 56–57 (discussing a famous example of James Rachels that compares a killing to an omission to save from death, along with the response of Frances Kamm).

6. *See supra* note 5; *see also* MOORE, *supra* note 2, at 514 (acknowledging that "[m]etaphysics is an arcane specialty").

7. For example, I think *aretaic* means pertaining to virtue or excellence. The adjective *nomic* seems to mean according to the laws of nature.

8. Moore attributes to Kant a version of deontology whereby morally forbidden actions can never be taken, not even to save the entire world, and contrasts himself as a "threshold deontologist[, which] means that over some threshold of truly awful consequences, [Moore] will potentially do virtually anything to avert such consequences. If [Moore] can locate and defuse a nuclear device at 42nd Street only by torturing the innocent child of the terrorist who planted it there, [Moore] torture[s]." MOORE, *supra* note 2, at 41.

9. Moore presents these as "two thoughts," but four distinct ideas are entailed. *Id.* at vii.

10. *See id.* ("This book defends [none] of these postulates . . . .").

emotional construct, a habit of speech or thought, or a way of summarizing seemingly inevitably recurrent sequences. Rather it is something like a metaphysical glue that actually exists in God's own true world.[12]

(b) There can rarely be moral responsibility without causation.[13]

(c) There can rarely be legal responsibility without moral responsibility.[14]

(d) Therefore, causation—true causation—is a necessary bedrock of criminal and tort law.

With that kind of build-up, the reader starts looking forward to a detailed exposition of Moore's concept of true causation. Instead comes a surprise: Moore proposes to correct others' conceptions of causation[15] without specifying his own (he says he is not quite ready yet).[16] So it is ultimately up to the reader to figure out what Moore thinks causation means or is. This is a tall order—at times I found myself wishing for an owner's manual—but the following seem to be the basic ideas.

1. Many academic tort lawyers believe that separating factual causation from proximate (or legal) causation has been modern tort law's finest achievement.[17] Moore is a radical dissenter on this point, insisting that factual and proximate cause must be remerged: "[Rather than positing that] there are two distinct causal enquiries, that of cause-in-fact and that of proximate causation . . . [we should realize that there should be only] one

11. *Id.*; *see also id.* at 257 ("The nature of causation—what causation is—is a matter of fact, inviting theoretical speculation. Such nature is not fixed by the conventions that have hitherto governed idiomatic English usage of the word, 'cause.'").

12. *See id.* at 90, 447 (explaining that most scholars, unlike David Hume, believe that causation has "some 'glue' to it").

13. *Id.* at vii. Moore's phrasing of this point says that "moral responsibility (like all moral properties) supervenes on natural properties like causation, intention, and the like." *Id.* Here, as is mostly true throughout the book, morality is controlled by metaphysics. But at times it is the other way around. *See, e.g.*, *id.* at 140 (acknowledging that no "metaphysical impossibility" prevents attributing causation to omissions and stating that "[t]he problem [with doing so] is moral" rather than metaphysical).

14. *Id.* at vii.

15. *See supra* note 3.

16. *See, e.g.*, MOORE, *supra* note 2, at x–xi (summarizing the causal-relation theories later surveyed without "formulating some precise test of causation to be given to legal factfinders"); *id.* at 496 (introducing singularist—as opposed to generalist—theories of causation without "offer[ing] knock-down arguments for the correctness of those theories"); *id.* at 504–06 (distinguishing between the three types of primitive, singularist causation theories and paying particular attention to metaphysical primitivism); *id.* at 512 (defending a reductionist, singularist theory, even though "[t]here is a gap between the evidence we possess . . . and the thing evidenced").

17. *See, e.g.*, RESTATEMENT (THIRD) OF TORTS: LIAB. FOR PHYSICAL & EMOTIONAL HARM §§ 26, 29 (Proposed Final Draft No. 1, 2005) [hereinafter RESTATEMENT (THIRD)] (pertaining to factual and proximate cause, respectively).

enquiry about substantiality of causal contribution."[18] (Right away Moore and I are off on the wrong foot.[19])

2. Causation moves forward in time; it never backs up. Something that came into existence at noon could not have been a cause of an event at 11 a.m.[20]

3. Moore posits that *A* does not cause *B* unless there is some kind of metaphysical glue extending from *A* to *B*.[21] This is one-way glue: if *A* causes *B*, *B* cannot cause *A*.[22]

4. Causation is transitive: if *A* causes *B* and *B* causes *C*, then *A* causes *C*.[23]

5. But Moore asserts that causation's transitivity is finite; it peters out as the links in the causal chain grow more numerous or attenuated. *A* may cause *B*, *B* may cause *C*, *C* may cause *D*, and so on through the alphabet, but at some point (around *LMNOP*, perhaps) *A* loses causal potency. Causation is a *scalar* relation that diminishes over time and space; it goes gradually from more to less and eventually to none.[24]

6. Just as physical and temporal remoteness (of putative effect from putative cause) can destroy the existence of causation, according to Moore so too can indirectness. When Jones persuades *X* to kill Smith, Jones is not a cause of Smith's death.[25]

7. Moore posits that the mysterious causal glue entails some kind of *action* by putative cause upon putative effect.[26] In Moore's world pure omissions cannot be causes of anything.[27] Thus, for example, "it is both morally and metaphysically absurd to think" that one who declines to rescue a drowning baby has been a cause of the baby's death.[28]

---

18. MOORE, *supra* note 2, at xii; *see also id.* at 104 (discussing three unified approaches to causation in the law).

19. *See, e.g.*, DAVID W. ROBERTSON, WILLIAM POWERS, JR., DAVID A. ANDERSON & OLIN GUY WELLBORN III, CASES AND MATERIALS ON TORTS 169–72 (3d ed. 2004) [hereinafter ROBERTSON ET AL., CASEBOOK] (emphasizing the wisdom of keeping the five now-traditional elements of a negligence cause of action separated from one another).

20. MOORE, *supra* note 2, at 134.

21. *See id.* at 90 (defining skepticism about causation, using the views of David Hume—whose "analysis [took] 'the glue' out of the causal relation"—and Herbert Hart as illustrations); *id.* at 94–95 (rejecting the Coasean analysis of tort law because, in Moore's opinion, "the best goal for tort law to serve is that of corrective justice"—not of achieving efficiency, and "[t]he corrective-justice view of tort law demands a robustly metaphysical interpretation of cause").

22. *Id.* at 134.

23. *Id.* at 121–23.

24. *Id.* at 71, 102, 105, 121–23, 224.

25. *Id.* at 13.

26. *See, e.g.*, *id.* at 257–58 (implying repeatedly that causation entails action); *id.* at 454 (emphasizing that omissions cannot "push, pull, or make things happen").

27. *See, e.g.*, *id.* at 436–37 (remarking that, following the generic meaning of the term, an omission is not eligible to serve as a cause); *id.* at 460–61 (arguing that double-prevention actors do not cause the harms that their actions make possible).

28. *Id.* at 480; *see also id.* at 55 ("[A]n omissive letting die is not a kind of killing."); *id.* at 62 ("[A]n omission to save someone from death is not a causing of death."); *id.* at 304 ("Absent events

8. According to Moore, tort law's dominant test for factual causation—the but-for or counterfactual dependency test—is innately too indeterminate to be useful.[29] (By this time Moore and I are about ready to part company.[30])

III. What If You Disagree?

What if you read Moore with care and an open mind but still end up unable to accept some of the foregoing postulates? Moore's answer, delivered politely and with considerable charm: You're wrong; you just don't get it. Moore carefully and cheerfully acknowledges that his postulates derive from his own intuitions about the way the world works.[31] And he is becomingly reluctant to overtly rank his intuitive skills against yours:

> There is an old saying in philosophy to the effect that one person's *reductio ad absurdum* is another person's valid inference.[32]
>
> It is difficult to argue for conclusions one finds intuitively obvious, and I must confess that I find the moral non-equivalence of acts and omissions to be obvious. So I shall cease beating (what for me is) a dead horse.[33]
>
> [T]hose [of us] who have been horrified, ashamed, or numbingly distressed, by some awfulness of which [we] were the author . . . know that causing things matters to responsibility in a way that requires no other argument. Those with either better characters or more fortunate opportunity sets will lack the relevant experience that makes this intuitively so plain to the rest of us.[34]

---

cannot serve as relata (of either kind) of the singular causal relation."). *But see id.* at 399–400 (noting, however, that the counterfactual theory of causation allows for omissions to be considered causes).

29. *Id.* at 85–86, 89, 286–88.

30. *See, e.g.*, David W. Robertson, *Causation in the* Restatement (Third) of Torts*: Three Arguable Mistakes*, 44 WAKE FOREST L. REV. 1007, 1009–11 (2009) [hereinafter Robertson, *Three Mistakes*]; David W. Robertson, *The Common Sense of Cause in Fact*, 75 TEXAS L. REV. 1765, 1768–73 (1997) [hereinafter Robertson, *Common Sense*] (both extolling the but-for test).

31. *See, e.g.*, MOORE, *supra* note 2, at 38 (using intuitive examples to illustrate that "we generally are governed by consequentialist reasons, save when we are permitted or obligated by agent-relative reasons"); *id.* at 58 ("It is difficult to argue for conclusions one finds intuitively obvious, and I must confess that I find the moral non-equivalence of acts and omissions to be obvious."); *id.* at 85 ("The third and fourth sets of problems stem from the inability of the counterfactual test to match what for most of us are firm causal intuitions."); *id.* at 140 ("The easiest, most intuitive way to draw this distinction is by using causation to mark the difference."); *id.* at 347 ("I take it my strong sense of a category mistake (occurring when one speaks of propositions causing anything) is part of a genuinely realist view about causation."); *id.* at 431 ("My own view here too is that the culpable causation of harm should be sufficient for liability."); *id.* at 443 ("What the foregoing examples do show is how unpromising is Austin's suggestion; to my ear at least, each of such pairs form equally idiomatic English."); *id.* at 467 ("Anyone who believes in the kind of moral luck defended in Chapter 2 should find this conclusion [that two simultaneous shooters should be held equally culpable for murder via vicarious attribution] absurd.").

32. *Id.* at 414.

33. *Id.* at 58.

34. *Id.* at 435.

At the end of the day, though, the reader is left with no doubt of Moore's confidence that his intuitions are probably better—more highly trained, more mature, more carefully vetted—than the reader's. The book's ultimate message, after all, is that Moore is tapped into metaphysical truth about causation and that, wherever the law diverges from that truth, the law has no choice but to shape up.

In other words, Moore claims a kind of entitlement to insist that you agree with him about causation and morality, at least in substantial part, or get lost. The perceived entitlement has a simple but powerful basis in the axiom that all other aims and interests are inferior to the pursuit of truth. Whereas many tort law analysts believe that tort law does well enough if it uses a concept of causation that its constituents recognize as respectable and can generally agree with,[35] Moore insists that the law utterly fails if its concept of causation does not match the metaphysically correct concept, *viz.*, God's own true causation.[36] As we will see in the next two Parts, this large ambition—to rethink whatever parts of the law fail to match up with the true nature of causation—pushes Moore to resect significant parts of traditional tort law, in turn requiring him to invent a bizarre prosthetic replacement.

IV. Moore's Take on Traditional Tort Law

To assess Moore's proposed revision of traditional tort thinking, we first need to get a sense of his version of the traditional. It has four key features.

A.

First is a seeming disregard for realities of litigation. Moore imagines that courts are interested in questions like "What caused the fire?"[37] or "Did defendant's roof cause the workmen to be injured?"[38] In actual tort courts, though, such broad causation questions never need to be answered and indeed do not even get asked. Tort plaintiffs whose pleadings do not identify

35. *See, e.g.*, Robertson, *Three Mistakes*, *supra* note 30, at 1010 ("In defining factual causation as but-for causation, tort law exhibits the conspicuous virtue of cleaving to the views of its constituency.").

36. *See* MOORE, *supra* note 2, at xiii (stating that the book was mostly written to correct "errors of legal theory" such as "doubt [regarding] the relevance of causation to legal and moral judgments"); *id.* at 4–5 ("[J]ustice is achieved only if the morally responsible are held liable to punishment or tort damages."); *id.* at 94–95 (criticizing an efficiency-based philosophy of tort law and advocating the corrective-justice view); *id.* at 256–57 (disparaging the "ordinary-language approach" to tort law as unhelpfully restrictive). Moore's functional disdain for the expectations of tort law's constituents is in sharp contrast to his view of insurance law, where he says "even . . . doctrines that get the metaphysics of causation wrong" are appropriate if they further the "expectations of [the] parties to insurance contracts." *Id.* at 547. Moore does not directly explain why insurance law and tort law should be so different in this respect, but presumably his answer would be that tort law is bottomed on morality—which for Moore must strive to reflect God's own morality, in turn entailing a tether to God's own causation metaphysics—while insurance law is not bottomed on morality, or at least not so much.

37. *Id.* at 264.

38. *Id.* at 124.

defendants' putatively wrongful conduct are summarily thrown out of court. Once the plaintiff has identified the conduct of the defendant that the plaintiff is prepared to try to prove was wrongful, the cause-in-fact question is almost always whether the plaintiff's injury would have been sustained if the defendant had corrected the allegedly wrongful conduct to the extent necessary to eliminate the wrongfulness that the plaintiff's suit is based on.[39] There is no occasion for asking "What caused plaintiff's harm?" or "Did the defendant cause plaintiff's harm?" The question is always a much narrower one: Did the conduct of defendant that plaintiff has singled out as sufficiently unacceptable to justify liability cause the harm?

But Moore's self-declared "armchair"[40] perspective leaves him free to contemplate causation in a broad, discursive way that is in marked contrast to the tight focus of a properly formulated tort case. We can see in Moore's discussion of three hypothetical cases the large differences between his broad focus and the pragmatism-driven traditional one. In the first case—which merely sets the stage for the discussion—the defendant uses unsuitable bolts to install a warehouse roof that then falls in on workmen during a "stiff (but not unusual) breeze."[41] Here the analysis and the imposition of liability are unproblematic.

In the second case, the roof is blown away by a windstorm that "is that kind of extraordinary event we call an 'act of God,' so that the roof does not simply fall but flies over one hundred feet before it injures its victims."[42] Moore says in this case the windstorm "will be an intervening cause relieving the defendant of causal responsibility for the injury."[43] But here Moore fails to focus on a crucial variable: Would adequate bolts have kept the roof from flying away? If not, traditional tort law will, as Moore projects, exonerate the defendant on causation grounds (although the proper explanation will be that the defendant's wrongful conduct—unsuitable bolts—was not a cause in fact of the injuries[44] rather than the proximate-cause explanation Moore suggests). But if adequate (non-negligent) bolts would have kept the roof on the warehouse, then the defendant's wrongful conduct *was* a cause in fact of

39. I have previously outlined a five-step process to conceptualize the but-for cause-in-fact question. The third step creates a counterfactual hypothesis by changing the defendant's conduct "only to the extent necessary to make it conform to the requirements of law." Robertson, *Common Sense*, *supra* note 30, at 1770. For a more in-depth discussion of the five-step process and a demonstration using a hypothetical, see *id.* at 1768–73.

40. MOORE, *supra* note 2, at 36.

41. *Id.* at 124.

42. *Id.*

43. *Id.*

44. *See* Robertson, *Three Mistakes*, *supra* note 30, at 1024–25 & nn.116–17 (discussing City of Piqua v. Morris, 120 N.E. 300 (Ohio 1918), in which the court held that the sole factual cause of the plaintiff's injuries was a catastrophic "act of God" flood).

the harm, and whether the defendant will escape liability on proximate-causation grounds is difficult to predict.[45]

In Moore's third hypothetical case, the roof is again blown away by an unusually severe windstorm, but this time the defendant, wishing to harm the workmen, has "sent [them] out . . . to work where he hoped the forthcoming storm would blow off the roof,"[46] and the scheme succeeds. Moore's analysis of this case is completely aberrant: He says that here the defendant "has used the storm for his own ends and [so the storm] is not an intervening cause."[47] Well, maybe so, but imagining this case as just another in which the alleged wrongful conduct is the condition of the roof is quite strange. On traditional tort law's view, here both the windstorm *and* the roof bolts become mere background because the defendant's intentional act of sending the workmen into harm's way would be an actionable battery regardless of how high the wind or bad the bolts.[48] (Plaintiffs sometimes have reasons—generally having to do with insurance coverage—to plead intentional tort cases as negligence cases,[49] but even then nobody would plead this as a negligent roof-bolts case; it would be pleaded as a case of negligently sending workers into harm's way.)

In sum, in the second and third cases Moore uses a vague and sweeping "intervening cause" idea to treat in a somewhat cursory way issues for which traditional tort law has better—more precise and more powerful—tools. Moore is not looking at tort law as closely as I wish he would.

B.

The second key feature of Moore's version of traditional tort law is a misunderstanding of how the but-for test for factual causation actually works. This misunderstanding lies at the heart of Moore's rejection[50] of the but-for test. Moore explains his difficulty with the test as follows:

---

45. *See* ROBERTSON ET AL., CASEBOOK, *supra* note 19, at 180–81 (presenting cases coming out both ways on whether a plaintiff who has sustained a generally foreseeable type of injury will be defeated on proximate-causation grounds when the way in which the injury came about seems freakish).

46. MOORE, *supra* note 2, at 124.

47. *Id.* Why does Moore believe that a killer who uses a storm as his instrumentality is a cause of the death whereas a killer who persuades another person to do the deed is not? *Id.* at 12–13. Apparently solely because of a postulate that "free, informed, voluntary actions by [human beings necessarily] break causal chains." *Id.* at 123.

48. *See* RESTATEMENT (SECOND) OF TORTS § 13 (1965) (defining the tort of battery to include intentionally causing a harmful bodily contact); *id.* § 8A (defining intent to include purpose and substantially certain knowledge).

49. *See* Ellen S. Pryor, *The Stories We Tell: Intentional Harm and the Quest for Insurance Funding*, 75 TEXAS L. REV. 1721, 1723–24 (1997) (explaining that some plaintiffs choose to plead and prove negligence in intentional tort cases because standard liability policies do not cover harms intentionally caused by the insured).

50. The rejection is nuanced. Moore ranges over four possible treatments of the but-for test. (1) Does it capture the essence of causation, as Justice Kennedy once claimed? *See* Price Waterhouse v. Hopkins, 490 U.S. 228, 282 (1989) (Kennedy, J., dissenting) (stating that using

> Suppose a defendant culpably destroys a life-preserver on a sea-going tug. When a crewman falls overboard and drowns, was a necessary condition of his death the act of the defendant in destroying the life-preserver? If the life-preserver had been there, would [the life have been saved? There is] an indeterminacy of meaning in the [but-for] test . . . . There is a great vagueness in counterfactual judgments. The vagueness lies in specifying the possible world in which we are to test the counterfactual. When we say "but for the defendant's act of destroying the life-preserver" what world are we imagining? We know we are to eliminate the defendant's act, but what are we to replace it with: a life-preserver that was alternatively destroyed by the heavy seas; a defendant who did not destroy the life-preserver because he had already pushed the victim overboard when no one else was around to throw the life-preserver to the victim; etc, etc?[51]

Traditional tort law has an accepted answer to Moore's questions, one that blunts his "indeterminacy" criticism. The imagined counterfactual world must be the same as the actual world as shown by the evidence in the case in all respects save one: the defendant's wrongful conduct must be corrected to the extent necessary to make the conduct acceptable under plaintiff's theory of the case. This imagined correction of the defendant's conduct is the only allowable change, and this change must be done in an intellectually conservative way, employing as little creativity as possible.[52] To lend emphasis to this crucial point, let us phrase it this way: The *imagined* correction of the defendant's conduct must never be *imaginative*.

A good illustration of the courts' acceptance of these constraints on counterfactual world construction—constraints that say *curb the imagination, keep the changes as modest as possible*—is Judge Friendly's opinion for the Second Circuit in *Lekas & Drivas v. Goulandris*.[53] A ship loaded cheese at Salonika, Greece, intending to carry it via the Mediterranean, Gibraltar, and the North Atlantic to New York, a projected one-month, 5000-mile journey

---

"[a]ny standard less than but-for . . . represents a decision to impose liability without causation"). Moore answers of course not; "counterfactual theory cannot be a theory of causation." MOORE, *supra* note 2, at 411. (2) Is the but-for test an acceptable "heuristic for the existence of causation"? Moore says no. *Id.*; *see also id.* at 425 (indicating the but-for test fails even as "a mere heuristic for causation"). (3) So should it be jettisoned? No, says Moore. *Id.* at 426. (4) Instead, the but-for test is a valid "desert-determiner independent of causation." *Id.*

51. MOORE, *supra* note 2, at 85.

52. *See* ROBERTSON ET AL., CASEBOOK, *supra* note 19, at 124 (delineating the creation of a counterfactual situation where wrongful conduct is corrected only to the minimal extent necessary to conform to the requirements of the law); Robertson, *Common Sense*, *supra* note 30, at 1770 & n.21 (explaining that the mental operation involved in the creation of a counterfactual scenario must be "careful, conservative, and modest"); *cf.* RESTATEMENT (THIRD), *supra* note 17, § 26 cmt. f ("[W]hen the actor's conduct is tortious only because it is marginally more risky than nontortious conduct, the causal inquiry must be framed by the incremental risk of the tortious conduct, as distinguished from the risk posed by the entirety of the conduct.").

53. 306 F.2d 426 (2d Cir. 1962).

in cool weather.[54] But before the ship left port, World War II intervened, forcing the ship to take a 13,000-mile detour through the Suez Canal, around Africa's Cape of Good Hope, and thence eventually to New York.[55] Because of wartime conditions, this detour turned into a six-month journey that exposed the ship to extreme heat, and the cheese arrived at New York rotted and useless.[56]

The question arose whether the shipowner's negligence in stowing the cheese in the ship's poop, a poorly-ventilated spot, was a factual cause of the spoilage.[57] Answering yes, the trial judge constructed a counterfactual world in which the cheese remained in the poop and the ship took the short North Atlantic trip (and the cheese still spoiled).[58] This was a remarkably ambitious counterfactual construction. It imagined World War II out of existence. Moreover, it failed to do the one essential thing: change the defendant's putatively wrongful conduct.

Judge Friendly reversed the trial judge on this point, pointing out that the proper counterfactual world was one in which the cheese was stowed at a better ventilated spot on the ship and then taken on the six-month detour through some hellishly hot weather.[59] It was utterly clear that no cheese anywhere on that (unrefrigerated) ship could have survived the trip to hell and back, so the absence of factual causation was obvious.[60]

So, to answer Moore's questions in the life-preserver case, the counterfactual world that is constructed to apply the but-for test must posit that an undestroyed life preserver was at its accustomed place on the tug the moment after the defendant's actual destruction of it was finished. Nothing else in the real world must be changed. The evidence—and not anyone's imagination—must answer such questions as whether a fellow crew member would have been available to throw the preserver to the victim and, if so, whether this would have done any good. These are tough questions—and if

---

54. *Id.* at 427.

55. *Id.*

56. *Id.*

57. *Id.* at 430.

58. *Id.*

59. *Id.*

60. For another demonstration of the discriminatory power of the modest-change-only constraints, consider the disagreement between the majority and dissenting opinions in *Kernan v. American Dredging Co.*, 355 U.S. 426 (1958). A portable open-flame kerosene lantern, attached to a barge three feet above the water line (and serving no useful purpose there), ignited petrochemical vapors lying just above the surface of the river, causing a fatal fire. *Id.* at 427. The factual-causation question was whether the barge's violation of a Coast Guard regulation requiring the barge to carry such a light at least eight feet above the water line was a cause of the fire. *Id.* at 430. The dissenters thought probably not, believing that the proper counterfactual construction entailed leaving the three-foot light where it was (and where it would still have set fire to the river) and adding another one at eight feet. *Id.* at 442 n.1 (Harlan, J., dissenting). The majority's thinking was more prosaic (and hence better): Obviously the more straightforward imagined correction was moving the light up to the proper spot (thus removing the ignition source) rather than getting and installing a new light. *Id.* at 433 (majority opinion).

the plaintiff cannot convince the trier of fact that the undestroyed preserver would probably have saved him, he will lose the case[61]—but they relate to the availability of evidence of facts in the real-world case, not to any "indeterminacy" flaw in the but-for test itself.

C.

The third key feature of Moore's version of traditional tort law is an assumption that tort law's proximate-cause requirement is based on moral grounds.[62] I think that assumption is probably wrong. It is true that one can find plenty of authority associating the proximate-cause requirement with a morality-based need to delimit a tortfeasor's appropriate sphere of responsibility.[63] But the premise has often been eloquently questioned,[64] and the *Palsgraf* decision[65]—typically cited as exemplifying a morality-based

---

61. The general rule requiring plaintiffs to establish but-for causation by a preponderance of the evidence is appropriately relaxed in certain limited types of cases in which the burden seems unduly harsh. *See* Robertson, *Common Sense*, *supra* note 30, at 1775 (identifying eight situations in which courts may employ alternative approaches to causation). If the court in the destroyed-life-preserver case perceives that injustice would result from requiring the plaintiff to show that an undestroyed life preserver would have saved the seaman's life, it should turn to an insight that Judge Calabresi has called the "causal link" concept:

> When a defendant's negligent act is deemed wrongful precisely because it has a strong propensity to cause the type of injury that ensued, that very causal tendency is evidence enough to establish a prima facie case of cause-in-fact. The burden then shifts to the defendant to come forward with evidence that its negligence was not such a but-for cause.

Liriano v. Hobart Corp., 170 F.3d 264, 271 (2d Cir. 1999) (emphasis omitted); *see also* Zuchowicz v. United States, 140 F.3d 381, 390–91 (2d Cir. 1998) (noting that tort law recognizes a causal link between a particular act and an outcome where the act was deemed wrongful because it increased the likelihood of the harm experienced).

62. *See, e.g.*, MOORE, *supra* note 2, at 97–100 (arguing that proximate-cause tests should serve the function of sorting offenders between more blameworthy and less blameworthy causers of harm); *id.* at 156–61 (discussing tort and criminal law's use of an actor's mental state to classify the actor's culpability). For me, Moore's largest misstep is the conflation of the cause-in-fact and proximate-cause inquiries, so that the factual causation question becomes entangled with relatively amorphous normative considerations. *See supra* notes 17–19. Entailed in that misstep is the matter treated in this subpart: an arguable mistake as to the nature of the norms that are brought to the mix by the proximate-cause ingredient.

63. *See, e.g.*, Martin A. Kotler, *The Myth of Individualism and the Appeal of Tort Reform*, 59 RUTGERS L. REV. 779, 806 (2007) (discussing proximate-cause rules that "allow an admittedly negligent defendant to escape all liability because of the intervening culpable misconduct of the plaintiff or a third person").

64. *See, e.g.*, *In re* Kinsman Transit Co., 338 F.2d 708, 725–26 (2d Cir. 1964) (reasoning that the "existence of a less likely additional risk that the very forces against whose action [the defendant] was required to guard would produce other and greater damage than could have been reasonably anticipated should inculpate him further rather than limit his liability," but acknowledging that the line must be drawn somewhere); Fazzolari v. Portland Sch. Dist., 734 P.2d 1326, 1331 & n.9 (Or. 1987) (noting some contrary arguments to the *Palsgraf* rule's demand of a closer link between negligence and its premise of fault; for example, that a defendant whose conduct has been substandard should bear the resulting loss rather than a blameless plaintiff, or that the extent of liability should not be limited where aggregate risks may make conduct negligent even though a specific, isolated risk might not be foreseeable or unreasonable).

65. Palsgraf v. Long Island R.R. Co., 162 N.E. 99 (N.Y. 1928).

approach to delimiting tort defendants' responsibilities[66]—need not and perhaps even cannot sensibly be read that way. It really seems almost absurd to say that a railroad has no moral responsibility to an intending passenger, waiting for a train with a paid-for ticket in her hand, to avoid injuring her by precipitating an explosion that causes a heavy machine belonging to the railroad to fall on her. So I do not think Helen Palsgraf's case failed because the defendant lacked moral responsibility to her; I think it failed because Judge Cardozo and three other judges thought the appropriate persons to sue the railroad were those most immediately threatened by the conduct of the railroad in precipitating the explosion (none of whom, as it happened, were hurt).

What the proximate-cause limit seems to be, rather than a moral-responsibility doctrine, is a justice-rationing device—meant largely for the courts' own protection—that is closely analogous to the standing-to-sue requirement.[67] Some of the U.S. Supreme Court's standing-to-sue cases use concepts and vocabulary that are easily associated with the everyday stuff of proximate cause. For example, in *Ass'n of Data Processing Service Organizations v. Camp*,[68] the Court made standing turn on whether the plaintiff's challenge to a provision of federal law fell "within the zone of interests to be protected or regulated" by the provision.[69] In *Barlow v. Collins*,[70] the Court indicated that standing to invoke a statute's protection turned on whether there was legislative "intent to protect the interests of the class of which the plaintiff is a member,"[71] and it suggested that the plaintiff's asserted grievance must flow more or less "directly" from the challenged governmental action.[72]

Conversely, proximate-cause discussions in tort cases sometimes sound quite a bit like discussions of standing. In *Edwards v. Honeywell, Inc.*,[73] a

66. *See, e.g.*, MOORE, *supra* note 2, at 164 & nn.27–30 (citing *Palsgraf* as an implementation of Bingham's Harm-within-the-Risk theory, in which "[t]he question is whether defendant breached his duty, and this question is to be framed in terms of whether 'the specific consequence comes within the limits of defendant's responsibility for his wrong'" (quoting Joseph W. Bingham, *Some Suggestions Concerning "Legal Cause" at Common Law* (pt. 1), 9 COLUM. L. REV. 16, 25 (1909))).

67. Analogizing the proximate-cause limitation to the standing-to-sue requirement is a new idea for me. The presentation of the idea here is a sketch of a fuller treatment that I hope to publish before too long. At this point I am not claiming that the standing analogy justifies the limitation any better than the morality-based idea, but only that it seems to explain it better.

68. 397 U.S. 150 (1970).

69. *Id.* at 153.

70. 397 U.S. 159 (1970).

71. *Id.* at 167.

72. *Id.* at 163; *see also* Bennett v. Spear, 520 U.S. 154, 162 (1997) ("Numbered among these prudential requirements is the doctrine of particular concern in this case: that a plaintiff's grievance must arguably fall within the zone of interests protected or regulated by the statutory provision or constitutional guarantee invoked in the suit.").

73. 50 F.3d 484 (7th Cir. 1995). Judge Posner phrased his ruling in "no duty" terms, but this was functionally a proximate-cause case. *See* ROBERTSON ET AL., CASEBOOK, *supra* note 19, at 199–200 (explaining Posner's selection of the "no-duty" articulation as a rhetorical choice, not an analytically driven one).

fireman in a burning residence died when the floor unexpectedly collapsed.[74] The collapse was unexpected because, unbeknownst to the firemen, the company with which the homeowner had a fire-alarm contract had dithered in notifying the fire department, so that the fire was further advanced than the firemen had reason to think based on their normal experience.[75] Judge Posner granted summary judgment that the fireman's widow had no rights against the fire-alarm company, explaining that the most appropriate plaintiff to sue the fire-alarm company was the homeowner and the most appropriate defendants for the plaintiff to sue were the homeowner and the fire department.[76] Posner said that given "the plethora of potential [more appropriate] defendants," "the cost of administering" the fire-alarm company's responsibilities to the firemen would be too excessive to justify allowing the case to proceed.[77] This is not precisely standing-to-sue talk, but it is not too far off. Moreover, Posner cited a proximate-cause decision with facts and reasoning that are strongly suggestive of a standing-to-sue rationale.[78]

D.

The fourth noteworthy feature of Moore's take on traditional tort law can charitably be put as inattention to small details and less charitably as getting lots of little (and sometimes not so little) things wrong. Here is a partial list:

- On page xii: "[Traditional tort law posits] that there is *a* risk that makes one negligent so that in each case it should be asked whether the harm that happened was within that risk." But compare Moore's emphasis on a singular risk to the Restatement (Third) of Torts: "An actor's liability is limited to those physical harms that result from the *risks* that made the actor's conduct tortious."[79] And from the comments to that Restatement provision: "[T]he jury should be told that, in deciding whether plaintiff's harm is within the scope of liability, it should go back to the *reasons* for finding the defendant engaged in negligent or other tortious conduct."[80] The above difference may look small, but it is crucial; one cannot understand why

74. *Edwards*, 50 F.3d at 485.

75. *Id.* at 486.

76. *Id.* at 491–92.

77. *Id.* at 491.

78. In *Mang v. Palmer*, 557 So. 2d 973 (La. Ct. App. 1989), a 16-year-old boy had to walk home from work, using a busy highway overpass, because he could not get through on the telephone to his mother to alert her to come pick him up—her phone, supplied by the defendant, was defective. *Id.* at 973–74. The boy was injured by overpass traffic. *Id.* at 974. Phrasing its decision in traditional proximate-cause language, the court said that the mother had the right to sue the phone company for providing a defective phone but the boy did not. *Id.* at 975.

79. RESTATEMENT (THIRD), *supra* note 17, § 29 (emphasis added).

80. *Id.* cmt. d (emphasis added).

proximate cause is such a recurrently difficult issue without knowing that there is normally a bundle of risks rather than just one.

- On page 70: "[W]hen one party sets a chain of events in motion that increase the probability of some harm but another party intervenes by intentionally causing the harm the first party was trying to cause or risked causing, [traditional tort law concludes] that only the later party caused the harm." Well, sometimes that is true and sometimes it is not. There are hundreds of cases that put the lie to Moore's categorical generalization here.[81]
- On page 86: "The defendant's fire arrives first and burns down the victim's building; the second fire arrives shortly thereafter, and would have been sufficient to have burned down the building . . . . [Traditional tort law holds] that neither fire caused the harm . . . ." This is a bad mistake on Moore's part. Traditional tort law holds that the defendant's fire was a cause in fact of the harm, although if the second fire was imminent enough, the value of the building at the time of defendant's fire might have been too small to be compensable.[82]
- On page 93: Moore asserts that the unimportance of factual causation to the economic-efficiency theory of torts has gone "unnoticed" by the legal economists. But Landes and Posner noticed and indeed seemed to emphasize this very thing: "[T]he idea of causation can largely be dispensed with in an economic analysis of torts . . . ."[83]
- On page 206: "[Traditional tort law's doctrine of transferred intent does] not transfer intentions across types of harms." But as the Restatement (Second) of Torts makes plain, intent is freely transferrable between battery and assault.[84]

Such mistakes, some quite trivial,[85] are understandable in the work of a prolific philosopher–lawyer who regards himself as a philosopher first and a

---

81. *See, e.g.*, d'Hedouville v. Pioneer Hotel Co., 552 F.2d 886, 890–94 (9th Cir. 1977) (holding a manufacturer of flammable carpet responsible for arson); Meyering v. Gen. Motors Corp., 275 Cal. Rptr. 346, 350–51 (Cal. Ct. App. 1990) (allowing suit against the manufacturer of a weak plexiglass sunroof for injuries from hooligans' throwing of a chunk of concrete through it).

82. *See* Robertson, *Three Mistakes*, *supra* note 30, at 1026–28 (attempting to reconcile differing treatments of a duplicated-harm situation in which an unsuccessful parachutist is shot at the moment of impact—Anderson v. Minneapolis, St. P. & S.S.M. Ry. Co., 179 N.W. 45 (Minn. 1920), seemingly would allow a claim for full damages, while Dillon v. Twin State Gas & Elec. Co., 163 A. 111 (N.H. 1932), would suggest no damages—as part of a discussion of the Restatement (Third) of Torts § 29). *See generally* Michael D. Green, *The Intersection of Factual Causation and Damages*, 55 DEPAUL L. REV. 671 (2006) (analyzing the intersection of causation and damages in tort law).

83. WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW 229 (1987).

84. RESTATEMENT (SECOND) OF TORTS §§ 18(1)(a), 20(1), 21(1)(a), 32(1) (1965).

85. This one is not even relevant, but I cannot resist flagging it: at page 291, Moore says that "rivers are legally *navigable* [for purposes of admiralty jurisdiction] even though in fact one cannot float a canoe on them." This is 180-degrees wrong. *See* LeBlanc v. Cleveland, 198 F.3d 353, 357

lawyer distantly second.[86] I do not cite these mistakes to suggest that Moore's legal work is untrustworthy—I intend no such suggestion. I cite them as evidence in support of an inference: Moore's relative lack of attention to the nitty-gritty operations of tort law signals that he does not vouch for his recommended changes to tort law thinking as operational improvements.[87] So it may not be a huge surprise if they turn out not to be operational improvements.

V. Proposed Alterations of Tort Law and Tort Thought

Moore recommends a handful of doctrinal changes. For example, he thinks the doctrine of transferred intent extends tort liability far beyond the legitimate reach of causation, so he would abolish that doctrine, deeming negligence liability a sufficient replacement.[88]

For another example, Moore concludes that liability for negligent entrustment—whereby, say, you may be liable to persons Mel Gibson shoots if you hand Mel a gun when he is obviously in one of his murderous moods—is not properly viewed as entailing liability for providing the gun but rather for failing to prevent the shootings.[89] Failing to prevent something is an omission, and, for Moore, "an omission is literally nothing at all."[90] So Moore thinks it may be sensible to abolish negligent-entrustment liability.[91]

Because of space limitations, I can only say that neither of the above doctrinal proposals seems a self-evident improvement. A third doctrinal proposal calls for slightly more attention. Moore believes that the proximate-cause limitation on liability for negligence might usefully be replaced by a system in which every blameworthy actor whose conduct was a true cause[92]

(2d Cir. 1999) ("'Those rivers must be regarded as public navigable rivers in law which are navigable in fact. And they are navigable in fact when they are used, or are susceptible of being used, in their ordinary condition, as highways for commerce, over which trade and travel are or may be conducted in the customary modes of trade and travel on water.'" (quoting The Daniel Ball, 77 U.S. (10 Wall.) 557, 563 (1870))).

86. This is pretty much Moore's self-description:

> This book is mostly written to correct . . . errors of legal theory. That there are also practical, doctrinal pay-offs is nice but not essential. Law schools were once defined as the unholy mixture of Plato's Academy and the training ground for young Hessian mercenaries. This book is on the Academy side of that line. I assume that it is worthwhile figuring out the function and nature of a property like causation, both in the law and out. There are practical benefits of achieving such understanding, but the reason for making the effort is, for me, mostly because it is there to be understood.

MOORE, *supra* note 2, at xiii–xiv.

87. *See id.* at xiii ("The main reason why the book gives only occasional attention to legal tests for causation stems from my interests in writing it, which are more theoretical than practical.").

88. *Id.* at 207.

89. *See id.* at 147–48 (arguing that negligent-provision-of-opportunity liability is not cause-based but true omission liability).

90. *Id.* at 53.

91. *Id.* at 147.

92. Here *true cause* would mean something like "more-or-less direct and active applier of causal glue."

of harm would bear some liability, with the actors' degrees of causal input being ranked and mulcted on some numerical basis.[93] Regardless of the purposes of the proximate-cause limit, this proposal is unattractive because it would bring an entire new level of complexity to multiparty torts litigation.[94] Moreover, if I am right that the proximate-cause limit is best viewed as a standing-to-sue doctrine,[95] this supplies an additional argument against Moore's proposal: Standing is innately a yes–no question.

Turning from doctrinal matters to Moore's central conceptual recommendation, we see him concluding that tort law needs to reclassify a significant number of its resolutions so as to rid itself of erroneous claims about causation and instead confess that these resolutions are instances of "non-causal liability."[96] How would this work? Moore provides a vivid demonstration: "*D* sees his old enemy drowning in the ocean, which makes *D* happy. Then *D* spies a lifeguard getting ready to save the drowning man. So *D* ties up the lifeguard, and *D*'s old enemy drowns."[97] Moore says this is a case of "non-causal" criminal and tort liability.[98]

That is a radical reclassification. Why do it? It is necessary, says Moore, simply because the law must never make metaphysically false claims about causation.[99] In Moore's metaphysics, *D* cannot have caused the drowning because he provided no active causal glue: He did not throw his enemy in the sea, or hold the enemy's head under, or really *do* anything other than happily watch him drown, and omissions can never be causal. The business of tying up the lifeguard was a metaphysical sideshow of some kind; it was just a way of facilitating *D*'s causational-metaphysics-entailed entitlement to happily stand by and watch.

But the entitlement to stand and watch is metaphysical only, and it quickly expires. Moore readily acknowledges that *D*'s conduct is condemned by morality, criminal law, and tort law.[100] *D* is as clearly liable as any tortfeasor ever was; we just cannot call the liability causal.

---

93. MOORE, *supra* note 2, at 279–80.

94. Undue complexity was the main reason the Supreme Court unanimously rejected a version of Moore's proposal in *Exxon Co. v. Sofec*, 517 U.S. 830, 837–39 (1996). For a striking demonstration of some of the kinds of complexity adopting the proposal could entail, see Otal Invs. Ltd. v. M/V Clary, 2008 WL 2844019, Nos. 03 Civ. 4304(HB), 03 Civ. 9962, 04 Civ. 1107, at *4–14 (S.D.N.Y. June 23, 2008), in which the court assigned each of three vessels involved in a marine collision a culpability percentage and a causation percentage and then averaged the two to yield a liability-determining percentage. *Otal* is presently on appeal to the Second Circuit.

95. *See supra* subpart IV(C).

96. *See* MOORE, *supra* note 2, at 145 (stating that "[o]ften I shall urge . . . non-causal" liability); *id.* at 251 (referring "again [to] a non-causal liability").

97. *Id.* at 62. For ease of presentation I have slightly altered the way the actors are identified.

98. *Id.* at 63.

99. *See id.* at 62–65 (contending that double-prevention scenarios may create moral responsibility but not causality).

100. *See id.* at 63 (stating that *D* "is plainly guilty of murder" and that many people "perceive, correctly, that [*D*] is morally responsible" for the drowning).

Notice what has happened. Moore began his journey with two postulates: (1) Non-causal moral violations, and hence non-causal tort liability, should be shrinkingly rare, and (2) metaphysically correct causation principles rule out causation for omissions, unduly indirect input, and too-remote input.[101] Very soon these postulates are at war with one another. Moore struggles for peace for a while, but his solution is ultimately to abandon the first postulate: Thus he presents us with a radically new version of tort law riddled right through with non-causal liability.

I do not understand Moore's choice. Why is it worse to recognize causation without active glue than to invent a category of moral responsibility for non-caused harm? But that philosophical question is an aside for me. For me the bottom line is that non-causal liability could not possibly be a useful addition to the tort law tool kit. The concept's label alone is enough to condemn it. To anyone who holds that the corrective-justice theory is an important part of the justification for tort law, *non-causal liability* is an oxymoron.[102] Moreover, the concept reeks of indeterminacy: When would non-causal liability be appropriate? The only general answer I can infer from Moore is, when the defendant's conduct was sufficiently blameworthy. And when would that be? I think we would pretty much be starting from scratch.

VI. Conclusion

Although I disagree with about half of it, I am not presently challenging the content of Moore's causation metaphysics. My challenge is to the idea that the law's implicit metaphysics must yield whenever another, truer, metaphysical system is revealed.

What tort law needs to keep trying to do is provide its constituents with simple, workable, and acceptable resolutions of their conflicts. Elsewhere I have argued that the but-for test comes closest to matching what thoughtful people believe causation means.[103] Moore says God knows they are wrong. I am sorry to hear that. But tort law cannot necessarily afford to aim at reflecting God's metaphysics of causation. It needs to stick with its more modest aim: trying to provide citizens who lose tort cases with explanations they can accept at some level as legitimate and thus learn to live with. I shudder to think what these citizens might make of a pervasive "non-causal liability."

---

101. *Id.* at vii.

102. *See* Robertson, *Three Mistakes*, *supra* note 30, at 1008 ("[T]he cause-in-fact requirement is the 'linchpin' of the corrective-justice theory." (quoting Larry A. Alexander, *Causation and Corrective Justice: Does Tort Law Make Sense?*, 6 LAW & PHIL. 1, 12 (1987))).

103. *See, e.g.*, *id.* at 1008–11 (asserting that the but-for test "exhibits the conspicuous virtue of cleaving to the views of its constituency").

VII. Postscript: Craft and Style

The book's subtitle says it is "an essay," but in fact it is a republication of a dozen or so essays that were originally published in the period from 1993–2009.[104] Moore says he has "sought to rewrite [them] to eliminate redundancies, supply cross-references and ease transitions,"[105] but this effort was not wholly successful. There is a great deal of repetition—the book could probably have shed a couple of hundred pages without losing anything—and there is also some outright self-contradiction. For example:

- Page 8 says it is inappropriate to say "Smith killed Jones" unless the causal mechanism was direct, immediate, and involved no intermediaries. But page 51 says "we kill quite often, [as when we] build tall buildings," knowing that some workers will die, and page 151 says that those who induce others to commit violent crime "quite literally rape, hit, maim, and take."
- Page 100 repeatedly refers to negligence as a "mental state," but page 178 says "negligence does not involve a state of mind."
- Page 480 says that "it is both morally and metaphysically absurd" to think omissions can be causative, but page 140 says "[t]he problem with [deeming omissions causative] is not . . . metaphysical impossibility; rather, the problem is moral."

There are also far too many typographical errors and similar errata for a book with such an exalted editorial pedigree. There are dozens, perhaps hundreds, of these (a list will be furnished on request). Some are merely irritating, such as the repeated phrase at page 117, note 27: "It might be otherwise if Betty's and Susan's omissions succeed on[e] another in time, where there is some authority that the later pre-empts the earlier (and so only the later pre-empts the earlier and so only the later causes the harm)."

But other errata are devilishly confusing or meaning distorting. On page 59, Moore says personal inconvenience can morally justify a physician's refusal to save the life of a nonpatient, adding: "[T]his is true even though that same level of inconvenience would hardly justify one in not completing a killing in which one is engaged." The word *not*, I believe, reverses Moore's intended meaning. Similarly, at page 68, the following passage appears: "Are accomplices who do no killing themselves as blameworthy as the principals who do? Despite the formal doctrines of modern Anglo-American criminal law equating accomplices to principals, I think the general answer is yes." I think Moore meant to say "the general answer is no."

For a third example of a meaning-distorting error, see page 247, note 94, where the U.S. Supreme Court is quoted as saying:

---

104. MOORE, *supra* note 2, at xviii–xix.
105. *Id.* at xi.

> We do not say that even the natural and probable consequences of a wrongful act or omission are in all cases chargeable to the misfeasance or nonfeasance. They are not when there is a sufficient and independent cause operating between the wrong and the injury. In such a case the resort of the sufferer must be to the originator of the original cause.

The concluding phrase should read "the originator of the intermediate cause."[106]

For the most part—errata aside—Moore's writing is clear and careful, but occasionally the reader runs up against near-impenetrability. Here is my favorite: "[I]f an omission, say, to rescue a stranger baby is something I am not deontologically obligated not to do, then I am permitted to justify such an omission on consequentialist grounds."[107] My suggested translation: *If I am not deontologically obliged to rescue a stranger baby, then I can justify failing to do so on consequentialist grounds.* My translation may not have commended itself to Moore because it moves the statement closer to revealing its tautological nature. If (as I suspect but cannot say for sure[108]) the sometimes-synonym "absolutely" can fairly be substituted here for "deontologically," the tautology is laid bare.[109]

106. For other examples of significantly confusing or meaning-distorting errata, see page 29, where in the middle of a hypothetical case in which *D* is a shooter and *V* is a shooting victim, somebody named "Bob" makes a sudden appearance, with no introduction or explanation; page 185, which cross-references "the succeeding sections of Part III" when evidently intending to say "of Chapter 8"; and page 213, which states that a good test for when a criminal statute can be used as a negligence-per-se rule would distinguish "those whose allegedly tortious behavior is, in fact, also criminal from those who allegedly tortious behavior is not" when probably intending to say "those whose allegedly criminal behavior is . . . also tortious from those whose allegedly criminal behavior is not."

107. MOORE, *supra* note 2, at 39.

108. *See supra* note 8.

109. There may be a special vocabulary that rescues Moore's statement from mere tautology, but I could not find it in the book.

* * *

Notes

A Dissentious "Debate": Shaping Habeas Procedures Post-*Boumediene*[*]

I. Introduction

The United States' formulation of detention policy for Guantanamo Bay enemy combatants has not exactly taken the path of least resistance. The rules governing military detentions at Guantanamo have emerged as the product of an expansive dialogue, vacillating and at times contentious, between the Executive, Legislative, and Judicial Branches. In *Boumediene v. Bush*,[1] the latest chapter in this ongoing discussion, the Supreme Court held by a vote of five to four that Guantanamo detainees possess a constitutional privilege of habeas corpus.[2]

But the widespread clamor over whether *Boumediene*'s brazen holding was the "correct" result has overshadowed a quieter but equally significant—and indeed, equally contentious—conversation that also was brought about by *Boumediene*: the Washington D.C. District Court's endeavor to design appropriate procedures to govern the Guantanamo detainees' habeas petitions. At first glance, the minutiae of habeas procedures may appear immaterial when compared to the more comprehensive constitutional issues confronted by the *Boumediene* Court. After all, the Court addressed the details of the proceedings only in passing: "We make no attempt to anticipate all of the evidentiary and access-to-counsel issues that will arise during the course of the detainees' habeas corpus proceedings."[3]

This Note demonstrates that to the contrary, the D.C. District Court's post-*Boumediene* mandate to design procedures for the Guantanamo detainee habeas proceedings, as well as its adjudicatory role in the disposition of these proceedings, substantially impact the interbranch national-security dialogue. And so far, the rumblings at the district court level have been anything but uneventful.

First, I must place this dialogue within its proper context. Part II provides this frame of reference by contouring the key exchanges between the Judicial, Legislative, and Executive Branches since the commencement of the War on Terror, which cumulatively have shaped the United States' noncitizen-detention policy to this point. By examining these decisions, and the reactions to these decisions, I find that *Boumediene* buttresses, rather than

---

* I extend my sincere appreciation to the Texas Law Review for its editing expertise and to Professor Robert Chesney for his insight and guidance.

1. 128 S. Ct. 2229 (2008).
2. *Id.* at 2240.
3. *Id.* at 2276.

introduces, a trend of judicial involvement in the post-9/11 national-security-policy conversation.[4] Part III presents a narrative of the D.C. District Court's development of habeas procedures from *Boumediene* to the present, focusing specifically on the two procedural issues on which the judges have been most divergent: the disclosure of exculpatory evidence and the regulation of discovery. An analysis of the judges' case management orders (CMOs), amended CMOs, and nondispositive orders illuminates the mini-dialogue that has accompanied the crafting of these procedures from the outset. Lastly, Part IV comments on the D.C. District Court's recent activity through a dialogic lens and posits that congressional involvement, particularly in the form of statutory habeas procedures, would ensure a more interbranch, systemically balanced construction of U.S. detention policy.

II. The Interbranch Guantanamo Detention Dialogue from *Rasul* to *Boumediene*

The current fractious national debate over U.S. detention policy at Guantanamo Bay stands in stark contrast to the unified support the government received for its military initiatives in the aftermath of 9/11. Less than a week after the Twin Towers collapsed, Congress enacted and the President signed the Authorization for Use of Military Force Against Terrorists (AUMF).[5] This joint resolution granted the President broad authority to "use all necessary and appropriate force" against "nations, organizations, or persons" who were involved in the 9/11 attacks or who harbored those involved.[6] The resolution passed overwhelmingly—by 98 to 0 in the Senate[7] and 420 to 1 in the House.[8] No legal objections hindered its passage, and public opinion expressed sweeping approval of the initiatives.[9]

---

4. *See, e.g.*, Hamdan v. Rumsfeld, 548 U.S. 557, 576–77 (2006) (holding that the jurisdiction-stripping provision of the Detainee Treatment Act does not apply to habeas petitions pending at the time the statute was enacted); Hamdi v. Rumsfeld, 542 U.S. 507, 509 (2004) ("[D]ue process demands that a citizen held in the United States as an enemy combatant be given a meaningful opportunity to contest the factual basis for that detention before a neutral decisionmaker."); Rasul v. Bush, 542 U.S. 466, 484 (2004) (holding that statutory habeas corpus "confers on the District Court jurisdiction to hear petitioners' habeas corpus challenges to the legality of their detention" at Guantanamo Bay).

5. Pub. L. No. 107-40, 115 Stat. 224 (2001) (codified at 50 U.S.C. § 1541 note (2006)).

6. *Id.* sec. 2(a). The full provision authorizes the President

> to use all necessary and appropriate force against those nations, organizations, or persons he determines planned, authorized, committed, or aided the terrorist attacks that occurred on September 11, 2001, or harbored such organizations or persons, in order to prevent any future acts of international terrorism against the United States by such nations, organizations or persons.

Id.

7. 147 CONG. REC. 17,045 (2001).

8. *Id.* at 17,156.

9. *See, e.g.*, Richard Morin & Claudia Deane, *Poll: Strong Backing for Bush, War; Few Americans See Easy End to Conflict*, WASH. POST, Mar. 11, 2002, at A1 (reporting that an ABC

Of course, the United States' counterterrorism operations resulted in the apprehension of suspected terrorists, and in early 2002 the first terrorism detainees were taken to Guantanamo Bay, Cuba.[10] In response to litigation on behalf of the detainees, the Executive Branch took the position that captives it determined were "enemy combatants" could be detained indefinitely under its "plenary authority to detain pursuant to Article II of the Constitution."[11] As the Supreme Court weighed in on the issue of military detention for the first time post-9/11 in *Hamdi v. Rumsfeld*,[12] a plurality chose not to adopt the Executive's broad reasoning; instead, the plurality upheld detention on statutory grounds under the AUMF.[13]

But the Court was even less deferential to the Executive Branch when it came to opportunities for detainees to challenge the United States' grounds for detention. Weighing the Executive's concern for the prevention of future terrorist activities against Hamdi's interest in being free from unwarranted physical detention,[14] four members of the Court held that Hamdi must be afforded a mechanism through which he received the "factual basis for his classification [as an enemy combatant], and a fair opportunity to rebut the Government's factual assertions before a neutral decisionmaker."[15] Although the Court's decision did not make clear which detainees would qualify for these procedural safeguards,[16] there was no question *Hamdi* put the onus on the Executive to ensure minimal safeguards exist for American citizens detained in the War on Terror.[17] Furthermore, in *Rasul v. Bush*,[18] an opinion handed down the same day as *Hamdi*, the Court held 6 to 3 that the federal courts had jurisdiction to consider *noncitizen*– as well as citizen–detainees' habeas corpus petitions.[19] The Judicial Branch had made its voice

---

News/*Washington Post* poll found that nine out of ten Americans supported Bush's use of the military in Afghanistan).

10. *See* James Dao, *U.S. Is Taking War Captives to Cuba Base*, N.Y. TIMES, Jan. 11, 2002, at A1 (reporting that twenty Taliban and Al Qaeda operatives, the first of hundreds of such detainees, were being sent to Guantanamo Bay by the U.S. military).

11. Hamdi v. Rumsfeld, 542 U.S. 507, 510–11, 516–17 (2004) (plurality opinion).

12. 542 U.S. 507 (2004).

13. *Id.* at 517 (plurality opinion). Interestingly, Justice Thomas, writing alone in his dissenting opinion, also justified the President's detention authority under AUMF and not under Article II. *Id.* at 587 (Thomas, J., dissenting).

14. *Id.* at 529–33 (plurality opinion).

15. *Id.* at 533.

16. While the plurality in *Hamdi* made clear throughout most of the opinion that its application was only for citizen–detainees, some key segments, including one addressing potential procedural concerns for enemy-combatant proceedings, are less clear. *See id.* at 533–34 (outlining potential elements of a general "enemy-combatant proceeding").

17. *See id.* (mandating that citizen–detainees be allowed an enemy-combatant proceeding and suggesting the various ways in which Guantanamo procedures would need to be altered to accommodate the "exigencies of the circumstances").

18. 542 U.S. 466 (2004).

19. *Id.* at 485.

heard in the decision of how the United States formulates military-detention policies post-9/11.

Whether or not noncitizens possessed the right to challenge their detention through military proceedings, the Executive Branch was mindful of the Court's directives. In July 2004, Deputy Secretary of Defense Paul Wolfowitz issued an order establishing military tribunals that would allow each foreign-national detainee to "contest designation as an enemy combatant."[20] These Combatant Status Review Tribunals (CSRTs), as initially designed, were not subject to an external judicial review process,[21] and the procedural avenues available for the detainees to challenge their detention were limited. For instance, the detainee was denied access to counsel,[22] the detainee could call witnesses but only if the witnesses were "reasonably available" (a determination left to the discretion of the tribunal),[23] hearsay evidence was admissible,[24] and there was a rebuttable presumption in favor of the Government's evidence.[25] But despite these procedural limitations, Wolfowitz's Order, issued in the wake of *Hamdi* and *Rasul*, signaled that the Executive Branch was not interested in defying the legal interpretations, or as cynics might say, policies, issued by its judicial counterpart.

Meanwhile, on a separate front, the Executive Branch sought a remedy for the Court's decision in *Rasul* that detainees, pursuant to federal statute, could file habeas petitions in U.S. courts. The Executive found, or at least thought it found, a solution on Capitol Hill. At the urging of the President,[26] Congress stepped in to pass the Detainee Treatment Act (DTA) in 2005, which, among other things, amended federal law by stripping federal courts of their jurisdiction to hear habeas petitions from Guantanamo detainees.[27] In the context of the interbranch dialogue, the passage of the DTA is noteworthy in that two branches (here, the two "political" branches) combined resources to negate the decision of the third.

---

20. Memorandum from Paul Wolfowitz, Deputy Sec'y of Def., Order Establishing Combatant Status Review Tribunal 1 (July 7, 2004), *available at* http://www.defenselink.mil/news/Jul2004/d20040707review.pdf.

21. *See id.* at 4 (providing no mechanism for federal judicial review and emphasizing that the Order was "intended solely to improve management within the Department of Defense").

22. *See id.* at 1–2 (appointing to the detainee a military officer as a "personal representative," rather than actual legal counsel).

23. *Id.* at 2–3.

24. *Id.* at 3.

25. *Id.*

26. *See* BENJAMIN WITTES, LAW AND THE LONG WAR 144 (2008) (describing Congress's national-security legislative efforts post-9/11 as "patchy and erratic" initiatives in response to Executive requests rather than independent endeavors).

27. Detainee Treatment Act of 2005 (DTA), Pub. L. No. 109-148, secs. 1001–1006, 119 Stat. 2680, 2739–44 (codified at 10 U.S.C. § 801 note, 28 U.S.C. § 2241(e), and scattered sections of 42 U.S.C. § 2000dd) (amended 2006).

Congress did not empower the Executive with supreme authority over Guantanamo detention when it passed the DTA. In fact, the Act explicitly invokes the help of the Judiciary: Detainees, though precluded from applying for habeas relief in federal court, could appeal the CSRT's ruling on their status as enemy combatants to the D.C. Circuit.[28] As the majority in *Boumediene* later lamented, however, the DTA appeals process did not adequately provide for an independent judicial determination of whether the Government was justified in classifying the particular appellant as an enemy combatant.[29] Rather, the D.C. Circuit under the DTA was permitted only "to assess whether the CSRT complied with the 'standards and procedures specified by the Secretary of Defense' and whether those standards and procedures are lawful."[30]

But before the Supreme Court confronted the legitimacy of the CSRT judicial review process, it faced an issue in *Hamdan v. Rumsfeld*[31] that on the surface was far more facile: whether the DTA of 2005, specifically its jurisdiction-stripping provision, prohibited all Guantanamo habeas petitions (an interpretation that would have disposed of those petitions that were pending at the time the Act was passed) or only future petitions.[32] In a 5 to 3 decision,[33] the Court allowed pending habeas cases to proceed.[34]

Hamdan would not have presented such a key issue if several petitioners had not already filed habeas claims with the D.C. District Court. The post-*Rasul* habeas petitions had been consolidated into two separate proceedings in front of two different judges: Judge Leon and Judge Green.[35] Foreshadowing the diversity that would define the D.C. District Court's post-*Boumediene* habeas procedures, the judges' rulings were in opposition with respect to the petitioners' habeas rights.[36] While these D.C. District Court

28. *Id.* § 1005(e)(2).

29. *See* Boumediene v. Bush, 128 S. Ct. 2229, 2265 (2008) (contrasting the DTA with other federal habeas statutes to demonstrate that the DTA did not allow the reviewing court "to inquire into the legality of the detention").

30. *Id.* (quoting DTA sec. 1005(e)(2)(C), 119 Stat. 2742).

31. 548 U.S. 557 (2006).

32. *Id.* at 574–77.

33. *See id.* at 635 (noting that Chief Justice Roberts took no part in the decision).

34. *Id.* at 584.

35. *Boumediene*, 128 S. Ct. at 2241. Judge Green was designated pursuant to the district court's Executive Session Resolution to coordinate and manage all the habeas proceedings, but Judge Leon, foreshadowing his post-*Boumediene* course of action, resolved to keep his habeas petitions for separate adjudication. *In re* Guantanamo Detainee Cases, 355 F. Supp. 2d 443, 451–52 & n.14 (D.D.C. 2005).

36. *Compare* Khalid v. Bush, 355 F. Supp. 2d 311, 317 (D.D.C. 2005) (Leon, J.) (asserting that the petitioners had not presented "at least one viable legal theory, accepting the facts as they plead them to be true, under which this Court could issue a writ of habeas corpus challenging the legality of their detention"), *with In re* Guantanamo Detainee Cases, 355 F. Supp. 2d at 454 (Green, J.) (interpreting *Rasul* to indicate that Guantanamo detainees possess substantive constitutional rights in addition to the rights conferred by the federal habeas statute).

decisions were pending on appeal, Congress had enacted the DTA.[37] But since *Hamdan* held that the DTA's removal of habeas rights from detainees did not apply to habeas petitions pending at the time of enactment, the cases were set to proceed on appeal—unless Congress responded. Continuing the dialogic interbranch volley that had dictated Guantanamo detention policy since 2004, Congress did just that by passing the Military Commissions Act (MCA) in 2006.[38] Section 7 of the MCA aimed to essentially extinguish all habeas petitions, including those pending, from the purview of the federal courts.[39]

The 2007 MCA set the stage for a dynamic shift in the military-detention dialogue. As predicted, the D.C. Circuit held that the habeas petitions before the court, which had appeared as a result of *Rasul* and survived the 2005 DTA as a result of *Hamdan*, now fell outside the jurisdiction of the federal courts pursuant to the MCA.[40] The circuit court also determined that Section 7 of the MCA did not invoke a constitutional analysis of the Suspension Clause because noncitizen detainees at Guantanamo were not entitled to the privilege of the writ of habeas corpus or the protections of the Suspension Clause.[41] Thus, if Guantanamo detainees were to be equipped with habeas-petition opportunities, the directive would have to come from the Supreme Court.

If the Executive and Legislative Branches expected the Supreme Court to rubber stamp the denial of detainees' habeas rights through the MCA and the creation of the CSRT process, they were sorely disappointed on both counts. Five Justices coalesced in *Boumediene* to hold that noncitizen Guantanamo detainees *do* possess a constitutional privilege of habeas corpus.[42] On the one hand, the Court agreed with the D.C. Circuit that if MCA Section 7 was presumed to be valid, Congress's intent to abolish the federal courts' jurisdiction over Guantanamo habeas petitions must be respected.[43] But by determining that noncitizen detainees at Guantanamo possess rights under the U.S. Constitution, the Court held Section 7 of the MCA unconstitutional.[44] The Court reasoned that Section 7 violated the

37. *Boumediene*, 128 S. Ct. at 2241.

38. Military Commissions Act of 2006 (MCA), Pub. L. No. 109-366, 120 Stat. 2600 (codified at scattered sections of 10 U.S.C.).

39. *See id.* sec. 7(b) (stating that the habeas jurisdiction-stripping provision of the DTA "shall take effect on the date of the enactment of this Act, and shall apply to all cases, without exception, pending on or after the date of the enactment of this Act").

40. Boumediene v. Bush, 476 F.3d 981, 987–88 (D.C. Cir. 2007), *rev'd*, 128 S. Ct. 2229 (2008).

41. *Id.* at 988–91 (citing U.S. CONST. art. I, § 9, cl. 2).

42. Boumediene v. Bush, 128 S. Ct. 2229, 2239–40 (2008).

43. *Id.* at 2242–44.

44. *Id.* at 2274.

Suspension Clause because the CSRT process did not create an adequate alternative in place of the federal habeas corpus jurisdiction it removed.[45]

Pursuant to the DTA, the CSRT review process had left out the federal district courts altogether.[46] The D.C. Circuit was given exclusive jurisdiction to review (1) whether the CSRT's enemy-combatant-status determination was consistent with the Department of Defense's standards and procedures,[47] and (2) whether the use of such standards and procedures properly conformed to the Constitution and laws of the United States.[48] A critical feature of any habeas proceeding, the Court proclaimed, is the ability of the detainee to rebut the factual basis for the Government's assertion that he is an enemy combatant.[49] Included in this right is a mechanism for the habeas court to admit and consider "relevant exculpatory evidence that was not introduced during the earlier proceeding."[50] The *Boumediene* Court's tone and its analysis of habeas principles are reminiscent of *Hamdi* and *Rasul*, demonstrating the Judiciary's relatively consistent contributions to the Guantanamo-detainee dialogue.[51]

The *Boumediene* majority provoked much public furor.[52] Journalists and politicos either assailed the Court for ignoring precedent in a fit of judicial imperialism[53] and for making our country less safe,[54] or expressed support toward the Court for providing detainees with constitutional protections.[55] But when the Court's holding is appraised within the dialogic framework outlined above, we see that the overtures from both proponents and critics about *Boumediene*'s break from precedent are exaggerated.[56] Tracing the back-and-forth exchanges between the branches, from the enactment of the DTA after *Hamdi* and *Rasul* to the enactment of the MCA

45. Robert M. Chesney, Boumediene v. Bush. *128 S. Ct. 2229. United States Supreme Court, June 12, 2008*, 102 AM. J. INT'L L. 848, 850 (2008).

46. DTA, Pub. L. No. 109-148, sec. 1005(e)(2)(A), 119 Stat. 2680, 2742 (codified at 10 U.S.C. § 801 note) (amended 2006).

47. *Id.* sec. 1005(e)(2)(C)(i), 119 stat. 2742.

48. *Id.* sec. 1005(e)(2)(C)(ii), 119 stat. 2742.

49. *Boumediene*, 128 S. Ct. at 2269–70.

50. *Id.* at 2270.

51. *See supra* notes 12–19 and accompanying text.

52. *See* Chesney, *supra* note 45, at 851 (reporting the range of reactions).

53. *E.g.*, John Yoo, Editorial, *The Supreme Court Goes to War*, WALL ST. J., June 17, 2008, at A23 (lambasting the Court for ignoring precedent, denying the President and Congress their constitutional powers, and for deciding an issue not yet before it).

54. *E.g.*, Fred Thompson, Boumediene: *A Supremely Problematic Court Decision*, REALCLEARPOLITICS, June 23, 2008, http://www.realclearpolitics.com/articles/2008/06/boumediene_a_supremely_problem.html (arguing that *Boumediene* incentivizes violations of the Geneva Convention).

55. *E.g.*, Editorial, *Justice 5, Brutality 4*, N.Y. TIMES, June 13, 2008, at A28 (applauding the Court's affirmation of detainees' habeas rights).

56. Other scholars have also taken note of this trend. *See, e.g.*, Chesney, *supra* note 45, at 851 (observing that contrary to the public opinion, *Boumediene* was not a "sharp departure" from the status quo).

after *Hamdan*, the *Boumediene* Court's opinion, though noteworthy as the latest edition in the detention-policy power struggle, was anything but capricious. Regardless of whether one believes that *Boumediene* was recalcitrant, all can agree that United States' military-detention policy in the War on Terror has formed not as a unilateral, preconceived plan but as a product of a multitude of interbranch decisions.[57]

The implications of granting habeas rights to Guantanamo detainees may indeed be considerable and long-lasting from the perspective of constitutional scholars, but more pressing are the implications for the district judges who have been designated to oversee the habeas proceedings. As opposed to the system under the MCA, which wholly excluded the district court from the review process, *Boumediene* elevated the district courts to sit at the head of the table.[58] The D.C. District Court did not voluntarily position itself in the midst of the habeas commotion. Rather, its authority over the disposition of these petitions was essentially sealed by the Supreme Court's majority in *Boumediene*, who advocated for consolidating the habeas petitions within the D.C. District to "reduce administrative burdens on the Government" and to accommodate the Government's interest in "avoid[ing] the widespread dissemination of classified information."[59] But although the Supreme Court made clear that the D.C. District Court would be the venue of choice, its recommendations for how to arrange the proceedings were, to say the least, wanting.[60]

III. The Dialogue at the Ground Level: The Shaping of the D.C. District Court's Habeas Procedures

A. (Limited) Guidance from Above

As it embarked on its new mission, the D.C. District Court was not entirely without assistance; it had a patchwork of Supreme Court suggestions

---

57. *See* LAURA K. DONOHUE, THE COST OF COUNTERTERRORISM 71–72, 83 (2008) (describing the development of the United States' policy on the detention of foreign nationals at Guantanamo Bay following the attacks on the World Trade Center and the Pentagon as primarily a mixture of legislative and executive decisions and claiming that in spite of the courts' efforts to "push back" against these decisions, ultimately their role has been limited); Chesney, *supra* note 45, at 848 ("*Boumediene* addressed the aftermath of a series of judicial decisions, legislative enactments, and policy developments relating to military detention at Guantánamo."); *cf.* Daphne Barak-Erez, *Terrorism Law Between the Executive and Legislative Models*, 57 AM. J. COMP. L. 877, 890 (2009) ("[T]he detention of terrorist combatants in the United States was based only on presidential orders or, as explained by the U.S. Supreme Court, on presidential orders supported by the vague authorization included in the AUMF.").

58. *See* Boumediene v. Bush, 128 S. Ct. 2229, 2275 (2008) ("Our decision today holds only that . . . the petitioners in these cases need not exhaust the review procedures in the Court of Appeals before proceeding with their habeas actions in the District Court.").

59. *Id.* at 2276. The majority declared that the Government can move for change of venue for habeas petitions filed in district courts other than the D.C. District. *Id.*

60. *See, e.g.*, *id.* ("We make no attempt to anticipate all of the evidentiary and access-to-counsel issues that will arise during the course of the detainees' habeas corpus proceedings.").

it could use to point itself in the right direction. In fact, Judge Thomas Hogan, a prominent influence in the development of the habeas procedural framework, meticulously paired almost every procedural phase in his CMO with a statement from either a Supreme Court or D.C. Circuit ruling.[61]

It might seem odd that *Hamdi*, handed down four years before *Boumediene* and before the advent of the CSRT system, was the most useful precedent to guide the judges in shaping the habeas procedures. The *Hamdi* plurality pondered what a proceeding would look like for detainees challenging their enemy-combatant status.[62] A central tenet of a detainee's proceeding is the right for the detainee to "receive notice of the factual basis for his classification."[63] Noting the difficulties inherent in balancing national-security concerns with individual liberties, the Court suggested various alterations to the normal criminal-defendant procedural protections: "[h]earsay . . . may need to be accepted as the most reliable available evidence from the Government in such a proceeding,"[64] and "the Constitution would not be offended by a presumption in favor of the Government's evidence, so long as that presumption remained a rebuttable one and fair opportunity for rebuttal were provided."[65]

By contrast, the *Boumediene* Court's guideposts were more general: "[i]f a detainee can present reasonably available evidence demonstrating there is no basis for his continued detention, he must have the opportunity to present this evidence to a habeas corpus court";[66] "[t]he extent of the showing required of the Government in these cases is a matter to be determined";[67] and "[t]he habeas court must have sufficient authority to conduct a meaningful review of both the cause for detention and the Executive's power to detain."[68]

Additionally, the D.C. District was guided by earlier Supreme Court habeas precedents even though they were not directly applicable to the Guantanamo habeas procedures. For example, Judge Hogan's decision to allow Guantanamo petitioners to obtain limited discovery by request was made after considering *Bracy v. Gramley*,[69] which declares that habeas petitioners are not entitled to the type of discovery normally afforded civil

---

61. *See, e.g.*, Case Management Order at 1, *In re* Guantanamo Bay Detainee Litig., Misc. No. 08-0442 (D.D.C. Nov. 6, 2008) (Hogan, J.) [hereinafter *Guantanamo Bay* CMO], *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2002cv0828-397 (designing the proceedings to "provide the petitioners . . . with prompt habeas corpus review" (citing *Boumediene*, 128 S. Ct. at 2275), while also "'proceed[ing] with the caution' necessary in [the] context" (quoting Hamdi v. Rumsfeld, 542 U.S. 507, 539 (2004) (plurality opinion) (first alteration in the original))).

62. *Hamdi*, 542 U.S. at 533–34.

63. *Id.* at 533.

64. *Id.* at 533–34.

65. *Id.* at 534.

66. *Boumediene*, 128 S. Ct. at 2273.

67. *Id.* at 2271.

68. *Id.* at 2269.

69. 520 U.S. 899 (1997).

litigants in federal court,[70] and *Harris v. Nelson*,[71] which grants a habeas district court the power to require discovery "when essential to render a habeas corpus proceeding effective."[72] The district court has also carefully heeded the Supreme Court's reminders to use caution when dealing with facts and intelligence that invoke national-security concerns.[73]

B. Examining the D.C. District's Divergent Habeas Procedures

By July 2008, the D.C. District faced the momentous task of adjudicating the hundreds of habeas petitions filed by Guantanamo detainees. The primary vehicle used to establish procedures for these cases is the Case Management Order.[74] With the Supreme Court's edict that "[t]he detainees . . . are entitled to a prompt habeas corpus hearing"[75] still ringing in its ears, the court sprang into action by resolving to consolidate all current and future habeas cases in front of one person—Judge Hogan.[76] But it was immediately apparent that the fifteen district judges were not going to proceed as a singular unit. Judge Hogan acknowledged as much in his first order pertaining to the consolidated petitions: "Excluded from reassignment are all cases over which Judge Richard J. Leon currently presides . . . ."[77] On the same day, Leon issued an order cementing his decision to keep the eighteen habeas cases initially assigned to his docket.[78]

Pursuant to the Executive Session Resolution, Judge Hogan was responsible for going through all the habeas cases and "identify[ing] and delineat[ing] both procedural and substantive issues . . . common to all or some of the[] cases and, to the extent possible, rul[ing] on procedural issues

---

70. *Guantanamo Bay* CMO, *supra* note 61, at 3 (citing *Bracy*, 520 U.S. at 904).

71. 394 U.S. 286 (1969).

72. *Guantanamo Bay* CMO, *supra* note 61, at 3 (citing *Harris*, 394 U.S. at 300 n.7).

73. The D.C. District Court has demonstrated a devout commitment to balancing national-security concerns with the interests of conducting legitimate and transparent proceedings. This is evidenced by the plethora of safety precautions the court has instituted to govern the prehearing and hearing phases of the proceedings and by its willingness to separate out the unclassified information from each case so that it can be made available to the citizenry. *See* Case Management Order at 2, Boumediene v. Bush, No. 04-1166 (D.D.C. Aug. 27, 2008) (Leon, J.) [hereinafter *Boumediene* CMO], *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2004cv1166-142; *Guantanamo Bay* CMO, *supra* note 61, at 3 (both requiring the Government to file both classified and unclassified factual returns).

74. *See* Marc D. Falkoff, *Back to Basics: Habeas Corpus Procedures and Long-Term Executive Detention*, 86 DENV. U. L. REV. 961, 1019–21 (2009) (detailing the common procedures for Guantanamo habeas cases created by Judge Hogan's November 2008 Case Management Order and adopted by other members of the D.C. District Court).

75. Boumediene v. Bush, 128 S. Ct. 2229, 2275 (2008).

76. Order at 1–2, *In re* Guantanamo Bay Detainee Litig., Misc. No. 08-442 (D.D.C. July 2, 2008) (Hogan, J.), *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2002cv0828-337.

77. *Id.* at 2 n.1.

78. Order at 1–2, Khalid v. Bush, No. 04-1142 (D.D.C. July 2, 2008) (Leon, J.) [hereinafter *Khalid* Order], *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2004cv1142-93.

that [were] common to the cases."[79] Hogan explicitly stated that the cases would only be transferred for the purposes of "coordination and management" and that the transferring judges would still rule on the merits of each case to which they were originally assigned.[80]

In November 2008, Judge Hogan released his CMO, which was significant in two major respects.[81] First, several key procedures, both in the prehearing and hearing stages of the proceedings, differed from those set out by Judge Leon.[82] Furthermore, Judge Hogan expressly stated that the judges, after getting back their original habeas cases, were free to alter the procedural framework "based on the particular facts and circumstances of their individual cases."[83] By allowing this mechanism for alterations, Judge Hogan was in effect admitting the general CMO was merely a starting point or a suggestion the other judges could adopt or ignore. Though some judges received Judge Hogan's CMO and implemented it almost wholesale,[84] other judges took the liberty of making relatively significant modifications.[85]

But just two weeks after Judge Hogan released his original CMO, he decided to change it. Staying all the habeas petitions under his control, he contemplated the Government's Motion for Clarification and Reconsideration[86] and released an amended CMO on December 16, 2008.[87] Again, judges were left with the option of following Judge Hogan's CMO or making alterations.[88] Interestingly, at least one judge—Judge Bates—who

---

79. Order at 2, *In re* Guantanamo Bay Detainee Litig., Misc. No. 08-4444 (D.D.C. July 3, 2008) (Hogan, J.), *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2008mc04444-1.

80. *Id.*

81. *See Guantanamo Bay* CMO, *supra* note 61, at 2–4 (stipulating procedures for the disclosure of exculpatory evidence (requiring the disclosure of all "reasonably available" evidence) and discovery (allowing for a limited form of automatic discovery)).

82. *See infra* notes 93–97, 124–28 and accompanying text.

83. *Guantanamo Bay* CMO, *supra* note 61, at 2 n.1.

84. *See, e.g.*, Order at 1, Abdullah v. Bush, No. 05-23 (D.D.C. Nov. 19, 2008) (Roberts, J.) [hereinafter *Abdullah* Order], *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2005cv0023-19 (confirming Judge Roberts's decision to adopt Hogan's CMO for all substantive purposes); Order at 1–2, Hamlily v. Bush, No. 05-0763 (D.D.C. Nov. 6, 2008) (Bates, J.) [hereinafter *Hamlily* Order], *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2005cv0763-90 (declaring Judge Bates's decision to apply Hogan's CMO subject to minor changes in the briefing schedule).

85. *See, e.g.*, Case Management Order at 1, 3–7, Al-Adahi v. Bush, No. 05-280 (D.D.C. Nov. 13, 2008) (Kessler, J.) [hereinafter *Al-Adahi* CMO], *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2005cv0280-201 (adopting Hogan's CMO "in large part" but providing additional or altered language for several key procedures, including the disclosure of exculpatory evidence, discovery, the presumption in favor of the Government's evidence, and hearsay).

86. Order at 1–2, *In re* Guantanamo Bay Detainee Litig., Misc. No. 08-0442 (D.D.C. Nov. 21, 2008) (Hogan, J.) [hereinafter *Guantanamo Bay* Order], *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2005cv2380-107.

87. Order at 2–3, *In re* Guantanamo Bay Detainee Litig., Misc. No. 08-0442 (D.D.C. Dec. 16, 2008) (Hogan, J.) [hereinafter *Guantanamo Bay* Amended Order], *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2002cv0828-420.

88. *See id.* at 1–2 (asserting that the November 6 order was amended only by the specific provisions enumerated in the December 16 order, which kept the opt-out mechanism in force).

chose to adopt all of Judge Hogan's original CMO drifted from Judge Hogan's amended CMO insofar as the language governing discovery procedures.[89]

Before reaching the details of the CMOs, a cursory glance at the development of these habeas procedures shows, at the least, that this process is malleable. Judging by the proportion of judges who exercised their right to amend Judge Hogan's original CMO, Judge Leon and Judge Sullivan arguably took the more prudential course of action by immersing themselves in their cases from the beginning and crafting their own set of procedures. If the other judges were going to go through the same individualized process regardless, it stands to reason that delaying this process until Judge Hogan released his CMO only served to delay the petitioners' opportunity to present their cases in court.[90]

In addition, by developing CMOs in a judge-by-judge fashion, rather than creating a uniform CMO, the D.C. District invites criticism of the legitimacy of the habeas proceedings. Detainees whose petitions are denied might have a better case on appeal under this fragmented system than if the procedures were all the same. The fact that the procedures are *not* the same only bolsters the argument that the ground-level, district court shaping of habeas procedures comprises a smaller but also contentious dialogue that meaningfully affects U.S. military-detention policy.

We turn now to the two habeas procedures that have varied the most among the D.C. District Court judges' CMOs: the disclosure of exculpatory evidence and discovery.

1. Exculpatory Evidence.—The *Boumediene* Court stressed the importance of allowing Guantanamo detainees the opportunity to present "relevant exculpatory evidence that was not introduced during the earlier proceeding."[91] And from the Court's perspective, the CSRT review process lacked this feature.[92] The D.C. District Court has heeded *Boumediene*'s

89. *See* Case Management Order at 2–3, Hamlily v. Bush, No. 05-0763 (D.D.C. Dec. 22, 2008) (Bates, J.), *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2005cv0763-116 (modifying Hogan's amended CMO by adding language in the discovery procedures to require the Government to turn over evidence indicating that the petitioner was interrogated with coercive tactics). Some judges have only recently altered the language and organization of their CMOs. *See, e.g.*, Order at 1–5, Al-Mithali v. Obama, No. 05-cv-2186 (D.D.C. Jan. 14, 2010) (Huvelle, J.), *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2005cv2186-209 (amending her previous CMO to incorporate changes to discovery, exculpatory evidence, and other procedures).

90. The concern for unnecessarily delaying the habeas proceedings was realized as evidenced by the three habeas decisions (affecting seven detainees) Leon issued before the end of 2008. Al Alwi v. Bush, 593 F. Supp. 2d 24 (D.D.C. 2008); Sliti v. Bush, 592 F. Supp. 2d 46 (D.D.C. 2008); Boumediene v. Bush, 579 F. Supp. 2d 191 (D.D.C. 2008).

91. Boumediene v. Bush, 128 S. Ct. 2229, 2270 (2008).

92. *Id.* at 2272. Some scholars have observed that the D.C. Circuit in *Bismullah v. Gates*, 501 F.3d 178 (D.C. Cir. 2007), *reh'g denied*, 503 F.3d 137, *reh'g en banc denied*, 514 F.3d 1291 (2008), *cert. granted, judgment vacated*, 128 S. Ct. 2960 (2008) (vacated for further consideration in light of *Boumediene*), expansively interpreted the MCA to bestow upon the circuit court a fact-finding

directive by making the disclosure of exculpatory evidence a central component in the habeas prehearing procedures. The judges have disagreed, however, about which standard for disclosure should apply.

Judge Leon's CMO was released over two months before that of any other judge. Aiming to move efficiently and effectively through the habeas cases assigned to his docket, Judge Leon abstained from the consolidation process and immediately began holding hearings to determine the parties' statuses.[93] Moreover, in contrast to his robed colleagues, he has not made any substantive changes to the procedures laid out in his first CMO.[94]

Section I(E) of Judge Leon's CMO from August 27, 2008, requires the Government to disclose the following exculpatory evidence on an ongoing basis: "any evidence contained in the material reviewed in developing the return for the petitioner, and in preparation for the hearing for the petitioner, that tends materially to undermine the Government's theory as to the lawfulness of the petitioner's detention."[95]

The corollary to this procedural safeguard in the normal criminal law framework is commonly referred to as the "*Brady* rule," which imposes a duty on the prosecution to disclose information to the criminal defendant that is "material either to guilt or to punishment."[96] Judge Leon's version of the exculpatory-evidence disclosure requirement is notable for its limitations on the scope of the Government's burden. The Government is required to turn over evidence it finds only while creating the factual return or while preparing for the habeas hearing *for the individual petitioner*.[97] So if the Government's attorneys, while working on a petitioner's return or hearing, find exculpatory evidence for a different petitioner, they are free to keep it to themselves.

Judge Hogan's original CMO, released on November 6, 2008, constructed an exculpatory-evidence procedure that was missing Judge

---

function. *E.g.*, Chesney, *supra* note 45, at 851–52 ("The D.C. Circuit in *Bismullah v. Gates* recently held, for example, that DTA review would not be limited to the record of information actually presented to the CSRT as the government had urged. Instead, the court would consider all information in the government's possession that in theory should have been assembled for the CSRT, regardless of whether it actually was collected previously.").

93. *E.g.*, *Khalid* Order, *supra* note 78, at 1–2 (scheduling a status conference to discuss common issues among eighteen habeas cases).

94. *Compare* Case Management Order at 1–4, Al Shurfa v. Bush, No. 05-0431 (D.D.C. Nov. 28, 2008) (Leon, J.), *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2005cv0431-77 (ordering the Government to provide the court with a brief stating, at a minimum, the factual and legal basis for the detention and containing provisions for the presentation and confidentiality of evidence at the habeas hearing), *with Boumediene* CMO, *supra* note 73, at 1–4 (requiring the government to, at a minimum, establish the factual and legal basis for the detention and containing provisions for the presentation and confidentiality of evidence at the habeas hearing).

95. *Boumediene* CMO, *supra* note 73, at 2.

96. Strickler v. Greene, 527 U.S. 263, 280 (1999) (quoting Brady v. Maryland, 373 U.S. 83, 87 (1963)).

97. *Boumediene* CMO, *supra* note 73, at 2.

Leon's explicit limitations. Under Judge Hogan's version, the Government is required to disclose "all reasonably available evidence in its possession."[98] Although one might argue that Judge Hogan's original version impliedly limits the Government's burden in the same manner as Judge Leon's (and evidently, the Government did argue this),[99] the more logical reading seems to require disclosure of evidence the Government discovers in *any* materials, whatever the subject matter. If interpreted this way, Judge Hogan's original CMO imposed a more significant burden on the Government's attorney. An additional requirement dictated by Judge Hogan's original CMO was that the Government must file a "notice certifying either that it has disclosed the exculpatory evidence or that it does not possess any exculpatory evidence."[100]

Even though Judge Hogan's exculpatory-evidence disclosure requirements were more burdensome on the Government than Judge Leon's, some judges felt they still were inadequate. Judge Kessler issued her first CMO a week after Judge Hogan's.[101] Judge Kessler announced she was implementing the generic CMO "in large part,"[102] but her textual alterations suggested otherwise. For Judge Kessler, exculpatory evidence included all reasonably available evidence not only in the Government's possession, but also "that the Government can obtain through reasonable diligence."[103] The inclusion of the latter phrase creates an additional burden on the Government, but the extent of the burden is ambiguous.

Imagine a scenario where the Government is reviewing a CIA memo containing exculpatory evidence about a petitioner, say Petitioner *A*, while preparing for Petitioner *A*'s habeas hearing. The Government would be required to disclose this information to Petitioner *A* no matter which of the three judges mentioned above was assigned the case—Judge Leon,[104] Judge Hogan,[105] or Judge Kessler.[106] Now, imagine the same scenario in which the

98. *Guantanamo Bay* CMO, *supra* note 61, at 2–3.

99. *See Guantanamo Bay* Amended Order, *supra* note 87, at 2 (making clear that the phrase "reasonably available evidence" is not constrained to mean evidence found solely while reviewing materials for the petitioner).

100. *Guantanamo Bay* CMO, *supra* note 61, at 3.

101. *Al-Adahi* CMO, *supra* note 85, at 1.

102. *Id.* at 1.

103. *Id.* at 3.

104. Because Judge Leon's version of the *Brady* rule would require the Government to disclose only the exculpatory evidence against the individual petitioner, Judge Leon would require the disclosure of exculpatory evidence found against Petitioner *A* during her habeas corpus proceeding. *See supra* notes 95–97 and accompanying text.

105. Judge Hogan would require this disclosure because Judge Hogan's version of the *Brady* rule would require the Government to disclose the same evidence as Judge Leon's understanding, and, in addition, any other exculpatory evidence in its possession that was reasonably available. *See supra* notes 98–100 and accompanying text.

106. Judge Kessler would require the Government's disclosure of all exculpatory evidence under Judge Hogan's version of the *Brady* rule and additionally require disclosure of other

attorneys are reviewing the CIA memo. This time, though, they are preparing for Petitioner *B*'s hearing. Under this scenario the attorneys are arguably only obligated to disclose the information to Petitioner *A* under the standards set out by Judge Hogan and Judge Kessler.[107] So where does Judge Kessler's "reasonable diligence" standard come into play? Imagine this time that the Government attorneys are preparing for Petitioner *A*'s hearing, but the CIA memo bears no mention of Petitioner *A*. But then, one of the attorneys gets a call from his supervisor informing him that British intelligence just came across exculpatory evidence regarding Petitioner *A*. The information would not be in the Government's "possession" (if the Government defines "possession" as "actual possession"), so if the Government attorneys were arguing in front of Judge Hogan or Judge Leon, disclosure would not be required.[108] But does asking the British officials for the information constitute "reasonable diligence?" One could make the case that Judge Kessler's standard requires that the information described in all three of the above scenarios be disclosed.

This argument is circumstantially supported by the additional language Judge Kessler inserted into her exculpatory-evidence procedure. Pursuant to her CMO, the Government must, by a particular date, file an additional notice, which "notif[ies] each Petitioner of the existence of any evidence within its actual knowledge but not within its possession or capable of being obtained through reasonable diligence."[109] Thus, if in the last scenario described above, even if the government attorneys were unable to convince the British officials to hand over the information, they still are obligated to tell the Petitioner of the *existence* of the information (not necessarily the details of the information).

Lastly, Judge Kessler explicitly includes within her definition of exculpatory evidence "any evidence of abusive treatment, torture, mental incapacity, or physical incapacity which could affect the credibility and/or reliability of evidence being offered."[110] Taken altogether, Judge Kessler's exculpatory-evidence procedure, compared to Judge Hogan's, and especially compared to Judge Leon's, sets a much higher bar in terms of the time,

---

exculpatory evidence discoverable through reasonable diligence. *See supra* notes 101–03 and accompanying text.

107. *See supra* text accompanying notes 95–103. Because Judge Leon is the only judge of the three who understands the *Brady* rule as not requiring the Government to disclose exculpatory evidence found during the preparation for a hearing of another habeas petitioner, only Judge Leon would not require the Government in this hypothetical to disclose evidence that was exculpatory for Petitioner *A* if it was found during the Government's preparation for Petitioner *B*'s habeas hearing.

108. *See supra* text accompanying notes 95–100. Because Judge Leon's and Judge Hogan's versions of the *Brady* rule would require only the exculpatory evidence in the possession of the Government to be disclosed, evidence in the possession of the British government of which the U.S. Government was aware would not be in the possession of the U.S. Government and thus not subject to mandatory disclosure.

109. *Al-Adahi* CMO, *supra* note 85, at 1.

110. *Id.*

resources, and materials the Government is expected to contribute to the habeas proceedings.

Perhaps Judge Kessler was indirectly making the point that CMOs with greater specificity yield better results. If so, Judge Hogan was listening. Seven days after Judge Kessler's CMO, Judge Hogan stayed all habeas due dates and went to work revising his CMO.[111] In Judge Hogan's amended CMO, released on December 16, 2008, he clarified his original definition of "exculpatory evidence" but did not incorporate Judge Kessler's more stringent standard.[112] The revised language defines "reasonably available evidence" as "evidence contained in any information reviewed by attorneys preparing factual returns for all detainees; it is not limited to evidence discovered by the attorneys preparing the factual return for the petitioner."[113] By including this language Judge Hogan officially distinguished his procedure from Judge Leon's[114] and precluded the Government from arguing that as a practical matter both standards were the same.

Just days after Judge Hogan's revised CMO was released, Judge Huvelle released her first CMO, which both expanded and narrowed Hogan's new exculpatory-evidence standard. Huvelle expanded the standard by requiring the disclosure of reasonably available evidence not just in the Government's possession but also "within its actual knowledge."[115] This standard almost certainly imposes a more substantial burden on the Government than the "reasonable diligence" standard devised by Judge Kessler.[116] Recall the situation above where the Government's attorneys are aware of exculpatory information possessed by British officials. Judge Huvelle's version of the standard makes any discussion about the viability of obtaining the information from the British officials irrelevant: if the Government attorneys know about it, then the petitioner has a right to know as well.

Like Judge Kessler, Judge Huvelle more specifically defines "reasonably available evidence" but provides even more detail:

> [A]ny evidence or information that undercuts the reliability and/or credibility of the government's evidence (*i.e.*, such as evidence that

---

111. *See Guantanamo Bay* Order, *supra* note 86, at 1–2 (staying all due dates for habeas cases under his control pending resolution of the Motion for Clarification and Reconsideration of the CMO from November 6, 2008).

112. *Guantanamo Bay* Amended Order, *supra* note 87, at 2.

113. *Id.*

114. *See supra* text accompanying note 95.

115. Order at 2, Abdessalam v. Bush, No. 06-cv-1761 (D.D.C. Dec. 19, 2008) (Huvelle, J.) [hereinafter *Abdessalam* Order], *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2008mc0442-1365.

116. *Compare id.* (requiring the Government to disclose "all reasonably available evidence in its possession or any evidence within its actual knowledge that tends to materially undermine the evidence that the government intends to rely on"), *with Al-Adahi* CMO, *supra* note 85, at 1 (requiring disclosure of all evidence in the Government's possession that is reasonably available and all evidence "that the Government can obtain through reasonable diligence").

> casts doubt on a speaker's credibility, evidence that undermines the reliability of a witness's identification of the petitioner, or evidence that indicates a statement is unreliable because it is the product of abuse, torture, or mental or physical incapacity).[117]

Judge Huvelle's—like Judge Kessler's—concern for evidence tainted by possible coercion or torture is clearly indicated in her disclosure requirement.

Despite the broad nature of Judge Huvelle's original exculpatory-evidence procedure, it narrows, perhaps inadvertently, Judge Hogan's amended standard in one small but crucial way. Judge Huvelle left unclear whether the reasonably available evidence standard applied to information the Government came across for *all detainees*, or only the particular detainee.[118] Judge Huvelle's standard at first allowed the Government to make its own interpretation of the scope of application.[119] This debate has been made moot, however, by Judge Huvelle's order of January 9, 2009, which officially expanded the definition to include evidence the Government is reviewing for all other detainees.[120]

Looking at the CMOs chronologically, from Judge Leon's August 2008 CMO to Judge Huvelle's December 2008 CMO, the exculpatory-evidence standards have gradually become more burdensome for the Government. This is problematic for those who ascribe to the view that similarly situated petitioners should be subject to the same procedural standards.[121] This trend may be partially attributed to the system the district court devised for establishing the CMOs in the first place. Judges like Kessler and Huvelle were able to wait and see what Judge Hogan produced, and then, if dissatisfied (which they apparently were), they could craft their own standard. It is this consecutiveness that lends the procedure-formulating process to the dialogic framework. The judges' CMOs and amended CMOs constitute an intrabranch conversation regarding which standard for exculpatory evidence is most appropriate.

2. Discovery.—As the U.S. Supreme Court has observed, discovery in ordinary habeas proceedings is not a privilege petitioners enjoy "as a matter

---

117. *Abdessalam* Order, *supra* note 115, at 2.

118. *See id.* ("The government shall disclose . . . all reasonably available evidence in its possession or any evidence within its actual knowledge that tends to materially undermine the evidence that the government intends to rely on in its case-in-chief.").

119. *See id.* (omitting direction as to whether or not the Government must provide evidence relating to all detainees or only the specific detainee).

120. Order at 1, Al-Mithali v. Bush, No. 05-cv-2186 (D.D.C. Jan. 9, 2009) (Huvelle, J.), *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2005cv2186-138.

121. *See* Teague v. Lane, 489 U.S. 288, 290 (1989) ("A new rule will not be announced in a given case unless it would be applied retroactively to the defendant in that case and to all others similarly situated. This . . . avoids the inequity resulting from an uneven application of new rules to similarly situated defendants.").

of course."[122] Instead, the determination of whether a petitioner is entitled to invoke discovery for good cause shown is left to the discretion of the court.[123] Though all the judges on the D.C. district bench agree that Guantanamo petitioners should be allowed an opportunity for discovery in some circumstances, they are sharply divided over the type of discovery the petitioners should be offered.

Continuing the relatively conservative approach he displayed with his exculpatory-evidence procedures, Judge Leon provides Guantanamo petitioners limited discovery opportunities "for good cause shown" and refrains from placing excessive compliance burdens on the Government.[124] In order to obtain discovery at all, however, a petitioner must first submit to the court, in writing, the specific reasons for the request.[125] Furthermore, petitioners' requests must comply with the following rules:

> Any request for discovery must: (1) be narrowly tailored; (2) specify why the request is likely to produce evidence both relevant and material to the petitioner's case; (3) specify the nature of the request . . . ; and (4) explain why the burden on the Government to produce such evidence is neither unfairly disruptive nor unduly burdensome to the Government.[126]

Despite the plethora of CMOs that have been released since Judge Leon's *Boumediene* CMO, many of which pose alternatives to Leon's limited conception of habeas discovery, Judge Leon has stood by his original decision and has not made any substantive changes.[127]

A comparison of the discovery procedures from Judge Leon's and Judge Hogan's CMOs reveals that the judges possess divergent views on the discovery privileges to which the petitioners are entitled. In Section I(E)(1) of Judge Hogan's original CMO, the Government is *required* to disclose the following information to the petitioner upon a request from the petitioner:

> (1) any documents or objects in its possession that are referenced in the factual return; (2) all statements, in whatever form, made or adopted by the petitioner that relate to the information contained in the factual return; and (3) information about the circumstances in which such statements of the petitioner were made or adopted.[128]

---

122. Bracy v. Gramley, 520 U.S. 899, 904 (1997). The D.C. District Court has referenced this observation to support its discovery requirements. *Guantanamo Bay* CMO, *supra* note 61, at 3.

123. *Bracy*, 520 U.S. at 904.

124. *See Boumediene* CMO, *supra* note 73, at 2 (allowing petitioners discovery only after they have submitted detailed requests showing why the information will benefit their case).

125. *Id.*

126. *Id.*

127. *See supra* note 94.

128. *See Guantanamo Bay* CMO, *supra* note 61, at 3 (stipulating that the government "shall disclose" the listed types of information to the petitioner).

As distinguished from the exculpatory-evidence procedures, these discovery measures concern information, whether exculpatory or not, the court has determined should be available for all petitioners.

In light of Supreme Court precedent acknowledging that judges have the capacity to deny habeas petitioners all types of discovery,[129] it is in some ways remarkable that Judge Hogan (and other judges)[130] promulgated procedures that incorporate a form of "automatic" discovery. Or maybe automatic discovery is not just appropriate but necessary if the petitioners are to have, as Justice O'Connor stated in *Hamdi*, a "fair opportunity to rebut the Government's factual assertions."[131] Evidence may be lacking if the petitioners were captured on a battlefield in a foreign country, and petitioners may have limited access to witnesses and evidence. Consequently, the most probable explanation behind Judge Hogan's decision to incorporate an automatic discovery requirement is that the unmistakable burden it places on the Government is outweighed by the interest of conducting a fair proceeding according to *Boumediene*'s instructions.[132]

Under Judge Hogan's original CMO, however, petitioners are allowed discovery opportunities that even stretch beyond those provided automatically. In Section I.E.2, he stipulates that the judge, "for good cause," may "permit the petitioner to obtain limited discovery" beyond that described in the automatic-discovery provision.[133] Qualifying for this additional discovery is dependent on the discretion of the judge, and like Judge Leon's blanket discovery procedure, a petitioner seeking this discovery must submit a discovery request that complies with various specifications.[134] These specifications largely resemble those from Judge Leon's CMO, with a few minor changes. Instead of requiring the petitioner to explain why the requested information is likely to produce evidence "both relevant and material to the petitioner's case,"[135] Judge Hogan requires an explanation of why the information is likely to produce evidence "that demonstrates that the petitioner's detention is unlawful."[136] It is difficult to say which standard is tougher on the petitioner—demonstrating unlawful detention seems to require evidence that is both relevant and material—but it nevertheless reveals Judge Hogan's affirmative decision to change Judge Leon's wording.

---

129. Bracy v. Gramley, 520 U.S. 899, 904 (1997).

130. *See infra* notes 140–41 and accompanying text.

131. Hamdi v. Rumsfeld, 542 U.S. 507, 533 (2004).

132. *See* Boumediene v. Bush, 128 S. Ct. 2229, 2247 (2008) ("The [Suspension] Clause protects the rights of the detained by a means consistent with the essential design of the Constitution. It ensures that, except during periods of formal suspension, the Judiciary will have a time-tested device, the writ, to maintain the 'delicate balance of governance' that is itself the surest safeguard of liberty. . . . The Clause protects the rights of the detained by affirming the duty and authority of the Judiciary to call the jailer to account.").

133. *Guantanamo Bay* CMO, *supra* note 61, at 3.

134. *Id.* at 3–4.

135. *Boumediene* CMO, *supra* note 73, at 2.

136. *Guantanamo Bay* CMO, *supra* note 61, at 3.

It could be argued Judge Hogan's difference in language is inconsequential and that any change in meaning is inadvertent. But this assertion can be refuted for two reasons. First, by structuring his discovery procedures into a two-tiered, more pro-petitioner approach, Judge Hogan signaled he has thought about and discarded Judge Leon's discovery framework, meaning it is not a stretch to presume that any modification he made to Judge Leon's language is there because he intended it to be there. The second reason is the subtle manner in which his changes shift the tone of the procedure. Viewed in isolation, Judge Hogan's wording may not attract much attention. But take, for instance, his fourth rule governing a petitioner's request for additional discovery: "Discovery requests shall . . . (4) explain why the requested discovery will enable the petitioner to rebut the factual basis for his detention without unfairly disrupting or unduly burdening the government."[137] Rule number four's direct counterpart in Judge Leon's CMO requires that the request "explain why the burden on the Government to produce such evidence is neither unfairly disruptive nor unduly burdensome to the Government."[138] Judge Leon emphasizes "burden," making it the subject of the sentence; Judge Hogan replaces "burden" with "the requested discovery" and emphasizes the phrase "will enable the petitioner."[139] This quiet exchange between the judges is subtle, but significant.

Judge Bates adopted Judge Hogan's discovery language into his CMO on the same day Judge Hogan's original CMO was released,[140] and Judge Roberts followed suit thirteen days later.[141] Finding fewer disagreements with the generic discovery rules than with the exculpatory-evidence procedure, Judge Kessler implemented Judge Hogan's rules, adding only the phrase "whether coercive or not" in the automatic-discovery provision requiring the Government to disclose information about the circumstances surrounding any statements the petitioner made or adopted that relate to anything in his factual return.[142]

In his revised CMO on December 16, 2008, Judge Hogan kept most discovery provisions the same, but in a marked effort to reduce the

---

137. *Id.*

138. *Boumediene* CMO, *supra* note 73, at 2.

139. *Guantanamo Bay* CMO, *supra* note 61, at 3; *Boumediene* CMO, *supra* note 73, at 2.

140. *Hamlily* Order, *supra* note 84, at 2.

141. *Abdullah* Order, *supra* note 84, at 1. Roberts did slightly modify the Section I.E.2 by using the explicit phrase "*additional discovery*." *Id.*

142. *See Al-Adahi* CMO, *supra* note 85, at 1, 3–7 ("If requested by a Petitioner, the Government shall disclose to him: (1) any documents or objects in its possession that are referenced in the factual return; (2) all statements, in whatever form, made or adopted by the Petitioner that relate to the information contained in the factual return; and (3) information about the circumstances—whether coercive or not—in which such statements of that Petitioner were made or adopted.").

Government's automatic-discovery burden,[143] he reworded the first two requirements. The first was changed from "any documents or objects in its possession that are referenced in the factual return"[144] to "any documents and objects in the government's possession that the government relies on to justify detention."[145] Similarly, the second category of items was scaled back to require disclosure of "all statements . . . made . . . by the petitioner that the *government relies on to justify detention*."[146] After reviewing the Government's contention that the initial phrase was too burdensome, Judge Hogan decided to scale back the language of his CMO—and by consequence, other judges' CMOs—to place less of an obligation on the Government.[147]

After Judge Hogan's revised CMO was released, Judge Bates reconsidered his earlier wholesale adoption of Judge Hogan's discovery procedure. In addition to the previous requirement that the Government disclose "information about the circumstances" under which a petitioner made statements, under Judge Bates's revised discovery procedure, the Government must also disclose information that includes "but [is] not limited to any evidence of coercive techniques used during any interrogation or any inducements or promises [that] were made."[148] Judge Huvelle, staying true to form, paid more attention to the details, evidenced by her automatic-discovery disclosure requirements for statements the Government planned to use in its case-in-chief:

> [T]he Government . . . shall disclose (1) the identity of the speaker; (2) the content of the statement; (3) the person(s) to whom the statement was made; (4) the date and time the statement was made or adopted; and (5) circumstances under which such statement was made or adopted (including the location where the statement was made).[149]

The discovery procedures issued by the D.C. District Court have mimicked the trend of the exculpatory-evidence procedures in that they have grown increasingly more demanding in terms of the information they expect the Government to produce. For reasons stated above, this increasing

---

143. *See* Order at 2, Zaid v. Obama, No. 05-1646 (D.D.C. Feb. 9, 2009) (Bates, J.) [hereinafter *Zaid* Order], *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2005cv1646-146 ("To be sure, [Hogan's] amended order reduced [the Government's] automatic discovery burden under section I.E.1(2)—but not as much as respondents claim.").

144. *Guantanamo Bay* CMO, *supra* note 61, at 3.

145. *Guantanamo Bay* Amended Order, *supra* note 87, at 2; *see also supra* text accompanying note 128 (quoting the requirements from the initial CMO).

146. *Guantanamo Bay* Amended Order, *supra* note 87, at 2; *see also supra* text accompanying note 128 (quoting the requirements from the initial CMO).

147. *See Zaid* Order, *supra* note 143, at 2 (acknowledging the slightly scaled-back disclosure requirements in Judge Hogan's revised CMO).

148. Case Management Order at 2, Zaid v. Bush, No. 05-1646 (D.D.C. Dec. 22, 2008) (Bates, J.) [hereinafter *Zaid* CMO], *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2005cv1646-96.

149. *Abdessalam* Order, *supra* note 115, at 2.

demand may be a natural consequence of the system to which the judges initially subscribed.[150] This variance, though, carries potentially damaging systemic consequences.

IV. The Judges Are Talking, But Should They Be Leading the Conversation?

Contrasting Leon's steady, conservative CMO with Huvelle's evolving, pro-petitioner CMO demonstrates the amount of discretion available to district judges. The malleable nature of district court jurisprudence, especially when precedent is lacking or vague, may be criticized for its tendency to create uncertainty.[151] The D.C. District Court, however, has not stepped outside the bounds of its authority. After all, the *Boumediene* Court offered only broad generalizations for what it envisioned the habeas procedures would look like.[152] But discretionary determinations are arguably more susceptible to the influence of personal predilections.[153] And systemically, we might be especially concerned about such a high level of discretion when important national-security issues are at stake.

One must keep in mind, though, that the Judiciary's leadership role in crafting habeas procedures is not without limit; it could be checked quite easily by congressional action. Since *Boumediene*, Congress has conspicuously abstained. And the more variant the Judiciary's procedural experimentation, the more noticeable the Legislature's absence from this process. Shifting the habeas-procedure dialogue from an intrabranch to an interbranch discussion—with Congress at the head of the table—would ensure an institutionally balanced and uniform formulation of the procedures governing Guantanamo detainees' legal efforts.[154]

One could argue that the current intrabranch dialogue—fickle as it may be—is an appropriate approach to the shaping of habeas procedures. This may be true for three reasons. First, the judges are skilled in procedural matters. As overseers of the pretrial and trial process for all types of

150. *See supra* subpart III(A).

151. *See, e.g.*, Tonja Jacobi & Emerson H. Tiller, *Legal Doctrine and Political Control*, 23 J.L. ECON. & ORG. 326, 328 (2007) (opining that appellate court standards, as opposed to rules, "offer little guidance as to expected behavior, thus generating some costs associated with uncertainty").

152. *See, e.g.*, Boumediene v. Bush, 128 S. Ct. 2229, 2270 (2008) (declaring that habeas courts must have the "means to correct errors"); *id.* (asserting that the district court must have the authority to admit and consider "relevant exculpatory evidence that was not introduced during the earlier proceeding"); *id.* at 2275 (cautioning that the Judiciary should have sensible rules for staying habeas proceedings for domestic exigencies).

153. *See* Jacobi & Tiller, *supra* note 151, at 328 ("[I]ncreased discretion . . . makes standards more easily susceptible to value judgments and fact shading by lower courts and thus also may produce policy errors when the lower court has policy preferences different from the higher court.").

154. *See* Cynthia J. Bowling & Margaret R. Ferguson, *Divided Government, Interest Representation, and Policy Differences: Competing Explanations of Gridlock in the Fifty States*, 63 J. POL. 182, 183 (2001) (noting that overcoming interbranch rivalry can theoretically create a more cooperative and effective government).

litigants, the D.C. judges possess unique knowledge of the discovery rules and the rules of evidence, not to mention the federal rules of civil and criminal procedure.[155] In some ways their experience in this area seems to make them the natural choice to be the creators of the habeas procedures, especially since two of the biggest procedures at issue post-*Boumediene* are the disclosure of exculpatory evidence and the rules governing discovery.

A second reason the D.C. District Court's procedure-shaping authority might actually help advance the broader dialogue lies in the very manner in which the procedures have unfolded. The judge-by-judge process, a process the court has ensured contains maximum transparency, has revealed a diverse set of procedural options. The fact that the judges have diverged, and in some instances diverged considerably, at least shows the judges are shaping these procedures with their eyes open. These procedures were crafted neither simultaneously nor independently. A review of the CMOs issued by the D.C. District Court in the last nine months shows us that most judges' procedures were crafted only after scrutinizing the initial procedures set out in Hogan's general CMO of November 2008.[156] In essence, the judges are communicating their military-detention policy preferences for the Guantanamo habeas proceedings through their original and revised CMOs.

Last, the D.C. judges are not talking entirely amongst themselves. In some respects, the mini-dialogue at the ground level has been a mutual conversation between the Judiciary and the Executive. Each proceeding has commenced with an invitation by the court for each party to submit briefs arguing for the appropriate standards and rules to govern the process.[157] Only those in chambers know whether and to what extent the judges incorporated these views into their CMOs, but the opportunity for external input is present. Government attorneys, as respondents, have served as the Executive's voice on the procedure-developing process. But this voice inevitably plays a subordinate role and can quickly be muffled by the court, a

---

155. *See Nonjudicial Activities of Supreme Court Justices and Other Federal Judges: Hearings Before the Subcomm. on Separation of Powers of the S. Comm. on the Judiciary*, 91st Cong. 60 (1969) (statement of L. Ray Patterson, Professor of Law) (asserting that the Judiciary is especially qualified in making rules of procedure, principally because they are constantly monitoring how well rules of procedure are working).

156. The judges were free to make these adjustments, as Judge Hogan himself noted in his original CMO. *Guantanamo Bay* CMO, *supra* note 61, at 2 n.1.

157. *See, e.g.*, Order at 2–4, Batarfi v. Bush, Civ. No. 05-0409 (D.D.C. July 31, 2008) (Sullivan, J.), *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2005cv0409-92 (asking the parties to address various procedural issues, including evidentiary standards, burdens of proof, and the overall structure of the habeas proceeding); Briefing and Scheduling Order at 3–4, Boumediene v. Bush, No. 04-1166 (D.D.C. July 30, 2008) (Leon, J.), *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2005cv0573-65 (obligating the parties to file briefs addressing relevant standards of proof, discovery, evidentiary concerns, and other procedural matters); Scheduling Order at 3, *In re* Guantanamo Bay Detainee Litig., Misc. No. 08-442 (D.D.C. July 11, 2008) (Hogan, J.), *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2002cv0828-348 (requiring the parties to file briefs concerning the scope of discovery, evidentiary standards, burdens of proof, and other procedural matters).

point emphatically reinforced by several judges openly chastising, and even forcibly withdrawing, government attorneys who have not complied with the court's CMO.[158]

These potential benefits, when weighed against alternatives, do not ultimately justify a judicially dominated procedure-shaping system. Although the Judiciary's authoritative role in legal proceedings is an important one, and should not be unnecessarily hampered, it is problematic that since *Boumediene*, the most poignant government voice regarding the structure of the Guantanamo habeas proceedings has come from the federal district bench and not from our nation's law-making authority. To be clear, it's not as if the Judiciary has spoken *more loudly* vis-à-vis Congress on detention procedures and policy; it's that Congress has been virtually silent.[159] Perhaps, one could say, this arrangement passes muster at least in the sense it has happened before: the Judiciary has not shied away from intervening in Guantanamo issues (e.g., *Hamdi*, *Rasul*, and *Hamdan*), and Congress's initiatives in the aftermath of these opinions were arguably more a result of deference to the Executive than independent will.

But *Boumediene*'s call for the D.C. District Court to take the reins of the habeas petitions constitutes a qualitatively different type of judicial intervention. In addition to merely having jurisdiction over habeas proceedings, the district court has the task of generating the procedural standards that govern these proceedings. The judges in this context are functioning not just in their normal capacity as fact finders but also as policy makers. The significance of this type of decision-making process lies not only in the fact that the judges are applying facts to procedures they have themselves generated; which set of procedural standards a given petitioner is afforded depends on the judge to whom the petitioner is assigned. Moreover, as we have seen, the available array of procedures is quite heterogeneous.

Putting Congress in charge of this dialogue would not necessarily provide a more workable system for Guantanamo habeas petitioners. The institutional deficiencies of the Legislature indeed create concerns of their own. But congressional leadership in the form of targeted legislation would help create a more systemically balanced process. A balanced option like this is becoming increasingly attractive as the tragic attacks of 9/11 become

158. *See, e.g.*, Order Denying Respondents' Motion for Reconsideration at 12, Al Odah v. U.S., No. 02-828 (D.D.C. Apr. 6, 2009) (Kollar-Kotelly, J.), *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2002cv0828-529 (ordering an attorney to withdraw from the case for repeatedly evading the court's orders); Order to Show Cause at 1–3, Batarfi v. Bush, No. 05-0409 (D.D.C. Mar. 13, 2009) (Sullivan, J.), *available at* https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2005cv0409-170 (expressing Judge Sullivan's frustration with the Government's noncompliance with the court's exculpatory-evidence procedures and ordering the Government to explain why it should not be held in contempt).

159. WITTES, *supra* note 26, at 144.

more distant and the complex legal dilemmas surrounding anti-terrorism efforts become more prominent.[160]

First, congressionally created habeas procedures would instill uniformity while still allowing judges flexibility in adjudicating specific petitioners' claims. As it stands, some detainees may have a greater chance of gaining a favorable judgment on the merits if the judge's exculpatory-evidence procedures require more exertion on the part of the Government. Petitioners might prefer Judge Huvelle or Judge Kessler—rather than Judge Bates[161]—to adjudicate their petitions, since Huvelle's[162] and Kessler's[163] discovery and exculpatory-evidence disclosure standards are, at least on the surface, relatively more favorable to petitioners. But perhaps the differences are merely formal, not substantive, and the Government's response to habeas petitions is already uniform.[164] Even assuming this is true, having formally varied standards within the D.C. District Court is still problematic. Such a system breeds uncertainty and skepticism. It personifies a court in disagreement about the way in which suspected terrorists are treated in our legal system, a disagreement in need of resolution.

Second, because of the practical reality of the congressional lawmaking process, a shift from judge-made to statutory habeas procedures would give the Executive an opportunity to play a more influential role in the process. The Government's interest in shielding certain information in the interests of national security is a legitimate concern that must be a critical factor in the design of concrete habeas procedures.[165] Although the D.C. District Court judges have certainly been cognizant of this concern,[166] the reality is that the Executive can insert itself into the process in a much more meaningful way as a collaborator with Congress than merely as a respondent in a habeas proceeding.

Last, congressionally created habeas procedures would serve as a catalyst for the development of a more stable and consistent legal framework

---

160. *See id.* at 146 ("[I]t is, in fact, a system that we're building. Its parts are interconnected and affect one another. Legislators cannot ignore the often perverse incentive structures legal rules create between and among these parts.").

161. *See supra* notes 140, 148 and accompanying text.

162. *See supra* notes 115–17, 142 and accompanying text.

163. *See supra* notes 102–03, 142 and accompanying text.

164. *See* Robert M. Chesney, *National Security Fact Deference*, 95 VA. L. REV. 1361, 1366–85 (2009) (describing the constitutional, functional, and practical arguments the government typically sets forth in detention proceedings).

165. *See* Boumediene v. Bush, 128 S. Ct. 2229, 2296 (2008) (Scalia, J., dissenting) ("Henceforth, as today's opinion makes unnervingly clear, how to handle enemy prisoners in this war will ultimately lie with the branch that knows least about the national security concerns that the subject entails."). *But see* Falkoff, *supra* note 74, at 1021 ("The mere fact that all evidence in cases of this nature necessarily touches on issues of national security should not be an adequate justification [for keeping the evidence from the petitioners].").

166. *See, e.g.*, *Zaid* CMO, *supra* note 148, at 3 (requiring that petitioner's counsel be cleared to access classified information before requiring that the Government release such information to the lawyer).

for the United States' broader antiterrorism efforts. Back-and-forth jabs between the Supreme Court, Congress, and the President, the type of dialogue we have witnessed since 2004, has its advantages but is not conducive to establishing durable long-term policies. And if the United States wants to take a step toward solidifying some of the legal areas currently draped in ambiguity, the development of clear habeas procedures for Guantanamo detainees would be a feasible start to such an undertaking.

The district court's adjudicatory process, if this were to happen, would not be torn asunder. In fact, statutory procedures would enable the district court to dispose of habeas cases more efficiently, reducing the amount of resources judges expend on crafting and tweaking their CMO procedures. Certainty could also come in the form of an appellate decision, but while this would help clarify the current heterogeneous guidelines, the systemic benefits described above would not come to fruition and the Judiciary-dominated system would still be in place. Congressional leadership on this issue, or at least a system in which Congress and the Judiciary work in tandem,[167] offers a more balanced system for adjudicating habeas claims.

V. Conclusion

In the post-9/11 setting, the Judiciary is no stranger to the military-detention policy dialogue. The Supreme Court has stated that its involvement in this conversation strengthens, not weakens, the United States' ability to counteract terrorism because it ensures that our actions are within the bounds of the law.[168] In the past five years Americans have watched the three branches jockey for position for the final word on the scope of the Executive's authority for detaining terrorist suspects: *Hamdi* and *Rasul*;[169] the creation of the CSRT process;[170] the passage of the DTA;[171] *Hamdan*'s response to the DTA;[172] Congress's response to *Hamdan* via the MCA;[173] and now, *Boumediene*.[174]

The *Boumediene* Court proclaimed that in the United States, liberty and security are reconciled "within the framework of the law," and that "[t]he Framers decided that habeas corpus, a right of first importance, must be a

167. The process by which Congress promulgates these statutory habeas procedures could resemble the current system for other federal rules, such as the Federal Rules of Civil Procedure. This sort of arrangement comprises judicial committees that construct rules and then make recommendations to Congress. Administrative Office of the U.S. Courts, Federal Rulemaking: The Rulemaking Process, http://www.uscourts.gov/rules/newrules3.html.

168. Hamdan v. Rumsfeld, 548 U.S. 557, 636 (2006) (Breyer, J., concurring).

169. *See supra* text accompanying notes 12–19.

170. *See supra* text accompanying notes 20–25.

171. *See supra* text accompanying notes 26–30.

172. *See supra* text accompanying notes 31–37.

173. *See supra* text accompanying notes 38–41.

174. *See supra* text accompanying notes 42–50.

part of that framework, a part of that law."[175] We can be less certain of the framers' thoughts on whether the district court is the governmental entity best suited to develop the parameters of this framework. Nevertheless, following *Boumediene*, the D.C. District Court has relentlessly assumed this responsibility, and in doing so has demonstrated that the ground-level communication inside the district court can meaningfully impact the broader interbranch dialogue. It is time now for this dialogue to include the Legislative Branch. The Judiciary has made it clear that Guantanamo Bay detainees have a right to the benefits of U.S. law through habeas corpus; Congress must now help determine how that law is developed.

—Colin C. Pogge

175. Boumediene v. Bush, 128 S. Ct. 2229, 2277 (2008).

* * *

Child Pornography, the First Amendment, and Mistakes of Age: An Age-Old Debate*

I. Introduction

For over three decades, legislatures and courts have commented on the evils that sexual exploitation inflicts upon children. The Supreme Court of the United States summarized its concerns about the creation of child pornography in *New York v. Ferber*.[1] It opined that sexually exploited children tend to be unable to establish healthy romantic relationships; have sexual dysfunctions; abuse drugs and alcohol; engage in prostitution; and sexually abuse children during their adulthoods.[2] It also acknowledged the close relationship between child pornography and sexual molestation of child subjects and that where "such performances are recorded . . . the child's privacy interests are also invaded."[3] The existence of visual recordings and the likelihood that they have been distributed primarily for the predilections

---

* J.D. candidate, The University of Texas School of Law, 2010; B.A. Trinity University, 2007.

I thank Professor Allison Benesch and Amber McKeon for their helpful comments on this Note.

1. 458 U.S. 747 (1982).
2. *Id.* at 758 n.9.
3. *Id.*

of pedophiles, the Court explained, psychologically harms a child well into his or her adult life.[4] Despite the stiff penalties for producing child pornography, these offenses are the fastest growing type of crime prosecuted by U.S. Attorneys.[5]

In response to charges of producing child pornography, defendants have invoked the First Amendment to ward off conviction.[6] They have argued that state and federal statutes that require no knowledge as to the age of the child subject in a pornographic production are overly broad because they substantially chill protected speech.[7] Though at least one court—the Ninth Circuit Court of Appeals—was sympathetic to defendants' First Amendment challenges,[8] others that have subsequently evaluated the merits of these arguments have repeatedly rejected them.[9] In 2009, for example, the Fourth and Eighth Circuit Courts of Appeals rejected similar assertions that the First Amendment requires a mistake of age defense to a charge of producing child pornography.[10]

Because courts have continually cast off First Amendment defenses to these charges, the resultant strict-liability standard as to a defendant's knowledge of a child subject's age has some chilling effects on constitutionally protected speech.[11] Conversely, the Ninth Circuit's reasonable mistake of age defense fails to adequately protect the interests of children against the long-lasting physical and psychological effects of being photographed or filmed while engaging in sexually explicit acts.[12] As a compromise between these two approaches, this Note proposes an intermediate standard that would require defendants claiming a mistake of age defense to show that they verified child subjects' ages with government documents or officials. By establishing a clearer standard than the Ninth Circuit's reasonableness test and by providing some defense to defendants, this intermediate standard would protect children from the harms of child pornography and quell First Amendment concerns.[13]

---

4. *Id.* at 759 n.10.

5. MARK MOTIVANS & TRACEY KYCKELHAHN, BUREAU OF JUSTICE STATISTICS, FEDERAL PROSECUTION OF CHILD SEX EXPLOITATION OFFENDERS, 2006, at 2 (2007), http://bjs.ojp.usdoj.gov/index.cfm?ty=pbdetail&iid=886.

6. *See infra* subpart IV(B).

7. *See infra* subpart IV(B).

8. *See, e.g.*, United States v. U.S. Dist. Court, 858 F.2d 534, 547 (9th Cir. 1988) (agreeing with the defendant that the First Amendment required a reading of a reasonable mistake of age defense into a federal statute criminalizing the production of child pornography).

9. *See, e.g.*, United States v. Malloy, 568 F.3d 166, 177 (4th Cir. 2009) (holding that excluding a reasonable mistake of age defense does not infringe upon a defendant's rights); Gilmour v. Rogerson, 117 F.3d 368, 373 (8th Cir. 1997) (disagreeing with the defendant that the First Amendment requires a mistake of age defense to charges of producing child pornography).

10. *Malloy*, 568 F.3d at 177; United States v. Wilson, 565 F.3d 1059, 1069 (8th Cir. 2009).

11. *See infra* subpart V(B).

12. *See infra* subpart V(A).

13. *See infra* subpart V(C).

Part II of this Note surveys state and federal statutes that criminalize the creation of child pornography. After Part III outlines the overbreadth doctrine with regard to obscenity and child pornography, Part IV describes the strict-liability and reasonableness approaches to the crimes of producing child pornography, looking primarily at decisions from the Fourth, Eighth, and Ninth Circuit Courts of Appeals and the Courts of Appeals of Maryland and Minnesota. Part V then compares the arguments made by both sides of the mistake of age debate, and it offers and defends an intermediate standard that strikes a better balance of free speech and child-protection concerns. Part VI concludes this Note with a few summarizing remarks.

II. Child Pornography

A. State Statutes on the Sexual Exploitation of Children

Some states have criminalized the use of minors as pornographic subjects,[14] and since the mid-1990s, all states have prohibited visually depicting actual children engaged in sexual activity.[15] Statutes regarding this form of sexual exploitation of minors vary in three primary respects. First, they impose different levels of mental culpability that the state must prove when prosecuting violators.[16] Some states require proof that producers

---

14. Charles L. Simmons, Jr., Note, *Maryland's Child Pornography Statute Holds Photographers Strictly Liable for the Use of Under-Age Subjects but Leaves Open the Possibility of the Mistake of Age Defense*, 25 U. BALT. L. REV. 109, 109–10 (1995).

15. *See id.* at 109 n.1 (citing each state's particular statute governing the sexual exploitation of children via visual depictions).

16. ALA. CODE § 13A-12-197 (Supp. 2008); ALASKA STAT. § 11.41.455 (2008); ARIZ. REV. STAT. ANN. § 13-3552 (2009); ARK. CODE ANN. § 5-27-303 (2006); CAL. PENAL CODE § 311.3 (West 2008); COLO. REV. STAT. ANN. § 18-6-403 (West Supp. 2009); CONN. GEN. STAT. ANN. § 53a-196a (West 2007); DEL. CODE ANN. tit. 11, § 1108 (2007); D.C. CODE § 22-2012 (2001); FLA. STAT. ANN. § 827.071 (West Supp. 2010); GA. CODE ANN. § 16-12-100 (2007); HAW. REV. STAT. ANN. § 707-750 (LexisNexis 2007); IDAHO CODE ANN. § 18-1506 (Supp. 2009); 720 ILL. COMP. STAT. ANN. 5/11-20.1 (West Supp. 2009); IND. CODE § 35-49-3-2 (2004); IOWA CODE ANN. § 728.12 (West Supp. 2009); KAN. STAT. ANN. § 21-3516 (2007); KY. REV. STAT. ANN. §§ 531.320–.370 (LexisNexis 2008); LA. REV. STAT. ANN. § 14:81.1 (Supp. 2010); ME. REV. STAT. ANN. tit. 17-A, § 282 (Supp. 2007); MD. CODE ANN., CRIM. LAW § 11-207 (LexisNexis 2002); MASS. GEN. LAWS ANN. ch. 272, § 29A (West 2000); MICH. COMP. LAWS ANN. § 750.145c (West Supp. 2009); MINN. STAT. ANN. § 617.246 (West 2009); MISS. CODE ANN. § 97-5-33 (Supp. 2009); MO. ANN. STAT. § 568.060 (West 1999); MONT. CODE ANN. § 45-5-625 (2008); NEB. REV. STAT. §§ 28-1463.01 to .05 (2008); NEV. REV. STAT. §§ 200.700–.760 (2007); N.H. REV. STAT. ANN. § 650:2 (2007); N.J. STAT. ANN. § 2C:24-4(b) (West 2005); N.M. STAT. ANN. § 30-6A-3 (LexisNexis Supp. 2009); N.Y. PENAL LAW § 263.05 (McKinney 2008); N.C. GEN. STAT. § 14-190.6 (2007); N.D. CENT. CODE §§ 12.1-27.2-02 to 04 (1997); OHIO REV. CODE ANN. § 2907.321 (West 2006); OKLA. STAT. ANN. tit. 21, § 1021.2 (West Supp. 2010); OR. REV. STAT. § 163.670 (2007); 18 PA. CONS. STAT. ANN. § 6312 (West Supp. 2009); R.I. GEN. LAWS § 11-9-1.3 (Supp. 2009); S.C. CODE ANN. § 16-15-335 (Supp. 2009); S.D. CODIFIED LAWS § 22-22-24.3 (2006); TENN. CODE ANN. § 39-17-902 (2006); TEX. PENAL CODE ANN. § 43.25 (Vernon Supp. 2009); UTAH CODE ANN. § 76-5a-3 (Supp. 2009); VT. STAT. ANN. tit. 13, § 2822 (Supp. 2009); VA. CODE ANN. § 18.2-374.1 (2009); WASH. REV. CODE ANN. § 9.68A.040 (West 2003); W. VA. CODE ANN. §§ 61-8C-2 to 3 (LexisNexis 2005); WIS. STAT. § 948.05 (Supp. 2009); WYO. STAT. ANN. § 6-4-303 (Supp. 2009).

intended to visually depict minors engaged in sexual acts or knew, or should have known, the age of the child involved.[17] Others impose strict liability on defendants if prosecutors merely prove the child was a minor.[18] Second, they vary in what age ranges constitute "children" or "minors."[19] And third, states disagree on whether a defendant's mistake of a child's age amounts to a defensible position to prosecutions.[20] In perceiving a dearth of effective deterrents to the use of children in pornography,[21] Congress also prohibited the production of child pornography in the 1970s.[22]

B. Federal Criminalization of the Production of Child Pornography

In 1977, Congress enacted the Protection of Children Against Sexual Exploitation Act,[23] which provided a federal mechanism for the prosecution of child exploitation through sexually exploitative live performances and visual depictions of children engaged or engaging in sexual conduct.[24] Senators sponsored the bill to protect children against the negative mental health consequences of being used in live shows and paraded around in the nude.[25] The sponsors also argued that sexual exploitation through visual materials increased the risk that children, still in their formative stages of development,

---

17. *E.g.*, ALA. CODE § 13A-12-197(a) (Supp. 2006); IND. CODE § 35-49-3-2.

18. *E.g.*, CONN. GEN. STAT. § 53a-196a(a).

19. *Compare, e.g.*, ALASKA STAT. § 11.41.455 (requiring the child subject to be under the age of eighteen), *with, e.g.*, IND. CODE § 35-49-3-2 (requiring the child subject to be under the age of sixteen).

20. *Compare* ALA. CODE § 13A-12-197(a) (providing a reasonable mistake of age defense as interpreted by the Court of Criminal Appeals of Alabama in Sherman v. State, 778 So. 2d 859, 861 (Ala. Crim. App. 2000)), MD. CODE ANN., CRIM. LAW § 11-207 (allowing a reasonable mistake of age defense after the Court of Appeals of Maryland's interpretation of the predecessor statute in Outmezguine v. State, 641 A.2d 870 (Md. 1994)), *and* VT. STAT. ANN. tit. 13, § 2822(b) (Supp. 2009) (providing explicitly that a reasonable mistake of age constitutes an affirmative defense to criminal prosecution under the statute), *with* OHIO REV. CODE ANN. § 2907.321(B)(2) (providing expressly that mistake of age is not a defense).

21. Sponsors of federal legislation on child pornography thought the current laws were insufficient to curb production of child pornography. 123 CONG. REC. 33,045 (1977).

22. Audrey Rogers, *Protecting Children on the Internet: Mission Impossible?*, 61 BAYLOR L. REV. 323, 326 (2009).

23. Pub. L. No. 95-225, sec. 2(a), 92 Stat. 7 (1978). As of 1982, only twelve states prohibited the use of minors to make pornography. New York v. Ferber, 458 U.S. 747, 749 n.2 (1982).

24. 123 CONG. REC. 33,045. The constitutionality of the Act was a primary concern of the sponsors:

> [M]ost importantly, [this legislation] is clearly constitutional. S. 1585 can be upheld by the courts and serve as the basis of successful prosecutions resulting in the reduction of sexual exploitation of children.
>
> . . . .
>
> . . . We must be ever watchful that in our efforts to control the most offensive pornography we do not infringe on these important constitutional rights. . . .
>
> . . . .
>
> I do not want to get into an area that is unconstitutional.

Id. at 33,045–46 (Statement of Sen. Bayh).

25. *Id.* at 33,046.

would be more likely to engage in prostitution and less likely to form "healthy, affectionate relationships" as adults.[26] They sought to deter exploitation by criminalizing the production of certain visual representations of nude children and "arous[ing] [a] collective conscience" that would foster policies more protective of children.[27]

Congress amended the statute with the Child Protection Act of 1984.[28] With that legislation, it found that the child pornography industry "had developed into a highly-organized, multi-million-dollar industry which operates on a nationwide scale" and that thousands of children were used in productions.[29] Consistent with the contentions of the supporters of the Protection of Children Against Sexual Exploitation Act,[30] Congress also found that "the use of children as subjects of pornographic materials is harmful to the physiological, emotional, and mental health of the individual child and to society."[31] In conjunction with these declarations, Congress enhanced the penalties for violations and—with potential First Amendment implications in mind—confined the statute's coverage from "visual or print medium" depicting children engaging in sexually explicit conduct[32] to "visual depiction[s]" of such activity.[33]

In its current form, the federal child pornography statute proscribes persuading or pressuring a minor to partake in sexually explicit activities for the purpose of producing a visual recording of that conduct.[34] Those convicted of producing visual depictions of actual children engaged in sexually

26. *Id.* at 33,047.

27. *Id.* at 33,046–47; 33,050; 33,057.

28. Pub. L. No. 98-292, 98 Stat. 204.

29. *Id.* sec. 2(1)–(2).

30. *See supra* text accompanying notes 25–27.

31. Sec. 2(3), 98 Stat. at 204.

32. *Id.* sec. 3(1)–(2).

33. *Compare* Protection of Children Against Sexual Exploitation Act of 1977, Pub. L. No. 95-225, 92 Stat. 7 (1978) ("Any person who violates this section shall be fined not more than $10,000, or imprisoned not more than 10 years, or both, but, if such person has a prior conviction under this section, such person shall be fined not more than $15,000, or imprisoned not less than two years nor more than 15 years, or both."), *with* § 3(4)–(5), 98 Stat. at 204 ("Section 2251 of Title 18 of the United States Code is amended by striking out '$10,000' and inserting '$100,000' in lieu thereof; [and] by striking out '$15,000' and inserting '$200,000' in lieu thereof."). Since 1984, the amendments to the federal statute have either been minor changes in the wording of the statute for clarity or in anticipation of Commerce Clause challenges to the statute. *See Timeline of Significant Events in Child Pornography Legislation*, THIRD BRANCH, Dec. 2009, http://www.uscourts.gov/ttb/2009-12/article06a.cfm (displaying the relevant child pornography legislation enacted since 1977); *see also* Mitchell P. Goldstein, *Congress and the Courts Battle over the First Amendment: Can the Law Really Protect Children from Pornography on the Internet?*, 21 J. MARSHALL J. COMPUTER & INFO. L. 141, 151–55 (2003) (chronicling the developments in federal child pornography laws). In 1986, Congress further amended the statute to criminalize the transportation of minors in interstate or foreign commerce for the purpose of having the child appear in pornography. Child Sexual Abuse and Pornography Act of 1986, Pub. L. No. 99-628, sec. 3, 100 Stat. 3510, 3510 (codified at 18 U.S.C. § 2251 (2006)).

34. 18 U.S.C. § 2251(a).

explicit conduct face at least fifteen years in prison for a first offense and could serve a life sentence if they have two prior convictions.[35] Though it requires proof that defendants must have "know[ledge] or . . . reason to know that such visual depiction will be transported or transmitted"[36] in interstate commerce, the statute does not require that a defendant have knowledge—or even reason to believe—that a child subject has yet to turn eighteen years of age.[37] Thus, the statute facially imposes strict liability on producers of visual depictions of children engaging in sexually explicit conduct.

In responding to federal charges of producing child pornography, defendants have voiced a variety of constitutional objections to federal prosecution of the sexual exploitation of children. They have claimed that Congress lacked constitutional authorization to enact the statute,[38] and that such charges violated the First Amendment,[39] the Equal Protection Clause,[40] the substantive due process guarantees of the Fifth Amendment,[41] and constitutional requirements that a *mens rea* element be proven in criminal cases.[42] Though case law has developed relatively uniformly with regard to most of these constitutional challenges,[43] federal circuits and states have split on whether the Free Speech Clause of the First Amendment requires a mistake of age defense to federal and state crimes of producing child pornography.[44] Last year's decisions by the Fourth[45] and Eighth[46] Circuits considering the First Amendment's application to the federal child pornography statute invite a reexamination of this split among federal and state courts.

III. The First Amendment

In the two centuries following the adoption of the First Amendment, the Supreme Court significantly stretched the scope of constitutionally protected

35. *Id.* § 2251(e).

36. *Id.* § 2251(a).

37. United States v. U.S. Dist. Court, 858 F.2d 534, 536 (9th Cir. 1988); *see also* 18 U.S.C. § 2256(1) (defining "minor" for the purposes of § 2251(a) as a person under the age of eighteen years).

38. *E.g.*, United States v. Malloy, 568 F.3d 166, 171 (4th Cir. 2009).

39. *E.g.*, United States v. Kantor, 677 F. Supp. 1421, 1423 (C.D. Cal. 1987).

40. *E.g.*, United States v. Freeman, 808 F.2d 1290, 1292–93 (8th Cir. 1987).

41. *E.g.*, United States v. Bach, 400 F.3d 622, 628 (8th Cir. 2005).

42. *E.g.*, United States v. Esch, 832 F.2d 531, 536 (10th Cir. 1987).

43. For example, on the Commerce Clause issue circuits have agreed that Congress had the constitutional authority to pass the Protection of Children Against Sexual Abuse Act of 1977. United States v. McCalla, 545 F.3d 750, 753–56 (9th Cir. 2008); United States v. Blum, 534 F.3d 608, 609–10 (7th Cir. 2008); United States v. Griffith, 284 F.3d 338, 345–48 (2d Cir. 2002); United States v. Buculei, 262 F.3d 322, 328–30 (4th Cir. 2001); United States v. Hampton, 260 F.3d 832, 834–35 (8th Cir. 2001); United States v. Galo, 239 F.3d 572, 575–76 (3d Cir. 2001).

44. *See infra* subpart IV(B).

45. United States v. Malloy, 568 F.3d 166 (4th Cir. 2009).

46. United States v. Wilson, 565 F.3d 1059 (8th Cir. 2009).

speech. The First Amendment stipulates that "Congress shall make no law . . . abridging the freedom of speech."[47] The framers of the Free Speech Clause supported the amendment as an essential element of promoting the free exchange of ideas in public debate—which lay at the heart of a democratic system of government[48]—emphasizing the protection of minority views.[49] To further facilitate the exchange of political ideas, the Supreme Court has expansively interpreted the word "speech" to include other expressive modes of communication, such as conduct that can be understood as conveying ideas.[50] In tandem with this recognition, the Court also acknowledged additional societal benefits of protecting the free flow of speech and speech-like conduct, including artistic activity.[51] It eventually recognized that the First Amendment protected artistic expressions including photography and film.[52]

A. Obscenity as Unprotected "Speech"

Consistent with its expansion of expressive conduct protected by the First Amendment, the Court has also narrowed categories of unprotected speech, such as obscenity. Courts in the United States originally used the *Hicklin*[53] rule from English common law as the obscenity test.[54] Under this rule, a legislature could prohibit materials that it found to "deprave and corrupt those whose minds are open to such immoral influences and into whose hands a publication of this sort may fall."[55] The Supreme Court explicitly rejected this rule in *Roth v. United States* as inconsistent with the First Amendment,[56] and offered a more specific test in *Miller v. California*.[57]

---

47. U.S. CONST. amend. I. The Supreme Court has incorporated the free speech protections of the First Amendment against the states. Gitlow v. New York, 268 U.S. 652, 666 (1925).

48. *See* DAVID M. RABBAN, FREE SPEECH IN ITS FORGOTTEN YEARS, 1870–1920, at 190 (1999) (stating that the American framers intended to overturn what they saw as "antirepublican" English speech laws).

49. *See id.* (asserting that the framers wanted greater speech rights than the "antirepublican" English common law provided).

50. *See, e.g.*, Texas v. Johnson, 491 U.S. 397, 406 (1989) (holding burning the American flag to be expressive conduct); Cohen v. California, 403 U.S. 15, 26 (1971) (holding that wearing a jacket with the words "Fuck the Draft" on it was constitutionally protected speech); Tinker v. Des Moines Indep. Cmty. Sch. Dist., 393 U.S. 503, 505–07 (1969) (holding that wearing of black armbands to protest the Vietnam War was protected speech).

51. *See, e.g.*, Ward v. Rock Against Racism, 491 U.S. 781, 790 (1989) ("Music, as a form of expression and communication, is protected under the First Amendment."); Joseph Burstyn, Inc. v. Wilson, 343 U.S. 495, 502 (1952) ("For the foregoing reasons, we conclude that expression by means of motion pictures is included within the free speech and free press guaranty of the First and Fourteenth Amendments.").

52. *Burnstyn*, 343 U.S. at 502.

53. Regina v. Hicklin, [1868] 3 L.R.Q.B. 360, 371.

54. Roth v. United States, 354 U.S. 476, 488–89 (1957).

55. *Hicklin*, 3 L.R.Q.B. at 371.

56. *Roth*, 354 U.S. at 489.

57. 413 U.S. 15, 24 (1973).

In addition to limiting the definition of obscenity to material that specifically depicts or describes sexual conduct, the *Miller* Court required a finding that the work as a whole appeals to prurient interest in sex (applying contemporary community standards), that the conduct is depicted in a patently offensive way, and that the "work, taken as a whole, lacks serious literary, artistic, political, or scientific value."[58] By providing a clearer standard, the *Miller* rule provided greater First Amendment protection to sexually explicit materials than did the *Hicklin* rule.

Even though some pornography may be considered not obscene under *Miller*, the Supreme Court has declared that sexually graphic depictions involving minors are constitutionally more permissible subjects of regulation than other genres of pornography.[59] In *New York v. Ferber*, the Court declared depictions of actual children engaged in sexual acts to be outside of the scope of the First Amendment[60] because they have de minimis social value,[61] are intrinsically related to child abuse,[62] and psychologically and emotionally harm children.[63] In clarifying the parameters of "child pornography," the *Ferber* Court explicitly separated such material from obscenity as defined by the *Miller* requirements.[64] The Court compared the two tests and held that, unlike obscenity, the test for child pornography did not require the material be considered as a whole, appeal to prurient interests, or portray sexual conduct in a "patently offensive manner."[65] Furthermore, the Court held that whether a work contains "serious literary, artistic, political, or scientific value" is irrelevant as to whether the material is considered child pornography.[66] The *Ferber* standard considers material to be child pornography—and outside the scope of the First Amendment—if it merely portrays, even in part, minors engaged in sexually explicit conduct, so long as prohibitory statutes specifically identify the acts that may not be depicted.[67]

---

58. *Id.*

59. New York v. Ferber, 458 U.S. 747, 756 (1982).

60. *Id.* at 763.

61. *Id.* at 762.

62. *Id.* at 759.

63. *Id.* at 759 n.10. The Court also noted that holding child pornography to be outside the scope of the First Amendment was consistent with its previous decisions. *Id.* at 763.

64. *Id.* at 764.

65. *Id.*

66. *Id.* at 761. *Ferber*'s standard is much more akin to the *Hicklin* rule because it merely requires depiction of sexually explicit conduct. However, the *Hicklin* rule and the *Ferber* rule may be distinguished because *Ferber* requires that "sexual conduct" be statutorily limited and described. *Id.* at 764. *Hicklin* provided no guidance in making content-based distinctions. *See supra* note 55 and accompanying text.

67. *Ferber*, 458 U.S. at 764.

B. The Overbreadth Doctrine

On top of precluding laws proscribing protected speech, the First Amendment also forbids laws that are overly broad in their application so as to inhibit speakers from engaging in protected speech activities. In deciding an overbreadth challenge to a criminal statute, a court must balance the challenged statute's statutory aims with free speech interests.[68] As the Supreme Court declared in *Broadrick v. Oklahoma*,[69] "the First Amendment needs breathing space,"[70] and legislation deterring individuals from engaging in constitutionally protected speech may consequently be considered unconstitutionally overbroad.[71] Overbreadth challenges to statutes criminalizing expressive conduct—as opposed to speech—must show that the statute's chilling of protected speech "not only be real, but substantial as well, judged in relation to the statute's plainly legitimate sweep."[72] Thus, courts considering overbreadth challenges to legislation criminalizing conduct must balance the "plainly legitimate sweep" of the legislation against the harms caused by inhibiting protected speech.[73]

Because *Miller* considers a broad class of adult pornography to be constitutionally protected speech,[74] and because *Ferber* considers production of any pornography employing child subjects to be unprotected speech,[75] the age of a subject participating in sexually explicit conduct for the purpose of producing pornography can determine whether the First Amendment offers full protection to those visual depictions or provides none whatsoever.[76] The fine line drawn by the Court in its *Miller* and *Ferber* decisions raises overbreadth concerns for laws prohibiting the production of child pornography,[77] as they might discourage individuals from engaging in constitutionally protected speech activities.[78] Given this fine line, defendants have challenged strict liability imposed by child pornography statutes as violating

68. *See* Virginia v. Hicks, 539 U.S. 113, 119–20 (2003) (describing the point at which the chilling effect on the First Amendment might outweigh the otherwise legitimate statutory aims of an anti-trespass law); *see also* Gideon Newmark, Comment, *The Strong Medicine of Overbreadth as Applied to Criminal Libel*, 59 CASE W. RES. L. REV. 553, 560–61 (2009) (explaining the Supreme Court's overbreadth doctrine and applying it to criminal libel statutes).

69. 413 U.S. 601 (1973).

70. *Id.* at 611.

71. *Id.* at 612–15.

72. *Id.* at 615.

73. *Id.*

74. *See supra* text accompanying notes 56–58.

75. *See supra* notes 59–67 and accompanying text.

76. United States v. U.S. Dist. Court, 858 F.2d 534, 538–39 (9th Cir. 1988).

77. *Cf. id.* ("[N]o one claims that [the material here] is obscene; the film would therefore enjoy the protection of the first amendment were it not for its depiction of a minor. . . . The age of the subject thus defines the boundary between speech that is constitutionally protected and speech that is not.").

78. *Id.* at 545. However, as the court goes on to note, current jurisprudence places a great deal of weight on the side of the state's interest in protecting children from exploitation, which diminishes the effect of overbreadth concerns. *Id.*

the First Amendment because it fails to provide "breathing space" for mistakes of a child subject's age.[79]

IV. Mistakes of Age

A. The Production-Versus-Distribution Distinction

The Supreme Court has not directly addressed whether the First Amendment requires a reasonable mistake of age defense to crimes of producing child pornography. In *United States v. X-Citement Video*,[80] it held that the federal child pornography statute requires prosecutors to show that distributors and retailers of child pornography had knowledge that material they sold contained sexually explicit images of minors.[81] In *X-Citement Video*, a defendant was indicted for selling and shipping multiple pornographic videotapes that contained a subject who was under the age of eighteen.[82] The defendant appealed, arguing that because the statute lacked a knowledge of age requirement, it was unconstitutionally overbroad.[83] The Supreme Court agreed that the statute would have free speech implications if it lacked such a requirement but construed the statute to include a knowledge of age requirement to avoid a constitutional collision with the First Amendment.[84] Because it did not specifically address the reasonable mistake of age defense for *producers* of child pornography, the Court left this issue open for lower courts to decide.[85]

To date, the Courts of Appeals of the Fourth, Eighth, and Ninth Circuits, and the Courts of Appeals of Maryland and Minnesota have addressed this exact issue.[86] Only the Ninth Circuit has resolved that the First Amendment requires a reasonable mistake of age defense.[87] Others have opted for a strict-liability standard. This Part argues that because strict-liability jurisdictions lack strong precedential support for the proposition that chilled speech is insubstantial, their position pits free speech rights against the interests of protecting the welfare of children. However, it also contends that the

79. *See infra* subpart IV(B).

80. 513 U.S. 64 (1994).

81. *Id.* at 65–66.

82. *Id.* at 66.

83. *Id.* at 67.

84. *Id.* at 71–73. The Court found that for distributors—unlike producers—"[t]he opportunity for reasonable mistake as to age increases significantly once the victim is reduced to a visual depiction, unavailable for questioning by the distributor or receiver." *Id.* at 72 n.2.

85. *See* Gilmour v. Rogerson, 117 F.3d 368, 372 (8th Cir. 1997) (accusing the Supreme Court of intentionally sidestepping the issue of whether the Constitution *requires* a knowledge of age element).

86. United States v. Malloy, 568 F.3d 166, 174–76 (4th Cir. 2009); United States v. Wilson, 565 F.3d 1059, 1067–69 (8th Cir. 2009); United States v. U.S. Dist. Court, 858 F.2d 534, 540–41 (9th Cir. 1988); Outmezguine v. State, 641 A.2d 870, 875–77 (Md. 1994); State v. Fan, 445 N.W.2d 243, 245–48 (Minn. Ct. App. 1989).

87. *See infra* subpart IV(B).

Ninth Circuit's approach may protect too much free speech at the expense of child-protection goals. After analyzing both sides' arguments, this Part proposes an intermediate standard that helps resolve the tension between strict liability and the reasonableness test, which requires that a defendant show "by clear and convincing evidence, that he did not know, and could not reasonably have learned, that the actor or actress was under 18 years of age."[88]

B. The Split in Authority

1. The Ninth Circuit.—The Ninth Circuit Court of Appeals was the first federal appellate court to address this issue[89] in its decision in *United States v. United States District Court for the Central District of California*.[90] In the underlying case, *United States v. Kantor*,[91] producers were charged under the federal child pornography statute for creating a film showing a sixteen-year-old girl engaging in sexually explicit conduct.[92] The defendants argued only that they reasonably mistook the child for an adult because of her fraudulent representations of her age.[93] They contended that the statute was overly broad because it failed to include a reasonable mistake of age defense.[94] When the District Court of Central California sided with the defendants, the government petitioned the Ninth Circuit for a writ of mandamus.[95]

The Ninth Circuit initially noted that the pornographic material at issue in the case was not obscene and that but for its employment of a minor, it would have been constitutionally protected speech.[96] In its review of relevant Supreme Court decisions, the Ninth Circuit deduced that a "speaker may not be put at complete peril in distinguishing between protected and unprotected speech. Otherwise, he could only be certain of avoiding liability by holding his tongue . . . ."[97] The court agreed with the defendant that a minor can look like an adult and vice versa,[98] and that no source of identifying a potential subject's age is infallible.[99] Instead of invalidating the entirety of

88. *Dist. Court*, 858 F.2d at 543.

89. *See supra* note 86.

90. 858 F.2d 534 (9th Cir. 1988).

91. 677 F. Supp. 1421 (C.D. Cal. 1987). The facts of this case are the same as the *Central District of California* case; the style of the case changed upon the government's petition for writ of mandamus. *Id.* at 1426–29.

92. *Dist. Court*, 858 F.2d at 536.

93. *Id.*

94. *Id.*

95. *Id.*

96. *Id.* at 538.

97. *Id.* at 539 (internal citations and quotation marks omitted).

98. *Id.* at 539–40.

99. *Id.* at 540.

the statute, the court read a reasonable mistake of age defense into the statute "to avoid [its] constitutional infirmity."[100]

In his dissent, Judge Beezer disagreed that the First Amendment required a mistake of age defense because society's interest in protecting children outweighed a minimal intrusion on free speech.[101] He contended that the federal child pornography statute was intended to protect children—even those who fraudulently convinced others that they were adults—from their own immaturity.[102] Judge Beezer would have imposed the burden on pornographers to take all steps "necessary to establish the age of the subjects they depict—or [to] employ different subjects."[103] He continued to rebut the majority's arguments by contending that pornographers can conduct accurate investigations "based on reputation, first-hand testimony, and especially, documents"—even if this investigation required viewing the original documents in person.[104] He also reiterated the point from *X-Citement Video* that producers are in a much better position than distributors to ascertain the age of the subjects of pornography.[105]

2. The Eighth Circuit.—After *District Court*, the Eighth Circuit was the next federal appellate court to address the issue,[106] which it took up in *Gilmour v. Rogerson*.[107] In *Gilmour*, the defendant took sexually explicit photographs of a seventeen-year-old girl after being informed by both the girl's boyfriend and driver's license that she was twenty-two years old.[108] The defendant appealed his jury conviction under an Iowa statute criminalizing the creation of child pornography[109] on the grounds that it violated the First Amendment.[110] In finding a mistake of age defense to be wholly inconsistent with child-protection principles—particularly that a child should be protected even from his or her own deceitful conduct[111]—the Eighth Circuit rejected the defendant's contention on at least three grounds.[112] First, the concern for chilling speech of distributors did not apply because the Iowa statute only imposed strict liability on producers, who were akin to statutory

100. *Id.* at 542.
101. *Id.* at 544 (Beezer, J., dissenting).
102. *Id.*
103. *Id.*
104. *Id.* at 546.
105. United States v. X-Citement Video, 513 U.S. 64, 76 n.5 (1994); *Dist. Court*, 858 F.2d at 547.
106. *See supra* note 86 and accompanying text.
107. 117 F.3d 368 (8th Cir. 1997).
108. *Id.* at 369.
109. *Id.* at 370.
110. *Id.* at 370–72.
111. *Id.* at 372.
112. *Id.*

rapists who received no mistake of age defense.[113] Second, the speech that would be chilled—adult pornography—was not highly valuable speech under the First Amendment.[114] Finally, the statute was not substantially overbroad under *Broadrick*, according to the court, since it targeted only speech plus conduct rather than pure speech.[115]

Judge Arnold dissented. He pointed out that there was no support for a conclusion that a mistake of age defense would encourage a minor to be increasingly deceitful about his or her age and thereby result in more production of child pornography.[116] He also argued that the analogy of a child pornography producer to a statutory rapist was inapposite because the latter had no constitutional right to engage in sex with anyone except his or her spouse, whereas the former had a right to take erotic photographs.[117] Judge Arnold concluded that strict liability substantially burdens protected speech because of the severity of the penalties for the crime.[118]

In 2009, the Eighth Circuit Court of Appeals reaffirmed its position in *Gilmour*. In *United States v. Wilson*,[119] the court applied its rationale about the First Amendment's application from *Gilmour* to the federal child pornography statute. In *Wilson*, Devin Wilson was convicted under the federal child pornography statute[120] for videotaping his casual sexual encounters with a sixteen-year-old girl.[121] The court restated its position in *Gilmour* in holding that a mistake of age was no defense to producing child pornography.[122] It incorporated the *X-Citement Video* logic, which Justice Beezer emphasized in his *District Court* dissent,[123] that a producer of child pornography was less deserving of First Amendment protection than a distributor because a producer is in a better position to verify an actor's age.[124] To support this proposition, the court also reiterated *Gilmour*'s analogy that a producer is like a statutory rapist, for whom a mistake of age

113. *Id.* at 372–73. The Eighth Circuit analogized a producer of child pornography to a statutory rapist, who may be held strictly liable as to the age of a minor with whom the statutory rapist has sexual intercourse, because a child pornographer and a statutory rapist both have personal contact with the minor. "In this information age, a prudent photographer or movie producer may readily and independently confirm the age of virtually every young-looking model." *Id.* (citing Outmezguine v. State, 641 A.2d 870 (Md. 1994)). In the Eighth Circuit's view, this similarity supported the denial of a reasonable mistake of age defense.

114. *Id.* at 373.

115. *Id.* at 372.

116. *Id.* at 374 (Arnold, J., dissenting).

117. *Id.*

118. *Id.* at 375.

119. 565 F.3d 1059 (8th Cir. 2009).

120. *Id.* at 1062.

121. *Id.* at 1063.

122. *Id.* at 1069.

123. *See* United States v. U.S. Dist. Court, 858 F.2d 534, 547 (9th Cir. 1988) (Beezer, J., dissenting) ("A producer is in the perfect position to establish the age of his subject; he may ask the subject directly where to obtain documents verifying the subject's age.").

124. *Wilson*, 565 F.3d at 1067.

defense is unavailable.[125] The court held that concerns of fraud perpetrated by a minor would only deter an insubstantial amount of protected speech, which did not outweigh the child-protection goals of the statute.[126]

3. The Fourth Circuit.—In *United States v. Malloy*,[127] Michael Malloy, a police officer who had sex with a fourteen-year-old student and videotaped the encounter, was charged under the federal child pornography statute.[128] Like the defendants in *District Court* and *Gilmour*, Malloy argued that the First Amendment required a mistake of age defense.[129] The court disagreed, holding that the government's interest in preventing the sexual exploitation of minors was significant and that there was no substantial chilling of protected speech.[130] The Fourth Circuit adopted *Gilmour*'s child-protection rationale—that proscriptions on child pornography aim to protect even those children that make self-destructive decisions, such as lying about their ages.[131] It reinforced this reasoning by highlighting the long-term physical and psychological harms to children described in *Ferber* and in congressional debate over the child pornography statute.[132]

The court also held that strict liability as to a child subject's age would not result in substantial self-censorship for four reasons.[133] First, it cited federal law that requires pornographers to verify an actor's age[134] by examining "an identification document" that "contain[s] . . . the performer's name and date of birth."[135] Second, the only type of pornography that would be chilled would be a small subset of pornography that depicted a "youthful-looking" actor.[136] Third, prosecution under the statute would be rare where the subjects were not unmistakably children because prosecutors must frequently prove a child's age by only showing the video or photos since they cannot always locate the actor.[137] Fourth, in light of a slight chance of prosecution, the significant profits from selling pornography would not inhibit much speech.[138] Considering these four points, the Fourth Circuit held that the

125. *Id.* at 1067–68; *see also supra* note 113.

126. *Wilson*, 565 F.3d at 1069. The Court of Appeals ruled on essentially the same grounds in *United States v. Pliego*, 578 F.3d 938 (8th Cir. 2009). The lower courts of the Eighth Circuit have consistently followed this ruling. *E.g.*, United States v. Heath, No. CR09-2003-LRR, 2009 U.S. Dist. LEXIS 37994, at *6–11 (N.D. Iowa May 1, 2009).

127. 568 F.3d 166 (4th Cir. 2009).

128. *Id.* at 169.

129. *Id.* at 171.

130. *Id.* at 176.

131. *Id.* at 175.

132. *Id.* at 175–76.

133. *Id.*

134. *Id.* at 175.

135. *Id.* (citing 18 U.S.C. § 2257(b)(1) (2006)).

136. *Id.* at 175–76.

137. *Id.* at 176.

138. *Id.*

First Amendment required no mistake of age defense and upheld Malloy's conviction.[139]

4. Minnesota and Maryland.—In addition to these federal courts, two states' courts of appeals have also considered whether the First Amendment requires a mistake of age defense to their own laws prohibiting the production of child pornography. In *State v. Fan*,[140] the Court of Appeals of Minnesota confronted a First Amendment overbreadth challenge after a jury convicted David Fan—who had hired a thirteen-year-old girl to dance nude on stage—under a state statute outlawing the use of minors in live sexual performances.[141] This statute expressly precluded a reasonable mistake of age defense.[142] After noting that an overbreadth challenge must show that a substantial amount of protected speech would be inhibited,[143] the court rejected Fan's defense, stating, "Although it is marginally possible that the statute could reach a valid first amendment application, the statute does not substantially prohibit constitutionally protected expression."[144]

The Court of Appeals of Maryland followed *Fan* in its 1994 decision in *Outmezguine v. State*.[145] Elan Outmezguine was convicted under a Maryland statute of taking nude photographs of a fifteen-year-old, Jessica H.[146] Though Outmezguine was in the home-cleaning-and-improvement business, he made money on the side by photographing nude dancers and models.[147] Outmezguine gave Jessica H. and her boyfriend some of his cleaning and home-improvement work so they could make some money.[148] After they asked to see Outmezguine's photography, Outmezguine offered Jessica H. (and she accepted) money for him to take nude photos of her.[149] Jessica H. later revealed this to a counselor who reported it to the police, who in turn charged Outmezguine under the Maryland child pornography statute.[150] Though Outmezeguine claimed that he did not know Jessica H.'s age, she testified that she told him she was fifteen and in high school.[151] After the lower courts rejected Outmezguine's First Amendment challenge,[152] the Court of Appeals of Maryland affirmed.[153]

---

139. *Id.*
140. 445 N.W.2d 243 (Minn. Ct. App. 1989).
141. *Id.* at 244–45.
142. *Id.* at 245–46.
143. *Id.*
144. *Id.* at 246.
145. 641 A.2d 870 (Md. 1994).
146. *Id.* at 872.
147. *Id.*
148. *Id.*
149. *Id.* at 872–73.
150. *Id.* at 873.
151. *Id.* at 873–74.
152. *Id.* at 874–75.
153. *Id.* at 880.

The *Outmezguine* court sought to strike the proper *Broadrick* balance between preventing the chilling of protected speech and achieving the statutory goals of protecting children.[154] It determined that because a photographer can require subjects "to produce a birth certificate, driver's license, or similar governmental identification card,"[155] protected speech would not be chilled when "[a] reasonable bona fide attempt to verify the authenticity of such documents will thus ensure that the subject being used is not a child."[156] It further noted that the value of the chilled speech was marginal in light of the strong historical protection the First Amendment contemplated for political speech.[157] In addressing the substantiality of the chilled speech, the court stated in a conclusory manner, "Certainly, the chilling effect would not be considered 'substantial' as is required under an overbreadth analysis."[158]

V. The Age-Old Debate

While strict liability for producing child pornography provides insufficient breathing space for speech protected by the First Amendment, the Ninth Circuit's leniency provides too much space, which compromises child-protection goals. With an intermediate standard for a reasonable mistake of age defense that requires producers of child pornography to establish that they verified actors' proof of age with original government documents, the First Amendment concerns of strict liability for child pornography would be avoided while still providing robust protection for the interests of children.

A. Interests of Child Pornography Statutes

Though the government's interest in protecting children from the harms of child pornography is quite compelling, a mistake of age defense would still facilitate this interest. Those arguing that the First Amendment requires no such defense have determined that the statutory aims of protecting children—even from their own deceitful behavior—outweigh the protected speech that would be chilled by strict liability.[159] They have relied on legislative debate and *Ferber*'s articulation of the harms produced by child pornography, particularly those relating to difficulties forming relationships, more incidents of sexual molestation, and increased risks of drug and alcohol

154. *See id.* at 878 (recognizing that a statute that regulates speech "could possibly have a chilling effect on protected expression if criminality is imposed without a requirement that the defendant have knowledge of the child's minority" and balancing the right to freedom of expression against the government's right to protect children from sexual exploitation).

155. *Id.*

156. *Id.*

157. *Id.* at 878–79.

158. *Id.* at 879.

159. *See supra* subpart IV(B).

abuse.[160] Neither the Ninth Circuit in *District Court* nor Judge Arnold in his dissent in *Gilmour* attempted to rebut the harmfulness of child pornography.[161] Instead, Judge Arnold argued that engrafting a mistake of age defense into the federal child pornography statute would not result in an increase in the production of child pornography.[162] The court's ruling against a mistake of age defense similarly did not contend that this defense would deter child pornography production.[163]

The prosecutions under state and federal child pornography statutes have suggested that those charged with child pornography would not be benefited in many circumstances by a mistake of age defense.[164] Indeed, the Ninth Circuit's test is a difficult standard to meet. First, defendants must establish the defense by clear and convincing evidence.[165] Second, they must show that they not only lacked actual knowledge of the child's age but also could not have reasonably learned of the child's age.[166] Thus, as Judge Arnold argued in *Gilmour*, a mistake of age defense would not necessarily result in increased production of child pornography,[167] particularly because the stiffness of the penalties under the federal statute would still act as a sizeable deterrent.

Moreover, many defendants in the previously discussed cases would not have been able to effectively satisfy these requirements. For example, the defendants in *Wilson*, *Malloy*, *Fan*, and *Outmezguine*, who did not request to see government-issued proof of the photographed or video-recorded child's age,[168] probably could not have proven that their failures to do so constituted a reasonable attempt to learn of the child's age. The only case in which a defendant has benefited from a reasonable mistake of age defense was *District Court*, where the defendants relied on government documents provided by the child subject.[169] Even though the Ninth Circuit's reasonable mistake of age defense would not address the concern articulated by the *Gilmour* and *Malloy* courts that children be protected from their own deceit,[170] the case for excluding the defense is weaker than courts have purported because the other child-protection interests of child pornography statutes can still be actualized with the defense.

---

160. New York v. Ferber, 458 U.S. 747, 758 n.9 (1982).

161. *See supra* sections IV(B)(1)–(2).

162. Gilmour v. Rogerson, 117 F.3d 368, 374 (8th Cir. 1997) (Arnold, J., dissenting).

163. *See supra* subpart IV(B).

164. *See supra* subpart IV(B).

165. United States v. U.S. Dist. Court, 858 F.2d 534, 543 (9th Cir. 1988).

166. *Id.*

167. *Gilmour*, 117 F.3d at 374 (Arnold, J., dissenting).

168. *See supra* sections IV(B)(2)–(4).

169. The defendant in *Gilmour* who had allegedly seen a driver's license presented by a seventeen-year-old girl may have also benefited from a reasonable mistake of age defense, *Gilmour*, 117 F.3d at 369, but this would have turned on whether the jury found Gilmour's testimony to be clear and convincing, *Dist. Court*, 858 F.2d at 543–44.

170. *See supra* section IV(B)(2).

B. *Substantial Chilling of Protected Speech*

In recognizing a reasonable mistake of age defense, the *District Court* court argued that some protected speech would inevitably be chilled.[171] Since the defendant's guilt necessarily turns on the age of a child, those contemplating photographing or videotaping others engaged in sexual conduct would decide against it both because children can look older than their age and no type of record or proof of age is completely infallible.[172] The massive penalties under child pornography statutes would substantially inhibit protected speech when adult subjects look younger than their age, and when adult subjects look their actual age but the pornographer fears the adult subject looked young enough to be a minor.[173] Though some of the arguments responding to *District Court* have provided significant counterweights to the Ninth Circuit's reasoning, many others have missed the mark.

For example, the *Malloy* court held that strict liability as to knowledge of the child's age for the crime of producing child pornography would not inhibit a substantial amount of protected speech because strict liability would only apply to a subset of pornography where a producer sought to employ "youthful-looking" actors.[174] *Malloy*'s reasoning is problematic because it relies on no authority that chilled speech is insubstantial because it merely comprises a subcategory of a type of speech. Moreover, this argument fails to grapple with *District Court*'s position that children can look more adult-like than some adults and vice versa, which implies that strict liability chills speech in other categories where the producers of the pornography are *not* employing particular actors *because* they look of barely legal age.[175] Consequently, strict liability may inhibit protected speech in multiple genres of adult pornography.

The Fourth Circuit also reasoned that where the actor in question looked like an adult, producers of pornography would not be deterred by the small risk of prosecution because of the potential profitability of the pornographic material and the fungibility of pornographic actors.[176] This reasoning is also troublesome because it contemplates only commercial, mass producers of pornography who hire actors and sell their productions where many non-

171. *Dist. Court*, 858 F.2d at 540.

172. *Id.*

173. *See supra* text accompanying notes 96–98.

174. *See supra* text accompanying note 136.

175. Although "barely legal" is a subcategory of adult pornography, M. Eric Christensen, Note, *Ensuring that Only Adults "Go Wild" on the Web: The Internet and Section 2257's Age-Verification and Record-Keeping Requirements*, 23 BYU J. PUB. L. 143, 155–56 (2008), there are many other different categories of adult pornography that may use youthful-looking actors. *See, e.g.*, Michael Flood, *Child Porn Indicative of Culture's Teen Fetish*, CANBERRA TIMES (Austl.), Oct. 14, 2004, at A19 (listing types of pornography that often portray women as children, including "barely legal," "teenerama," "seventeen," "schoolgirls," and "teeny vision").

176. United States v. Malloy, 568 F.3d 166, 176 (4th Cir. 2009).

commercial producers of child pornography have been charged.[177] It also under assesses the risk of prosecution by failing to take into account the prosecutorial provinces of both state and federal officials, and by assuming that prosecutors will not have access to the child in question a vast majority of the time.[178] However, this is not always the case. Given Judge Arnold's observation about the stiffness of the penalties under federal and state child pornography statutes,[179] noncommercial production might be significantly inhibited because there is little to no financial incentive for production or readily-accessible substitutes.

Because consensual sex is not constitutionally protected speech,[180] the statutory-rape analogy[181] does not translate to the First Amendment context. Opponents of the defense argue that producers of child pornography are in a better position than distributors to take reasonable steps to validate the age of their photographed or videotaped subjects, as could a statutory rapist.[182] Producers, as the Eighth Circuit noted in *Gilmour* and *Wilson*, are thus more akin to statutory rapists, who receive no reasonable mistake of age defense. Because documents are never infallible,[183] though, producers may not be in a much better position than distributors to verify child actors' ages. Moreover, the statutory-rape analogy makes little sense for overbreadth purposes because this doctrine has never been applied outside of the First Amendment.[184]

Finally, overbreadth jurisprudence obfuscates whether strict liability chills a sufficiently substantial amount of protected speech to require the defense. Courts have argued that the value of the chilled speech is marginal because adult pornography is not pure speech.[185] However, the first warrant to this argument is conclusory because *Broadrick*'s balancing test is implicated when the allegedly chilled "speech" is actually chilled "expressive conduct."[186] Thus, the argument that the speech of adult pornography lacks value under the overbreadth test *because it is expressive conduct* is circular. Courts have also argued that the chilled speech is not substantial because it does not hit the political heart of the First Amendment.[187] This argument

177. *See supra* subpart IV(B).

178. *See supra* subpart IV(B).

179. Gilmour v. Rogerson, 117 F.3d 368, 375 (8th Cir. 1997) (Arnold, J., dissenting).

180. *But cf.* Lawrence v. Texas, 539 U.S. 558, 584–85 (2003) (holding that there is a fundamental right to privacy that extends to consensual adult sexual relations and thereby undermining Judge Arnold's contention that there is no constitutional right to sexual relations with anyone).

181. *See supra* note 113.

182. *See* Outmezguine v. State, 641 A.2d 870, 876 (Md. 1994) (summarizing the state's argument for strict liability—that producers and photographers are in a better position to discover the ages of their subjects).

183. United States v. U.S. Dist. Court, 858 F.2d 534, 540 (9th Cir. 1988).

184. Newmark, *supra* note 68, at 560.

185. Gilmour v. Rogerson, 117 F.3d 368, 373 (8th Cir. 1997).

186. *See supra* text accompanying note 72.

187. *Outmezguine*, 641 A.2d at 878–79.

carries more precedential weight, as the Supreme Court has noted that different types of speech have different values.[188] This contention is also consistent with the history of the First Amendment, which was framed to protect unpopular political views.[189] However, it remains unclear whether strict liability chills enough speech to be considered substantial under *Broadrick*.

C. *An Intermediate Standard*

On one hand, the Ninth Circuit's reasonable mistake of age defense hardly protects children who are deceitful about their age—arguably those children that need to be protected the most. On the other, strict-liability jurisdictions have not persuasively justified leaving no breathing space for a distinct category of protected speech—though marginally valuable—given that the mistake of age defense can still facilitate child-protection interests. In the absence of proposals for a middle ground,[190] this Note proposes an intermediate mistake of age defense, which would provide more protection to deceitful children than the Ninth Circuit's test and chill less protected speech than the strict-liability standard. More specifically, state and federal courts[191] should consider permitting defendants to claim a mistake of age defense if they show by clear and convincing evidence that they (1) had no actual knowledge of the child's age; (2) requested and reviewed the child's proof-of-age documentation; and (3) verified the validity of the presented documents with government records or government officials, which happened to be defective or misleading.

These elements could be met in circumstances very similar to the facts of *District Court*. In *District Court*, the defendants had no actual knowledge of the child's age and requested and received the child's identification, which was fraudulent.[192] However, a defendant would not be able to succeed under the proposed intermediate standard for the mistake of age defense unless the defendant also showed by clear and convincing evidence that this inaccuracy

---

188. *Gilmour*, 117 F.3d at 373 (citing United States v. X-Citement Video, 513 U.S. 64, 84 (1994)).

189. *See supra* text accompanying notes 48–49.

190. *See generally* Jorn Axel Holl, Comment, *Judges, Congress, and the Sixteen-Year-Old Porn Star: Questions on the Proper Role of the First Amendment*, 75 IOWA L. REV. 1355, 1357 (1990) (arguing that the Ninth Circuit decided *District Court* incorrectly when it determined to add a reasonable mistake of age requirement based on First Amendment concerns); Simmons, *supra* note 14, at 121–25 (disagreeing with the *Outmezguine* court's analysis and concluding that it should have followed the Ninth Circuit's *District Court* opinion and permitted a reasonable mistake of age defense); Robert R. Strang, Note, *"She Was Just Seventeen . . . and the Way She Looked Was Way Beyond [Her Years]": Child Pornography and Overbreadth*, 90 COLUM. L. REV. 1779, 1803 (1990) (concluding that courts should recognize a mistake of age defense for distributors of child pornography, but not for producers).

191. State or congressional legislation that statutorily provides this defense should be considered, but whether judicial or legislative action is preferable is outside the scope of this Note.

192. United States v. U.S. Dist. Court, 858 F.2d 534, 536 (9th Cir. 1989).

was also confirmed by an original government document (or by a government employee that had access to those documents), including those on file with local government agencies such as an original birth certificate, passport, or driver's license.[193]

In lieu of the Ninth Circuit's current reasonableness rule, this intermediate standard would increase the general protection of children against sexual exploitation by specifying the particular instances under which defendants could claim the defense, as it would presumably include only the very rare circumstances where the original official government records were wrong with regard to a birth date or where government officials miscommunicated information from an accurate birth record. The intermediate standard would also offer more protection to those particular children who were deceitful about their age in at least two ways. First, by putting the risk of fraud or deceit on the pornographer, it would further encourage age verification with government documents, which are more likely to be accurate. Second, it would discourage children from lying about their age because vigorous age-verification practices would likely catch this fraud.

An intermediate standard would also be more palatable to the First Amendment. As the *District Court* court elucidated, strict liability would inevitably chill some constitutionally protected speech because of the fallibility of the sources of evidence that Judge Beezer argued could be used to verify age, such as "reputation, first-hand testimony, and . . . documents."[194] Under the strict-liability standard, defendants assume the risk of faulty or misleading documents and information provided by the government. By lifting this risk off of potential producers of child pornography, they might be more likely to seek out the proper verification if they know that it might shield them from criminal liability.

This intermediate standard would sit well with those who have rejected the defense because they have explained that a defendant's burden should be to do exactly what this immediate standard proposes. However, the strict-liability principle would not permit a defendant to raise a mistake of age defense even where the defendant met these burdens as these courts have

193. Such factual circumstances would indeed be very rare, but official government records are not infallible. *Id.* at 540; *see also, e.g.*, Julie A. Greenberg, *Defining Male and Female: Intersexuality and the Collision Between Law and Biology*, 41 ARIZ. L. REV. 265, 309 (1999) (suggesting that errors with regard to the sex of a baby can be made on birth certificates by medical assistants at the time of birth). The potential for errors also exists, for example, any time a change is made to an official document. Revisions to a birth certificate, for example, may be granted for a number of reasons, including correcting a factual mistake on a birth certificate, and sometimes when a person or child changes name or sex. *Id.* at 309–11. In addition, revisions could be made to a birth certificate when a child changes parents. Michael J. Ritter, Note, *Adoption by Same-Sex Couples: Public Policy Issues in Texas Law and Practice*, 15 TEX. J. C.L. & C.R. (forthcoming 2010) (manuscript at 12 & n.63, on file with author) (discussing state laws on the issuance of supplementary birth certificates after an adoption). Under a strict-liability scheme, the risk for errors in official government records could also serve as an additional chilling factor.

194. *Dist. Court*, 858 F.2d at 546–47 (Beezer, J., dissenting).

described them. For example, Judge Beezer articulated a comparable standard in his *District Court* dissent:

> Documents establish age. While documents may be counterfeited, the originals exist somewhere. It would be simple for a pornographer to write his subject's birthplace for a certified copy of the subject's birth certificate. A pornographer might even go see the original himself. By obtaining proof in this fashion, a pornographer could eliminate all doubt about the subject's age.[195]

The *Outmezguine* court similarly stated:

> A photographer or filmmaker is in a position to ascertain the true age of the individual being photographed or filmed by requiring that individual to produce a birth certificate, driver's license, or similar governmental identification card. A reasonable bona fide attempt to verify the authenticity of such documents will thus ensure that the subject being used is not a child.[196]

While what Judge Beezer and the *Outmezguine* court described might almost always be true, they do not sufficiently rebut the *District Court* majority's point that documents are *never infallible*.[197] The risk of miscommunication by a subject's birthplace or a problem with birth records would still fall on a defendant under the strict-liability rule and protected speech would still be chilled. By permitting defendants to raise a defense if they can show by clear and convincing evidence that they checked a child subject's government-issued identification for a child's age and subsequently verified the age with the original government document or with a government official, courts could provide the breathing space for protected speech that the First Amendment would prefer while protecting the interests of children at the same time.

VI. Conclusion

Though the federal and state governments have criminalized the production of child pornography, producers continue to sexually exploit children by visually recording their sexual acts. This criminalization is problematic because these statutes generally omit mistake of age defenses. This omission raises First Amendment free speech concerns because strict liability would inevitably chill some constitutionally protected speech—the production of adult pornography. As defendants have challenged criminal statutes without mistake of age defenses as unconstitutionally overbroad, different courts have ended up on different sides of the debate. The Ninth Circuit has held that the First Amendment requires a reasonable mistake of

195. *Id.* at 546.

196. Outmezguine v. State, 641 A.2d 870, 878 (Md. 1994).

197. Under the court's interpretation of the statute, a photographer who diligently checks identification could still violate the regulation if the identification provided is forged or inaccurate.

age defense under child pornography statutes; all other courts have opted for strict liability. Because neither side of this debate has sufficiently addressed opposing arguments, the current approaches have not satisfactorily achieved both child-protection and free speech goals. An intermediate standard that requires defendants to show by clear and convincing evidence that they had no actual knowledge of a child's age; requested and reviewed documents falsely stating that the child was an adult; and verified this false information with a government record or official, which happened to be incorrect, would check the inadequacies of both standards by boosting protection of deceitful children and providing breathing space for protected speech.

—Michael J. Ritter

* * *

Soul Searching and Profit Seeking: Reconciling the Competing Goals of Islamic Finance*

I. Introduction

The worldwide growth of the Islamic financial industry over the past decade has been steady.[1] Increasingly, the ethical principles that underlie the Islamic financial system have been a popular topic of debate.[2] While Islamic finance has much to offer with regard to financial management, ethical investing, and project finance, there are fundamental tensions within the system that may work to stifle its growth and foreclose opportunities to reach a broader and more diverse investor group.

Chief among these difficulties is the fact that the Islamic financial industry lacks what some consider a necessary cohesive and overarching governance structure. Different countries, and the various sects of Islam within those countries, each have their own interpretations of both the religious and financial teachings of Islam. This has caused the related problems of inconsistent enforcement, inaccurate risk estimation, and the generalized hesitancy among even Muslim investors to pursue Islamic financing options.

This Note, in Part II, provides a brief overview of the Islamic financial system's development, its current status, and its primary methods for financing and investing in compliance with Shari'a law. Part III then outlines several of the problems and issues that have prevented more widespread acceptance of Islamic finance as an alternative to conventional, Western financing techniques. These central issues are (1) a fragmented regulatory structure; (2) the regional differences among Muslim countries' interpretations of Shari'a law; (3) the heightened level of risk involved in anticipating future trends in the Shari'a compliance requirements; and (4) the lack of scholars specializing in Islamic finance.

In Part IV, I outline three general proposals aimed to address these problems. The proposals seek to prevent and manage the risks these problems create for the Islamic financial system. The proposals include (1) streamlining the educational system for Islamic financial experts; (2) creating new methods for avoiding conflicts of interests among the field's

---

* I would like to thank Professor Henry T.C. Hu for his guidance and advice during the preparation of this Note.

1. *See* Andreas Junius, *Islamic Finance: Issues Surrounding Islamic Law as a Choice of Law Under German Conflict of Laws Principles*, 7 CHI. J. INT'L L. 537, 538 (2007) (discussing the increasing number of Islamic financial institutions worldwide); Theodore Karasik et al., *Islamic Finance in a Global Context: Opportunities and Challenges*, 7 CHI. J. INT'L L. 379, 379 (2007) (noting, as of 2006, growth rates nearing 15% per year).

2. *See* Anita Hawser, *Back to Basics: Islamic Financing*, GLOBAL FIN., Nov. 2008, at 31, 31 (finding that Islamic financial institutions fare better than their conventional counterparts during times of economic distress due in large part to the system's emphasis on ethics).

influential scholars; and (3) changing how Shari'a compliance ratings are computed for Islamic financial institutions and companies.

Lastly, in Part V, I employ case studies of a recently developed and somewhat controversial Islamic financial product, the *tawarruq*, to illustrate the identified problems that have been associated with the Islamic financial system. Additionally, this study of *tawarruq* demonstrates how the implementation of even limited versions of the proposed modifications may help popularize the Islamic approach to investing and finance. Addressing this issue and exploring Islamic investing trends is particularly relevant given the system's short history, recently increasing popularity, and arguable viability as an alternative to the risky methods of conventional banking that have recently caused such severe economic turmoil.

II. A Brief Introduction to Islamic Finance

A. Shari'a Law

Islamic finance can be broadly described as a financial system that is intended to function in compliance with Shari'a law. For Muslims, Shari'a law serves as the principle source of guidance for all areas of their lives.[3] The term "Shari'a" can be roughly translated as "Islamic law" and is often interpreted by Muslims as "the totality of divine categorizations of human acts."[4] Shari'a law is thus an umbrella term that refers to four distinct sources of religious and legal tradition.[5] The primary materials from which Islamic law is derived include, in order of significance, (1) the Holy Qur'an, (2) the *hadith*, (3) *ijm'a*, and (4) *qiyas*.[6] Muslims believe the Qur'an contains the literal words of Allah as revealed to Muhammad.[7] The *hadith* are the recordings of the Prophet Muhammad's actions and words as documented by his contemporaries and later followers via oral tradition.[8] These examples set by Muhammad are clarified, expanded, and made applicable to present conditions primarily through a form of analogical reasoning known as *qiyas*.[9] And finally, when Islamic jurists reach a consensus on the proper application of *qiyas*, it results in a per se valid and binding religious law known as *ijm'a*,[10] popularly translated to mean "consensus of jurists."[11]

---

3. Gohar Bilal, *Islamic Finance: Alternatives to the Western Model*, FLETCHER F. WORLD AFF., Winter/Spring 1999, at 145, 146.

4. *Glossary*, *in* STRUCTURING ISLAMIC FINANCE TRANSACTIONS 226, 232 (Abdulkader Thomas et al. eds., 2005).

5. Haider Ala Hamoudi, *Jurisprudential Schizophrenia: On Form and Function in Islamic Finance*, 7 CHI. J. INT'L L. 605, 608 (2007).

6. Irshad Abdal-Haqq, *Islamic Law: An Overview of Its Origin and Elements*, 7 J. ISLAMIC L. & CULTURE 27, 36 (2002).

7. *Glossary*, *supra* note 4, at 231.

8. Bilal, *supra* note 3, at 146.

9. Hamoudi, *supra* note 5, at 608.

10. *Id.*

B. The Importance of Islamic Finance Today

Islamic finance is a relatively new brand of finance. Its roots are traced to a bank in Cairo, Egypt, founded in 1963.[12] Since then, Shari'a law has become increasingly popular and has become a player in both Muslim and Western countries' financial markets.[13] As of the close of 2007, there were $500 billion invested in Shari'a-compliant assets,[14] which reflects a growth rate greater than 10% per year for each of the past ten years.[15] In recent months, Islamic finance has received increased attention in mainstream media outlets in the context of the recent and worsening worldwide financial crisis.[16]

C. Departure from Conventional Banking

One of the world's foremost scholars in Islamic finance, Sheikh Muhammad Taqi Usmani, has written: "[T]he basic difference between capitalist and Islamic economy is that in secular capitalism, the profit motive or private ownership are given unbridled power to make economic decisions."[17] This idea may best be characterized as a distinction between a financial system driven purely by profits and one that contains the dual goals of religious piety and profit maximization.[18]

The most central practical differences between Shari'a and conventional finance revolve around the various restrictions on the types and methods of investments allowable under an Islamic approach. Muslims rely on Shari'a law for the proposition that investment in the following things, among others, are *haraam* (forbidden): the charging of *riba* (interest), engagement in excessively speculative ventures, contractual uncertainty or ambiguity, traditional insurance protection, and industries that deal in gambling, pornography, alcohol, tobacco, pork products, and even those that produce media products

---

11. *Glossary*, *supra* note 4, at 228.

12. SAYED KHATAB & GARY D. BOUMA, DEMOCRACY IN ISLAM 110 (2007).

13. *See* Karasik et al., *supra* note 1, at 379 ("[T]here are over 300 Islamic financial institutions in more than 75 countries . . . .").

14. Oliver Agha, *Islamic Finance in the Gulf: A Practitioner's Perspective*, 1 BERKELEY J. MIDDLE E. & ISLAMIC L. 179, 179 (2008).

15. *See* Juan Solé, *Islamic Banking Makes Headway*, IMF SURV. MAG., Sept. 19, 2007, http://www.imf.org/external/pubs/ft/survey/so/2007/RES0919A.htm (asserting that the Islamic banking industry has grown 10%–15% per year over the last ten years).

16. *See, e.g.*, Hawser, *supra* note 2, at 31 (reporting on Islamic finance's resilience during the credit crunch).

17. MUHAMMAD TAQI USMANI, AN INTRODUCTION TO ISLAMIC FINANCE, at xiv (2002).

18. *See id.* (discussing the attempt of Islamic finance to protect societal interests and divine restrictions within a market-based economy); *see also* Umar F. Moghul & Arshad A. Ahmed, *Contractual Forms in Islamic Finance Law and* Islamic Inv. Co. of the Gulf (Bahamas) Ltd. v. Symphony Gems N.V. & Ors.*: A First Impression of Islamic Finance*, 27 FORDHAM INT'L L.J. 150, 152 (2004) (recounting that Islamic finance is closely tied to Islamic religious principles).

such as gossip magazines.[19] Many scholars believe that in no instance should any of the forbidden products comprise more than 5% of the total revenues of any Shari'a-compliant business.[20]

The prohibition of interest is often considered the centerpiece of the Islamic banking system. The insistence on adherence to this rule is derived both from passages from the Qur'an and teachings of Muhammad.[21] The central Qur'anic passage on which Islamic finance is based reads:

> Those who devour usury will not stand except as stands one whom Satan by his touch hath driven to madness. That is because they say: "Trade is like usury," but Allah hath permitted trade and forbidden usury. Those who after receiving direction from their Lord, desist, shall be pardoned for the past; their case is for Allah (to judge); but those who repeat (the offence) are Companions of the Fire: they will abide therein (for ever).[22]

In addition to the widely accepted prohibition of interest, it is often necessary for Shari'a-compliant companies to appoint and maintain a Shari'a board that provides guidance to the company's leadership on matters of Shari'a law and compliance.[23] Each board should technically contain at least three Islamic scholars,[24] though there are various interpretations as to what this means, creating a problem that will be addressed later in Part III.

D. Islamic Financial Products

Islamic finance has demonstrated an ability to innovate and adapt to changing economic times, and it has experienced a wave of innovation over the past two decades. The most common forms of financing in Shari'a-compliant industries are addressed below. While the following list of financial products is in no way exhaustive, these are collectively considered

---

19. Shahzad Q. Qadri, *Islamic Banking: An Introduction*, BUS. L. TODAY, July–Aug. 2008, at 59, 59; *see also* Angela Jameson, *Conventional Insurance in Conflict with Islam*, TIMES ONLINE, Mar. 1, 2008, http://business.timesonline.co.uk/tol/business/industry_sectors/banking_and_finance/article3463702.ece ("Conventional insurance products are in conflict with Islamic beliefs for three reasons. Insurance involves an element of uncertainty, gambling and the charging of interest, which are prohibited by the Koran."). This is not intended to be a complete list of all prohibited items in the Islamic financial system. However, these remain the most central and visible restrictions on Islamic investment as it currently exists. *Cf.* Qadri, *supra*, at 59 (noting that "a consensus among Muslim scholars" confirms the existence of the religious prohibitions on these economic activities).

20. *See* Zaineb Sefiani, *Inside View: Islamic Finance*, FUNDS EUR., July 2009, http://www.funds-europe.com/July-2009/INSIDE-VIEW-Islamic-finance/menu-id-228.html (describing the first step in screening investments for Shari'a compliance as filtering out companies that derive more than 5% per year from forbidden business sectors).

21. Agha, *supra* note 14, at 180.

22. HOLY QUR'AN 2:275 (Abdullah Yusuf Ali trans., 2000).

23. Neil Miller & David Baylis, *Sharia-Reliant*, LAWYER, May 28, 2007, http://www.thelawyer.com/sharia-reliant/126147.article.

24. *Id.*

the most widely used.[25] The descriptions of these products illustrate the methods that have been employed to avoid formalized interest payments.

1. Mudharabah.—*Mudharabah* refers to a profit-sharing contractual agreement typically between a financial institution or investor and an entrepreneur who is seeking funding for a project or endeavor.[26] The investor or institution gives money to engage in the entrepreneur's business activity, and the entrepreneur provides the labor and expertise.[27] Prior to the beginning of the business activity, the two parties determine a ratio at which they will share profits, and all profits that are made in the venture are shared according to that ratio.[28] Likewise, losses from the business venture are also shared, and because the investor or bank is exposed to this risk, this justifies the party's claim to part of the profits in the event of a gain.[29]

2. Musharakah.—Along with the *mudharabah*, joint-venture financing, or *musharakah*, is among the more common methods used to engage in project finance, real-estate purchases, letters of credit, and other investment projects in primarily Muslim countries.[30] *Musharakah* is literally translated as "sharing."[31] Each partner in a *musharakah* arrangement inherently has the right to equal management authority over the venture, even though their respective investments may be unequal.[32] However, much like in Western business law, the parties are able to deviate from this presumption via express written contract.[33]

3. Murabaha.—The *murabaha* form of Islamic finance has many variations, but at its root it consists of a bank or financial institution buying an asset and selling it back to the customer, who will make either a single

---

25. *See, e.g.*, Qadri, *supra* note 19, at 59–60 (describing *mudharabah*, *wadiah*, *musharakah*, *murabaha*, and *ijarah* as key basic concepts in Islamic banking); Abdulkader Thomas, *Introduction: The Origins and Nature of the Islamic Financial Market*, *in* STRUCTURING ISLAMIC FINANCING TRANSACTIONS, *supra* note 4, at 1, 7–8 (naming *mudharabah*, *musharakah*, *ijarah*, and *sukuk* as core financing mechanisms); Ahmad Lutfi Abdull Mutalip & Mohd Herwan Sukri Mohammad Hussin, The Emergence of Islamic Financing Based on the *Syariah* Concept of *Tawarruq* 1 (2008) (unpublished manuscript, on file at http://www.azmilaw.com.my/Article/Article_8_&_9/Article_9_Tawarruq_00093603_.pdf) (naming *bai' bithaman ajil* as a traditional type of sale and *tawarruq* as a newer type of sale that is gaining popularity). *See generally* USMANI, *supra* note 17 (discussing *musharakah*, *murabaha*, and *ijarah* as principal parts of the Islamic financial system).

26. Qadri, *supra* note 19, at 59.

27. *Id.* at 59–60.

28. Bilal, *supra* note 3, at 156.

29. Qadri, *supra* note 19, at 59.

30. *Id.* at 60.

31. Muhammad Taqi Usmani, *The Concept of* Musharakah *and Its Application as an Islamic Method of Financing*, 14 ARAB L.Q. 203, 203 (1999).

32. Husam Hourani, *The Three Principles of Islamic Finance Explained*, INT'L FIN. L. REV., May 2004, at 46, 47.

33. Qadri, *supra* note 19, at 60.

deferred payment or multiple deferred payments over time.[34] The purchase and sale price, along with the explicit profit margin, are agreed upon up-front at the time of the original sales agreement. The profit margin in this sense is acceptable because the bank is viewed as being compensated for the time value of the money.[35]

4. Ijarah.—Literally translated as "compensation," "substitute," "consideration," "return," or "counter value,"[36] *ijarah* is a contract that involves the lease or transfer of ownership of a service for a specified period in exchange for prearranged consideration.[37] *Ijarah* is best likened to a simple lease form, and it is almost uniformly accepted as Shari'a-compliant as "a convenient means for people to acquire the right to use any asset that they do not own, as all people might not be able to own the tangible assets for use."[38] Islamic finance views dealing in assets or other intangibles that one does not own harshly,[39] but due to its structure, *ijarah* avoids this pitfall.

5. Bai' Bithaman Ajil *(BBA).*—A contract of *bai' bithaman ajil* involves a deferred-payment sale or a "credit sale."[40] This product differs from a concurrent purchase and delivery of an asset (such as in a *murabaha* agreement) because it allows for a deferred delivery or payment of existing assets.[41] Interestingly, a BBA does not typically require the lender, or lessor, to disclose a specific profit margin up-front, which differentiates this from many Shari'a-compliant arrangements.[42] This is because the vast majority of BBAs involve significantly long-term ventures.[43]

6. Wadiah.—Arabic for "custody," a *wadiah* is a contract between an account holder and a bank where the account holder places his funds in trust with the bank, and the bank, in return, keeps and invests the funds,

---

34. *Id.*

35. *Id.*

36. Ijarah, http://nurhisyammuhasabah.blogspot.com/2009/02/ijarah.html (Feb. 5, 2009, 13:31 EST).

37. *Glossary*, *supra* note 4, at 228.

38. MUHAMMAD AYUB, UNDERSTANDING ISLAMIC FINANCE 279 (2007).

39. *See generally* Mohammad Nejatullah Siddiqi, Economics of *Tawarruq*: How Its *Mafasid* Overwhelm the *Masalih* (Mar. 10, 2008) (unpublished manuscript, on file at http://konsultasimuamalat.com/home/index.php?view=article&catid=1%3Alatest-news&id=45%3Aeconomics-of-tawarruq--how-its-mafasid-overwhelm-the-masalih&format=pdf&option=com_content) (arguing that *tawarruq*, a tool to make purchases on deferred payment, is not Shari'a compliant because the macroeconomic harms it causes outweigh the benefits).

40. Abdulkader Thomas, *Changes and Challenges*, *in* STRUCTURING ISLAMIC FINANCE TRANSACTIONS, *supra* note 4, at 222, 225.

41. Andreas A. Jobst, *Derivatives in Islamic Finance*, *in* THE CREDIT DERIVATIVES HANDBOOK 16, 19–20 (Greg N. Gregoriou & Paul Ali eds., 2008).

42. *Id.* at 19 n.8.

43. *Id.*

guaranteeing repayment of any part of the funds on request.[44] While this may sound much like a traditional bank account, the major difference is that depositors are not entitled to any rewards or interest payments for entrusting their money to the bank or financial institution.[45] This allows the parties to the contract to avoid committing the forbidden act of interest charging and payment. In lieu of interest payment, most financial institutions engaging in *wadiah* contracts will pay "gifts" to their depositors periodically, and many depositors expect this payment, though there is no formal requirement for the bank to do so.[46]

7. Sukuk.—Frequently referred to as "Islamic bonds," *sukuk* is more accurately translated as "Islamic investment certificates."[47] This is a more precise definition, considering that the primary difference between *sukuk* and bonds is that *sukuk* do not draw traditional interest.[48] The legal structure of *sukuk* is most analogous to U.S. trust certificates.[49] A traditional bond is a contractual debt obligation, and the bond issuer is obligated to pay bondholders both interest and principal at agreed-upon intervals.[50] With a *sukuk* issuance, the holders each hold "an undivided beneficial ownership interest in the underlying assets" and are thus entitled to share both in the *sukuk* revenues as well as the proceeds of the realization of the *sukuk* assets.[51] However, as the *sukuk* system has progressed and grown, borrowers would usually promise to buy back the assets irrespective of whether the assets made money.[52] This has led to some controversy surrounding *sukuk* issuances, to be addressed below.[53]

8. Tawarruq.—Because not all funding endeavors can actually be supported by physical assets, the *tawarruq* financing method has grown in popularity.[54] The basic structure of this transaction type occurs where an

---

44. HOLGER TIMM, THE CULTURAL AND DEMOGRAPHIC ASPECTS OF THE ISLAMIC FINANCIAL SYSTEM AND THE POTENTIAL FOR ISLAMIC FINANCIAL PRODUCTS IN THE GERMAN MARKET 40 (2004).

45. *Id.*

46. Qadri, *supra* note 19, at 60.

47. Abdulkader Thomas, *Opportunities with* Sukuk *and Securitisations*, *in* STRUCTURING ISLAMIC FINANCE TRANSACTIONS, *supra* note 4, at 154, 154.

48. Jeff Black, *An Unhealthy Interest?*, MIDDLE E., July 2008, at 44, 44.

49. Thomas, *supra* note 47, at 154.

50. Tamara Box & Mohammed Asaria, *Islamic Finance Market Turns to Securitization*, INT'L FIN. L. REV., July 2005, at 21, 21.

51. *Id.* at 21–22.

52. *See* Kit R. Roane, *Fatally Flawed Bonds*, PORTFOLIO.COM, Sept. 23, 2008, http://www.portfolio.com/news-markets/top-5/2008/09/23/Changes-in-Islamic-Finance (noting Usmani's assessment that some 85% of *sukuk* issuances at the time guaranteed return of the principal regardless of default); *see also infra* notes 86–87 and accompanying text.

53. *See infra* notes 82–91 and accompanying text.

54. MAHMOUD A. EL-GAMAL, ISLAMIC FINANCE: LAW, ECONOMICS, AND PRACTICE 69 (2006); *see also* Salah Al-Shalhoob, Organised *Tawarruq* in Islamic Law 3 (unpublished

individual or company buys a commodity from a financial institution and then resells the commodity to a third party for cash.[55] Like the other above-mentioned instruments, in a *tawarruq* the bank is repaid over a fixed period of time with an inbuilt profit for the use of that commodity.[56] A case study of *tawarruq* development is included in Part V below.

The Islamic financial products discussed above each have numerous variations that have developed in response to changes in investor needs and conventional financial products. Many of the hybrid products incorporate elements of multiple financing methods, often achieving a result that parallels conventional investments.

III. Mixing Business and Religion: Recipe for Conflict

The restrictions placed on Shari'a-compliant investment funds cause rifts and various points of contention when applied within the framework of conventional banking and finance. Scholars and commentators, Muslim and non-Muslim alike, have pointed to numerous issues with the current way Islamic finance is structured.[57] The nature of a religiously based financial system that crosses geographic, ethnic, cultural, and doctrinal lines creates problems of consistency and levels of risk that have scared away some potential investors.[58]

A. Fragmented Regulatory Structure

There are numerous bodies that oversee the Islamic financial system, interpret Shari'a law, and issue recommendations, *fatwas*, and other forms of guidance on how to invest in accordance with the will of Allah. Some commentators have argued that this system has hindered the success and advancement of Islamic finance and its adaptation to modern conventions.[59]

---

manuscript, on file at https://eprints.kfupm.edu.sa/14894/1/organised_tawarruq_in_Islamic_law_(Conf_23_Apr_2007).pdf) (reporting that the Saudi Arabia British Bank introduced *tawarruq* in 2000 to facilitate trading on the international commodities market by Islamic investors); Mutalip & Hussin, *supra* note 25, at 1 (reporting that many Malaysian institutions offer *tawarruq* in order to avoid scholarly criticism of other forms of financing).

55. Stella Cox & Abdulkader Thomas, *Liquidity Management: Developing the Islamic Capital Market and Creating Liquidity*, *in* STRUCTURING ISLAMIC FINANCE TRANSACTIONS, *supra* note 4, at 171, 174–75.

56. *Id.*

57. *See, e.g.*, Agha, *supra* note 14, at 182 ("[U]ncertainty is considered a problem under Islamic Finance."); Kilian Bälz, *Islamic Finance for European Muslims: The Diversity Management of Shari'ah-Compliant Transactions*, 7 CHI. J. INT'L L. 551, 556 (2007) ("[F]or centuries, the Islamic Shari'ah has been a discursive legal system with a fair degree of pluralism (or uncertainty) in terms of black letter rules."); Rodney Wilson, *Capital Flight Through Islamic Managed Funds*, *in* THE POLITICS OF ISLAMIC FINANCE 129, 129–31 (Clement M. Henry & Rodney Wilson eds., 2004) (discussing the capital flight from Islamic countries due to the structure of Islamic finance).

58. *See, e.g.*, *infra* notes 80–92 and accompanying text (discussing a drop in issuances of *sukuk* following a speech by Sheikh Muhammad Taqi Usmani).

59. *See, e.g.*, sources cited *supra* note 57.

1. Accounting and Auditing Organization for Islamic Financial Institutions.—Among the several agencies that pass guidelines and regulate Islamic financial systems internationally, the Accounting and Auditing Organization for Islamic Financial Institutions (AAOIFI) is arguably the most important. Founded in 1991, the AAOIFI describes its role as "an Islamic international autonomous non-for-profit corporate body that prepares accounting, auditing, governance, ethics and Shari'a standards for Islamic financial institutions and the industry."[60] While it is located in Bahrain, the AAOIFI has two-hundred members from forty-five countries, and these members primarily include central banks and Islamic financial institutions.[61] Seven countries have adopted and implemented the AAOIFI's standards, and six others have issued guidelines and laws based on the AAOIFI's standards.[62]

Some commentators have noted that the AAOIFI has focused more on the "big ticket" market and given less attention to the retail market of Islamic finance.[63]

> In addition, the AAOIFI does not develop products; it simply establishes a framework within which products can be developed. Furthermore, even institutions that heavily rely on the AAOIFI will normally also retain an internal or external Shari'ah board that supervises compliance with these principles. As a result, standardization in the fashion carried out by the AAOIFI may help provide some orientation, but it will not solve the problems encountered when structuring a concrete product.[64]

2. Islamic Financial Services Board.—The Islamic Financial Services Board (IFSB) in Malaysia describes itself as "an international standard-setting organisation that promotes and enhances the soundness and stability of the Islamic financial services industry by issuing global prudential standards and guiding principles for the industry, broadly defined to include banking, capital markets and insurance sectors."[65] The IFSB has published standards and other documents to help guide institutions and countries in developing standards for Shari'a compliance.[66]

---

60. Accounting & Auditing Org. for Islamic Fin. Insts., AAOIFI Overview, http://www.aaoifi.com/overview.html.

61. *Id.*

62. *See id.* (noting implementation of AAOIFI standards in Bahrain, Jordan, Lebanon, Qatar, Sudan, Syria, and Dubai International Financial Centre, and of AAOIFI-based guidelines in Australia, Indonesia, Malaysia, Pakistan, Saudi Arabia, and South Africa).

63. *E.g.*, Bälz, *supra* note 57, at 556.

64. *Id.*

65. Islamic Fin. Servs. Bd., http://www.ifsb.org.

66. Islamic Fin. Servs. Bd., Standards Development, http://www.ifsb.org/standard.php.

3. Islamic International Rating Agency.—The Islamic International Rating Agency (IIRA) is the only ratings agency that provides a rating system that encompasses the "full array of capital instruments and specialty Islamic financial products."[67] The IIRA provides traditional ratings of Shari'a-compliant institutions, similar to more commonly used ratings agencies like Moody's and Standard & Poor's. However, in 2005 the IIRA became the first agency to also offer a Shari'a Quality Rating[68] (and currently remains the only one to do so). The Quality Rating gives investors a sense of the level of Shari'a compliance of certain Islamic financial institutions.[69]

4. International Islamic Financial Market.—The International Islamic Financial Market (IIFM), located in Bahrain, was created out of the efforts of a number of central banks and government agencies, including Bahrain, Brunei, Dubai, Indonesia, Malaysia, Pakistan, Sudan, and Saudi Arabia, among others.[70] Among the IIFM's primary objectives is to "encourage self-regulation for the development and promotion of the Islamic Capital and Money Market segment."[71] The IIFM issues trade guidelines, best-practice procedures, and standardized financial contracts in its efforts to promote Islamic financial innovation.[72]

5. International Islamic Fiqh Academy.—The International Islamic Fiqh Academy, located in Saudi Arabia, is an organ of the Organisation of the Islamic Conference.[73] Among its stated objectives is to "study contemporary problems from the Sharia point of view and to try to find the solutions in conformity with the Sharia through an authentic interpretation of its content."[74] Rulings of the Fiqh Academy have been considered generally decisive on a number of recent Islamic finance issues.

This structure of various regulatory bodies based in different countries, with different membership requirements and methods for evaluating financial products, poses some difficulties in advancing a cohesive message. First, most of these organizations have been formed in the past ten to fifteen years

67. Islamic Int'l Rating Agency, Corporate Profile, http://www.iirating.com/profile.asp.

68. *See id.* (emphasizing that the IIRA "is the sole rating agency established to provide capital markets and the banking sector in predominantly Islamic countries with a rating spectrum that encompasses the full array of capital instruments and specialty Islamic financial products, and to enhance the level of analytical expertise in those markets").

69. Islamic Int'l Rating Agency, Products & Services, http://www.iirating.com/service.asp.

70. Int'l Islamic Fin. Mkt., http://www.iifm.net.

71. Int'l Islamic Fin. Mkt., Objectives, http://www.iifm.net/AboutUs/CorporateProfile/Objectives/tabid/61/Default.aspx.

72. *Id.*

73. USMANI, *supra* note 17, at xi.

74. Organisation of the Islamic Conference, Subsidiary Organs, http://www.oic-oci.org/page_detail.asp?p_id=64#FIQH.

and are still evolving.[75] There are regular instances when the notion of Shari'a-compliance is defined differently by these organizations or is overbroad, leaving other scholars and financial officials unsure of where the lines are drawn. Second, because these organizations are international in nature and largely advance guidelines rather than binding laws, it has proven difficult in some instances to reconcile their directives with the disparate bodies of national law among different countries.[76]

B. Regional Differences

In addition to the structural and organizational gaps of the Islamic financial system, conflicting issues arise when national norms and interests are advanced and given precedence over international standards and guidelines on Islamic financial methods. One example of how differently various countries, even Muslim countries, can interpret Shari'a compliance can be found in the respective cases of Malaysia and Saudi Arabia. As one well-known Islamic finance practitioner casually characterized the range in implementation of Shari'a principles in the financial sector: "I would put Malaysia on a ten, in terms of permissiveness, Saudi Arabia at about a one, GCC [Gulf Cooperation Council] countries at about 4.5, Dubai exception, maybe five, London, . . . at about six, Pakistan, maybe at 9.5."[77]

Though these majority-Muslim countries are on the front lines of Islamic financial-product development, their differing approaches create two different-looking Islamic financial systems.[78] Not surprisingly, these systems are sometimes inconsistent with regard to acceptable financial practices and interpretation of Shari'a law.[79]

C. Risk of Changing Rules and Varying Legal Opinions

The risk of unexpected rule changes is one of the central and most widely discussed obstacles to expanding Islamic finance into the mainstream and non-Muslim population. Because many Islamic countries do not endorse

---

75. *See, e.g.*, Islamic Int'l Rating Agency, *supra* note 67 (recording that the IIRA began operating in July 2005).

76. *See generally* UBIQ CONSULTANCY, AN INTRODUCTION TO ISLAMIC FINANCE 3 (2007), http://www.ubiqconsultancy.com/docs/islamic_finance.pdf ("Malaysia is seen as more efficient and progressive, to the conventional banker, yet too liberal to Gulf Shariah scholars."); Qadri, *supra* note 19, at 59–60 (recommending that financial institutions comply with the rulings and opinions of several foreign supervisory agencies and entities); Boey Kit Yin, *Malaysia: A Natural Destination*, ACQUISITIONS MONTHLY (ISLAMIC FIN. 2009), Nov. 1, 2009, at 19, 20 ("Malaysian Sharia standards have always been perceived as too liberal in the [Middle East].").

77. Agha, *supra* note 14, at 187.

78. *See id.* at 188 (comparing the permissiveness of Malaysia in accepting the parties' determination that something is Islamic with that of Saudi Arabia, where it must conform with the basic principles in substance and form).

79. *See* UBIQ CONSULTANCY, *supra* note 76, at 3 ("Malaysia has one centralized standard setting board (IFSB), which harmonizes the various interpretations of Shariah law, whereas the Gulf banks[] release their own individual, and generally more conservative, interpretations.").

the notion of binding precedent,[80] meaning that clerics and scholars can change their opinions and disagree with past decisions, there is some degree of uncertainty as to whether a financial method or instrument currently considered Shari'a-compliant will remain so for the length of any given project or investment plan.[81]

A recent and notable example of this shift in opinion is the case of *sukuk* issuances. As discussed above, *sukuk* are the rough equivalent of a Shari'a-compliant bond issue, except without traditional interest payments.[82] In November 2007, an Islamic finance scholar, Sheikh Muhammad Taqi Usmani, questioned whether the issuance of *sukuk* was technically in compliance with the fundamental prohibition against interest.[83] Usmani stated in a policy paper, "The time has come to revisit this matter, and rid *sukuk* of these blemishes."[84] These "blemishes" include, among other things, the now-common practice of marketing asset-backed returns on the basis of the LIBOR rate benchmark, which is a "corruption" according to Usmani.[85] Usmani's discussion of *sukuk* also called into question "a popular type of *sukuk* that promised to pay back the face value of the bond at maturity or in case of default."[86] Because Islamic financial principles require risk sharing, many scholars agreed that this guarantee "ran counter to the spirit of Islamic finance."[87] Up to the time that Usmani released this statement, *sukuk* had been considered the backbone of Islamic finance and had allowed the system to grow and expand into more traditional investment arenas.[88]

After Usmani's pronouncement, *sukuk* issuances dropped off dramatically.[89] While many acknowledge that at least some of this decline may be attributed to the overall decline of worldwide financial markets, it is

---

80. *See* Box & Asaria, *supra* note 50, at 22 (noting Saudi Arabia's "lack of a system of binding precedent"); John H. Vogel, *Securitization and Shariah Law*, *in* ISLAMIC FINANCE NEWS: LEADING LAWYERS 2009, at 63, 64 (S. Slvaselvam et al. eds., 2009), *available at* http://www.islamicfinancenews.com/fla-mag/legal09/legal09.html ("[D]ecisions of courts are not often reported and, even if reported, are generally not considered to establish binding precedent for subsequent decisions.").

81. *See* Agha, *supra* note 14, at 189 ("[T]here is no requirement that just because a structure has been done in the past, that that structure will be considered viable a year from now.").

82. Box & Asaria, *supra* note 50, at 21–22.

83. Black, *supra* note 48, at 44.

84. *Id.*

85. *Id.* LIBOR stands for the London InterBank Offered Rate, which is the interest rate at which banks offer to lend to each other. British Bankers' Ass'n, The Basics, http://www.bbalibor.com/bba/jsp/polopoly.jsp?d=1627. This rate is commonly used as a benchmark for the quotation of interest rates for other types of loans, such as commercial loans. *Id.*

86. Roula Khalaf, *Islamic Finance Must Resolve Inner Tensions*, FIN. TIMES, Mar. 30, 2009, http://www.ft.com/cms/s/0/4894b482-1d44-11de-9eb3-00144feabdc0.html.

87. *Id.*

88. Roane, *supra* note 52.

89. *See* Syed Imad-ud-Din Asad, *An Overview of the* Sukuk *Market*, INVESTOR'S BUS. & FIN. J., Mar. 2009, http://www.jang.com.pk/thenews/investors/mar2009/p2.htm (noting a 66% decline in *sukuk* issuances worldwide in 2008).

likely that Usmani's comments also contributed to the trend.[90] Commentators, scholars, and investors were widely surprised and alarmed by how a single speech could set back progress and investment in a product that had proven so successful in recent years.[91]

Because of the religion-based, nonbinding legal nature of Shari'a-compliant financing, this danger of disagreement among religious leaders carries special weight that might not be present in other legal or regulatory systems. The purpose of engaging in Islamic finance is to make profits while at the same time adhering to the principles and directives of the Islamic faith. For non-Muslims, who lack this second prong of religious conviction, conventional financing would be equally appealing in many ways. However, those who engage in Islamic finance for religious purposes view their religious convictions as intertwined with their conduct both in their personal and professional endeavors.[92]

Some of the problems associated with these inconsistencies have recently gained international attention in the wake of the 2009 Dubai World Nakheel *sukuk* crisis.[93] Dubai World is a global holding company that manages investments for the government of Dubai.[94] In November 2009, the company announced its intention to delay upcoming payments on *sukuk* issued by Nakheel, Dubai World's real estate subsidiary.[95] This was the first large-scale potential *sukuk* default, and there were concerns among creditors as to how payments would be distributed and which creditors would receive priority.[96] The absence of a consistent and reliable system for determining the effect of market fluctuations and other common trends, in this case a near-default, results in tremendous uncertainty for investors.

---

90. Robin Wigglesworth, Sukuk*: Defaults Destabilise a Reviving Market*, FIN. TIMES, Dec. 7, 2009, http://www.ft.com/cms/s/0/90e4c3d6-e2be-11de-b028-00144feab49a,dwp_uuid=aee0503e-e09e-11de-9f58-00144feab49a.html.

91. *See, e.g.*, Haider Ala Hamoudi, *Baghdad Booksellers, Basra Carpet Merchants, and the Law of God and Man: Legal Pluralism and the Contemporary Muslim Experience*, 1 BERKELEY J. MIDDLE E. & ISLAMIC L. 83, 101 (2008) (expressing surprise at the influence of the speech, despite Usmani's having "no current position of official authority in any nation"); Black, *supra* note 48, at 45 (noting the "furore" over Usmani's speech in the financial press).

92. *See* Robert R. Bianchi, *The Revolution in Islamic Finance*, 7 CHI. J. INT'L L. 569, 573 (2007) (discussing the concept that individual Islamic bankers are specially obligated to be honest and moral because "their religion holds them to a higher standard").

93. *See* Vikas Bajaj & Graham Bowley, *Arab Emirates Aim to Limit Dubai Crisis in Pledge to Banks*, N.Y. TIMES, Nov. 30, 2009, at A1 (discussing the global implications of a Dubai World default); Heather Timmons, *Dubai Crisis Tests Laws of Islamic Financing*, N.Y. TIMES, Dec. 1, 2009, at B4 (discussing creditors' reactions to Dubai World's request to delay payment on debt).

94. Dubai World, http://www.dubaiworld.ae.

95. Timmons, *supra* note 93.

96. Timmons, *supra* note 93 (citing Zaher Barakat, professor of Islamic finance at Cass Business School in London, who expressed concern over the inconsistent rules about repaying creditors in the event of *sukuk* default).

D. Too Few Qualified Scholars

Two issues closely related to the problems associated with perceived inconsistency are the methods by which decisions and rulings are made within Shari'a boards, and the composition of these boards.[97] Each Islamic financial institution maintains a Shari'a board, which is comprised of Islamic religious scholars as well as financial experts and businesspersons. Because the field of Islamic finance is a relatively recent development, there are only a select number of scholars that have the necessary educational and professional background to sit on the boards of Islamic institutions and regulatory agencies.[98] Most of these individuals serve on multiple Shari'a boards.[99] There is no formalized prohibition against serving on numerous boards, even when those interests might conflict.

This pattern arguably promotes consistency because a limited group of people can better coordinate the direction of Islamic financial development. But having only a small group of eligible people may hinder expansion and create greater opportunities for self-dealing than might occur otherwise. Additionally, the scholars are in such high demand that it might prove difficult for innovators or entrepreneurs to verify the Shari'a-compliant status of new products or proposals because the Shari'a board members' time is so valuable.[100]

IV. Change Is Needed

Islamic finance has prospered in recent years and made adjustments to accommodate the evolving state of worldwide financial affairs. Even in the wake of the recent financial crisis of 2008, Shari'a-compliant banks have fared relatively well in comparison to conventional ones.[101] However, the Islamic financial system may currently be at a crossroads. Since Sheikh

---

97. *See id.* at 575 (reporting that the Organisation of the Islamic Conference is "racing to develop common ethical standards for Shari'a advisory boards," implying that the boards employ differing, possibly inconsistent, ethical standards in making their decisions). Furthermore, the interconnectedness and opacity of current system create serious ethics issues:

> The same religious scholars frequently advise competing businesses, government regulators, private entrepreneurs, Muslim-run companies in their own regions, and non-Muslim-owned multinational corporations headquartered in Europe, North America, and the Far East. The financial *ulama* often serve their clients not only as outside auditors, but also as permanent consultants or even as regular employees. These inherent conflicts of interest and temptations for self-dealing compromise advisors and clients alike.

Id.

98. *See Savings and Souls*, ECONOMIST, Sept. 6, 2008, at 81, 81 (recognizing that a small group of fifteen to twenty scholars repeatedly sits on the boards of Islamic financial institutions, largely due to the positions requiring knowledge of Islamic law and Western finance, fluency in English and Arabic, and investor and customer comfort with recognized names).

99. *Id.* at 83.

100. *Id.*

101. *See* Hawser, *supra* note 2, at 31 (discussing how Islamic financing's focus on real assets led to Islamic banks requiring fewer financial lifelines than conventional banks).

Usmani's warning about potential problems with Shari'a-compliant bonds, the incidences and amounts of *sukuk* issuances have fallen,[102] and questions have been raised about the sustainability of the system.[103] Recent concerns about the stability of Islamic investment giant Dubai World have also sparked criticism of the Islamic banking system, leading some to question whether Shari'a-compliant instruments actually offer more security than traditional tools.[104] With people everywhere looking for a safe place to invest, the current economic crisis is an important opportunity for Islamic finance to move more directly into the mainstream.

Some proposals for popularizing Islamic finance toe the line between actual and technical Shari'a compliance.[105] Some allege that these types of approaches emphasize profit maximization at the expense of marginalizing the religious and ethical aspects of Islamic banking.[106] There are dangers in abandoning the Shari'a roots of Islamic finance, and it remains important to adhere to traditional values even after implementing changes. In fact, some critics have suggested that the Islamic funds that have suffered the most in the recent downturn are those that have strayed from strict Shari'a compliance.[107] Therefore, the subparts that follow examine three proposals in areas in which Islamic finance could make minor adjustments and produce positive results that would benefit both Muslim and non-Muslim countries and companies. These proposals include both short- and long-term approaches designed to capitalize on the financial crisis and popularize Islamic finance, specifically in Western countries. These proposals recommend the following: (1) uniform educational requirements should be implemented for Shari'a board members and board advisors; (2) conflicts of interest on Shari'a boards should be better regulated; and (3) the current

---

102. *See supra* note 89 and accompanying text.

103. *See* Joanna Slater, *Dubai Deals 'Sukuk' Setback: Islamic Bonds Suffer from Bad Headlines Just as They Were Recovering*, WALL ST. J., Dec. 2, 2009, at C2 (reporting observations by a credit analyst, a CEO of an investment firm, and others concerned about the impact a potential default on Dubai World bonds may have on the overall *sukuk* market).

104. *See* John Foster, *How Sharia-Compliant Is Islamic Banking?*, BBC NEWS, Dec. 11, 2009, http://news.bbc.co.uk/2/hi/business/8401421.stm (noting the substantive similarities between traditional investments and those allegedly Shari'a-compliant instruments while expressing skepticism about Islamic finance).

105. *See* Hamoudi, *supra* note 5, at 605–06 (decrying the excessively formalistic analysis applied by proponents of Islamic finance as an approach that divorces the methods of Islamic finance from the goals of Islamic finance).

106. *E.g.*, Siddiqi, *supra* note 39, at 5–6.

107. *See* John Foster, *How Islamic Finance Missed Heavenly Chance*, BBC NEWS, Dec. 1, 2009, http://news.bbc.co.uk/2/hi/business/8388644.stm (concluding that many of the problems behind Dubai's near-default may be linked to the innovative instruments that have been deemed Shari'a compliant, but that have actually just "circumvent[ed] the principles of [Shari'a] law"); Andrew Ross Sorkin, *A Financial Mirage in the Desert*, N.Y. TIMES, Nov. 30, 2009, http://www.nytimes.com/2009/12/01/business/01sorkin.html (describing the Dubai World *sukuk* as investments that "looked like bonds, walked like bonds and talked like bonds").

ratings system for evaluating the level of Shari'a compliance should include elements similar to a traditional peer-review rating system.

These proposals may offer a bridge to a more widespread acceptance of Shari'a law. In the wake of the financial crisis, new investment opportunities could prove beneficial for both struggling Western investors as well as Islamic financial institutions, which are growing but still comprise only a small portion of the international financial markets. The following proposals should help to assuage the fears of potential investors regarding risk and volatility. These fears may be especially pronounced in countries such as the United States, where the notion of incorporating religious standards into the business arena has not traditionally been part of the professional culture.

A. Streamlining Educational Requirements for Shari'a Law Experts

As discussed above, one of the perceived problems of Shari'a-compliant investing is that investors cannot necessarily rely on consistency in rulings and judgments. This is true in two senses. First, the different members of the Shari'a boards and their advisors in national governments, international regulatory agencies, and individual companies may each interpret Shari'a law differently.[108] Not only are there varying Qur'anic interpretations, but the secondary nature of the *hadith*, *qiyas*, and *ijm'a* increase the probability for widely divergent understandings of meaning.

Second, some investors view the ever-changing rules as an indication that the additional hassle and compliance measures are not necessarily worth the benefit, which may or may not be formal Shari'a compliance.[109] Studies have shown that Islamic finance still has difficulty attracting certain classes of Muslims in a wide range of geographically and socioeconomically diverse groups.[110] One reason for this may be the perceived increase in transaction costs from doing business in compliance with Shari'a law, without any guarantees that the "compliant" status of the transactions would not later be changed.[111]

Thus, increasing efficiency and consistency is a common goal advanced by many scholars and commentators. Though the Islamic financial community has made great strides in the past two decades with the creation of the AAOIFI and the IIRA, the need for further improvement is still evident. Scholars and experts have advanced various notions for achieving a more

---

108. *See supra* notes 75–79 and accompanying text.

109. *See* Mahmoud A. El-Gamal, *Limits and Dangers of* Shari'a *Arbitrage*, *in* ISLAMIC FINANCE: CURRENT LEGAL AND REGULATORY ISSUES 117, 121–22 (S. Nazim Ali ed., 2005) (noting that due to the close connection and similarity to conventional financing, most potential customers either continue to use conventional finance or avoid all forms of organized finance rather than using the Islamic finance industry).

110. *See id.* at 7 (analyzing the mostly negative responses of different Islamic schools of thought to various finance transactions); *infra* subpart V(B) (illustrating the range of responses to one Islamic financial instrument by various scholars, organizations, and nations).

111. *Savings and Souls*, *supra* note 98, at 82.

cohesive message and streamlined approach. However, one area in need of change that has been largely neglected is the possibility of creating a standardized curriculum for individuals who will sit on or advise Shari'a boards, whether it be at the institutional, national, or international-agency level.

Only in the past few years have educational institutions responded to what has been called one of the sector's greatest challenges—the need for more specialists and experts on Islamic finance.[112] For example, the University of Reading in England recently "launched a master's degree in investment banking and Islamic finance, making it one of a growing number of universities in the west to offer a postgraduate course in a sector that has boomed in the Muslim world in recent years."[113]

As these formalized programs emerge and grow, it is important to structure them in a way that is most useful to the field and its long-term goals. First, educational requirements for sitting on and advising Shari'a boards and leading, managing, or engaging in ratings of Islamic financial institutions should be standardized to include both business and Islamic-religion-and-history curriculums. Though it appears that the newly developed programs at some institutions currently incorporate these issues in the course of study, it is less clear whether there is consistency among institutions. Though a strict version of standardization would be difficult, if not impossible, organizations such as the IIRA and AAOIFI could begin shaping how future Shari'a compliance will look by creating, funding, and publicizing Shari'a-compliant financing as a course of study. One option for this would be for these reputable and well-known agencies to allow universities in different countries to grant certification to engage in Shari'a-compliance evaluation. This way, these leading agencies would be able to set uniform requirements, curriculum, and accepted practices that the next generation of Shari'a scholars would adhere to.

One criticism of this proposal might be that making educational standards stricter could decrease the pool of individuals eligible and qualified to participate in shaping Islamic financial policy and products. However, it is equally feasible that a formalized program might actually increase the number of people interested in and eligible for careers in Shari'a law and Islamic finance. Additional benefits might include making the process for selection of Shari'a board members more transparent and democratic. The individual programs as instituted by universities could be managed and the curriculum controlled largely by the current leaders of the international Islamic financial institutions and agencies. Also, individuals who receive formalized training in Shari'a-compliant financial systems would be eligible to do more than merely serve as Shari'a board members. They would be

112. Shyamantha Asokan, *New Class for Islamic Finance*, FIN. TIMES, Jan. 19, 2009, at 13.
113. *Id.*

qualified to serve in management positions even in non-Muslim countries and for non-Islamic companies.

Though several of the international Islamic financial agencies have programs in place for training individuals for careers in Islamic finance, these sparse programs focus more on the basics of the system rather than on the larger overarching policy.[114] While these programs are a step in the right direction, the programs as they are currently administered do not mirror a university-level educational program. Instead, many of the opportunities for formalized and even certified Islamic-law training are more informal, of relatively short duration, and with relatively few restrictions on eligibility.[115] My proposal is to create a more formalized program with long-term training in the style of traditional university education systems. This would be an improvement over the current systems because of the opportunity to expose future leaders to a standardized curriculum. This might create a greater likelihood for consensus among leaders and promote a more uniform worldview. Having consistent standards across geographic and cultural boundaries may help mitigate the differences that have proved divisive in the past.

Some of the admitted challenges of such a proposal include the time lag for implementation and the variations that would inevitably still remain even after the creation of a formalized educational program.

B. Avoidance of Shari'a Scholar Conflicts

"The OIC is racing to develop common ethical standards for Shari'ah advisory boards and to set up training programs that can staff these boards with certified experts in Islamic finance."[116] Streamlining the educational requirements for Shari'a board membership and other leadership positions in Islamic financial institutions also advances a related goal: reducing the opportunity for conflicts of interests among leaders in the Islamic financial community. In 2009, the top five Islamic scholars held nearly a third of all 956 available Shari'a board positions, and "the top three each sit on more than 60 boards each."[117] It is arguable that by increasing funding and

---

114. *See, e.g.*, Islamic Fin. Serv. Bd., *supra* note 65 (noting that the IFSB "coordinates initiatives on industry related issues, as well as organises roundtables, seminars and conferences for regulators and industry stakeholders"); Islamic Int'l Rating Agency, *supra* note 67 (explaining that the IIRA holds seminars to educate individuals about Islamic rating analysis).

115. *See, e.g.*, Accounting & Auditing Org. for Islamic Fin. Insts., Certified Shari'a Adviser and Auditor (CSAA) Program, http://www.aaoifi.com/csaa2.html (offering a certification program that includes education in the application of Islamic jurisprudence to Islamic finance); Islamic Fin. Servs. Bd., *supra* note 65 (offering workshops and seminars about Islamic finance and related laws, and opening most of these short educational events to anyone who registers).

116. Bianchi, *supra* note 92, at 575 (citing Rifaat Ahmed Abdel Karim, Address at the Islamic Financial Services Board 3d Summit on Aligning the Architecture of Islamic Finance to the Evolving Industry Needs (May 17–18, 2006)).

117. Robin Wigglesworth, *Scholars: Sharia Compliance Rulings Reverse Trend*, FIN. TIMES, Dec. 7, 2009, *available at* http://www.ft.com/cms/s/0/8ec1abb4-e2be-11de-b028-00144feab49a.html.

opportunities for universities to offer more Islamic finance certification programs, the number of Islamic finance experts will continue to grow in both Muslim and non-Muslim countries. However, it might also be prudent and effective for international Islamic finance regulatory bodies to implement conflicts-of-interest guidelines that are uniform across jurisdictions and must be adhered to by Islamic financial institutions and bodies.

The problem as it now exists is that the number of Shari'a law and financial-compliance experts is small. Therefore, all companies and financial institutions needing to meet their Shari'a board standards of having at least three Islamic law experts are required to compete for the valuable time of the same select individuals.[118] This leaves open the opportunity for conflicts of interests when those select scholars may have incentives that are not completely in line with conventional Islamic-finance theory.

For instance, the conflicts in this setting can be compared with those of U.S. boards of directors, which Sarbanes–Oxley (SOX) attempted to address.[119] SOX sought to make boards more independent from management and reduce the number of activities in which they may have a perceived conflict of interest.[120] Likewise, the problems with conflicts of interest on Shari'a boards might be addressed in the same way: with sweeping and enforceable laws that mandate certain measures to prevent conflicts of interest. The most problematic conflicts might stem from the number of loyalties and obligations that the world's top Shari'a law experts have via their participation in so many companies and other financial institutions.

To address this issue, new legislation might seek to limit the number of boards on which individuals can sit. It might also be modeled on SOX and similar acts by requiring that a certain number of board members be independent of management and having more wide-ranging disclosure of potential conflicts.

Critics may argue that such a proposal requires the passing of legislation-like guidelines that face the same problems as the other guidelines in the Islamic financial arena: enforcing adherence across national and cultural boundaries. However, because some agencies already exist and their "stamp of approval" instills confidence in investors, requiring the AAOIFI or the IIRA to oversee and enforce such new rules may help mitigate this danger.

---

118. Miller & Baylis, *supra* note 23.

119. *See* Christine Walsh, *Ethics: Inherent in Islamic Finance Through Shari'a Law; Resisted in American Business Despite Sarbanes-Oxley*, 12 FORDHAM J. CORP. & FIN. L. 753, 754, 773 (2007) (explaining that SOX tries to foster ethical business and eliminate conflicts of interest between auditors and management by requiring enhanced disclosure, increased accountability, and independent review and monitoring of companies).

120. *Id.*

C. *Changes to Ratings System*

The IIRA, as discussed above, created a ratings agency that has proven highly successful over the past four years. Part of its appeal is (1) that it is the only entity that offers such a system specifically tailored to Islamic financial institutions and investors, and (2) it offers a "Quality Rating" that measures the degree of compliance with Shari'a law.[121] Slight modifications to this ratings system may work to create a more investor-friendly environment and, at the same time, advance the overall goal of streamlining Islamic leaders' and scholars' opinions on certain products.

These alterations would include adding a type of peer-review process to the Quality Rating calculation. If companies could rate each others' Shari'a-compliance levels, that may benefit both companies and investors. First, companies who engage in Shari'a financing are more familiar with the intricacies of the business than outsider agencies or evaluators. Second, if companies were encouraged to be members of a peer-review-style model of Shari'a-compliance rating, they would have incentive to act reasonably and rate truthfully because their own interests might be at stake. This type of approach might give investors a more accurate picture of the compliance level of potential companies and institutions.

V. Case Study: *Tawarruq*

A. *A Closer Look at* Tawarruq

Though a rudimentary form of *tawarruq* has existed for many years, the modern form of *tawarruq* is an Islamic finance instrument that was developed in Saudi Arabia less than a decade ago.[122] As discussed briefly above, *tawarruq* can be generally described as an arrangement involving the purchase of an asset at an agreed price and a subsequent sale of that asset to a third party for monetary gain.[123] A common illustration of a *tawarruq* arrangement might look as follows:

- *F* (financial institution) buys ten tons of iron for $20,000,000 from the international market (1 ton = $2,000,000).
- *F* then offers to *C* (client) a ton of iron for $3,000,000 to be paid in installments within ten years.
- At the same time *C* buys the iron, *F* offers to sell (through itself or via another agent) *C*'s iron on behalf of *C* in the international market for $2,000,000 (the same price *F* paid originally).

121. *See supra* notes 67–69 and accompanying text.

122. Al-Shalhoob, *supra* note 54, at 2.

123. Mutalip & Hussin, *supra* note 25, at 1.

- Result: *F* credits $2,000,000 in *C*'s account.[124]

Tawarruq has proven controversial among leading Islamic scholars, and there is an ongoing debate in the Islamic financial community over its Shari'a-compliant status.[125] Though highly popular when first introduced, *tawarruq* has been the subject of criticism from a small but vocal group of Islamic financial scholars.[126]

B. Tawarruq *as an Illustration of Islamic Finance Obstacles*

Among the more common complaints is that *tawarruq* has the potential to create a debt larger than the cash that it transfers.[127] Critics thus allege that despite involving the purchase and sale of real assets, *tawarruq* still violates Islamic principles. It is argued that a single asset can enable multiple *tawarruq* arrangements, essentially severing the tie to the "real sector of the economy."[128] This system of *tawarruq* treads dangerously close to conventional financial methods that freely allow the selling and trading of debt. Opponents also argue that as a matter of public policy, the Islamic community should resist *tawarruq* financing even if the arrangement is construed to be technically Shari'a compliant.[129] The mere fact that it appears to encourage inequity, inefficiency, and high risk is enough for some to deem it a disfavored method.[130]

Opponents to *tawarruq* typically rely on *hadith* directives to support their argument that such an arrangement is *haraam*. First, one *hadith* warns against having "two conditions relating to one transaction" in a financial contract—a phrase whose meaning is not altogether clear.[131] Some argue that this means that *tawarruq* is impermissible since it involves two transactions in one agreement. However, the more traditional view held by the vast majority of scholars is that this warning acts as a prohibition against uncertainty in contracts, meaning that naming two conflicting terms (such as price) in one contract promotes uncertainty.[132]

---

124. Al-Shalhoob, *supra* note 54, at 2. The given illustration is an edited version of an example provided in Al-Shalhoob's work.

125. MICHAEL AINLEY ET AL., FIN. SERVS. AUTH., ISLAMIC FINANCE IN THE UK: REGULATION AND CHALLENGES 18 (2007), *available at* http://www.fsa.gov.uk/pubs/other/islamic_finance.pdf.

126. *See, e.g.*, Siddiqi, *supra* note 39, at 1 (stating that the harmful consequences of *tawarruq* are much greater than the benefits generally cited by its advocates).

127. *Id.*

128. *Id.* at 3.

129. *See* ISLAMIC FIN. PROJECT, HARVARD LAW SCH. ISLAMIC LEGAL STUDIES PROGRAM, TAWARRUQ: A METHODOLOGICAL ISSUE IN SHARI'A-COMPLIANT FINANCE 5 (2007), http://ifptest.law.harvard.edu/ifphtml/ifpseminars/WorkshoponTawarruq.pdf (reporting that although many Shari'a scholars do not dispute the permissibility of *tawarruq*, many also support restrictions on *tawarruq* practice at the institutional level).

130. Siddiqi, *supra* note 39, at 6.

131. Al-Shalhoob, *supra* note 54, at 9.

132. *Id.*

Second, just as many argue that *tawarruq* should be disfavored on public-policy grounds, some scholars believe that *tawarruq* is, at its core, indistinguishable from conventional finance where interest is permitted. Indeed, some even argue that providing a loan with interest is actually less harmful to the client than in the case of *tawarruq* finance since "the financial institution lends to the client directly instead of engaging in a long procedure to buy commodities and sell[] them again in the same market."[133]

The AAOIFI has taken the position that if the commodity in the transaction is sold back to the original seller (from whom it was purchased on a deferred-payment basis), then the transaction is invalid.[134] However, the transactional structure is deemed permissible if the commodity is sold to a third party.[135] Though these positions are fixed by the AAOIFI, there are numerous variables to the transaction that could make a *tawarruq* agreement less aligned with conventional Shari'a law. Some of these questionable modifications include the case where a bank would appoint someone as an agent to buy the commodity on its behalf and then sell it to himself, or where the *tawarruq* is carried out "through . . . national or international commodity [exchanges], wherein only brokers are doing . . . agency services and the goods always remain where they were without transfer of ownership from the seller to the buyer."[136]

In addition to the AAOIFI, recently two other Islamic councils have formally considered the issue of *tawarruq*'s Shari'a compliance. The Fiqh Academy of the Organisation of Islamic Conference in Saudi Arabia forbade all *tawarruq* transactions outright.[137] Additionally, the Muslim World League issued two rulings: one that permitted *tawarruq* on the condition that there is no sale to the original seller and another that directly forbade the modern practice of having a bank sell commodities in global markets.[138] Meanwhile, the *Syariah* Advisory Council of Bank Negara Malaysia, which is responsible for determining Shari'a compliance of financial methods in that country, deemed *tawarruq* arrangements fundamentally permissible under Islam.[139]

Because *tawarruq* is a relatively new innovation that has produced conflicts among Islamic scholars, this product is useful as an illustration of some of the fundamental problems that arise in the Islamic financial community. Initially, *tawarruq* was widely supported among Islamic

133. *Id.* at 10–11. Al-Shalhoob explains the view that *tawarruq* and conventional financing of a loan with interest (usury in Islamic finance) are functionally equivalent, and that of the two, conventional finance is at least less costly to the borrower. *Id.*

134. AYUB, *supra* note 38, at 350.

135. *Id.*

136. *Id.*

137. EL-GAMAL, *supra* note 54, at 72.

138. *Id.*

139. Mutalip & Hussin, *supra* note 25, at 1.

scholars.[140] However, as different financial institutions and businesses have applied *tawarruq* and created hybrid forms of the arrangement, Islamic scholars have increasingly expressed doubt as to whether *tawarruq* is truly Shari'a compliant.

This reflects the sort of inconsistency that is likely to crop up in the Islamic financial system. Several notable, well-respected Islamic scholars have issued opinions on *tawarruq* that conflict with one another.[141] This conflict among scholars also exists at the national and international organizational level. While Malaysia has shaped policy in support of *tawarruq*, Saudi Arabia has largely forbidden some of *tawarruq*'s most familiar and useful forms.[142] The AAOIFI policy on *tawarruq* is at odds with national policy in some countries.[143] Additionally, some institutions are widely perceived as more "legitimate" (such as the Organisation of the Islamic Conference) than others, leading to further confusion among investors.[144]

C. *A Way Forward: Applying Modification Proposals to* Tawarruq

The broad proposals outlined in Part IV above may be more clearly explained when applied to a particular ongoing controversy in the Islamic financial system, such as *tawarruq* financing. The three proposals as applied to *tawarruq* are outlined below and illustrate the potential usefulness of these targeted reforms.

1. Educational Reform.—Creating a standard curriculum for individuals pursuing careers as Islamic financial experts may help ease the problems associated with new financial products such as *tawarruq*. At a macro level, there is a split between Shari'a boards of various nations and international institutions as to whether the general structure of *tawarruq* is permissible. Then, between those two poles, there are numerous different opinions as to

140. *See* Liau Y-Sing, *Islam Allows Organised* Tawarruq *Asset Sales—Scholar*, ARABIAN BUSINESS.COM, June 4, 2009, http://www.arabianbusiness.com/557758-islam-allows-organised-tawarruq-asset-sales---scholar (recording that the opposition among some scholars to *tawarruq* is a recent development and noting that the majority of scholars still sanction its use).

141. *See id.* (citing instances of disagreement among Islamic scholars regarding the compatibility of the *tawarruq* practice with Islamic values).

142. *See supra* notes 137–39 and accompanying text.

143. *Compare supra* notes 134–36 and accompanying text (describing the AAOIFI policy), *with supra* note 137 and accompanying text (describing the policy in Saudi Arabia), *and supra* note 139 and accompanying text (describing the policy in Malaysia).

144. *See* Juan Solé, *Introducing Islamic Banks into Conventional Systems* 5 (Int'l Monetary Fund, Working Paper No. 07/175, 2007), *available at* http://ssrn.com/abstract=1007924 ("[T]he Islamic Fiqh Academy, inaugurated . . . under the auspices of the Organization of the Islamic Conference, has earned the respect of Muslim scholars around the world."). *But see* Habhajan Singh, *Shariah Scholars Turn to AAOIFI over* Tawarruq, MALAYSIAN RESERVE, June 1, 2009, http://islamicfinanceasia.blogspot.com/2009/05/shariah-scholars-turn-to-aaoifi-over.html (reporting that AAOIFI is better regarded than the Fiqh Academy with respect to Islamic finance because its board is comprised of financial experts rather than experts across various fields).

the permissibility of slight variations on the *tawarruq* structure. While some degree of inconsistency in opinions is to be expected along the spectrum according to the structure's characteristics, the fundamental disagreement over the core structure poses a significant problem for investors, financial institutions, and companies that wish to remain Shari'a compliant.

A uniform educational system that emphasizes core principles and offers thorough analysis of *hadith* might result in the next generation of Shari'a law experts approaching financial innovation in a more unified manner. Such education would necessarily have to promote somewhat broad, general concepts that advance dominant approaches to Shari'a interpretation. For instance, in the case of *tawarruq*, the above-mentioned *hadith* restricting transactions with "two conditions" is largely considered not to forbid contracts involving two transactions but instead to discourage vagueness in contract drafting.[145] If this popular interpretation of the *hadith* message could be clarified at a theoretical level and taught as the general standard, then a central conflict among scholars might be resolved.

However, this is only feasible at the broad and general level of interpretation. To continue with the same example, even if the next generation of scholars were educated to believe that vagueness in contracting is bad and that contracts are not necessarily limited to a single transactional issue, conflicts would inevitably remain with regard to degree. Around the margins and along the spectrum of permissibility, there would predictably be conflicting authority. However, a unified stance as to the permissibility of the general *tawarruq* structure might encourage greater investment and use of the method. Investors and financial institutions would be reassured that *tawarruq* agreements are in compliance with Shari'a law and could at that point assume whatever degree of risk they desire by varying their contractual preferences. In the current state of affairs, parties are unsure even as to whether the transaction could *ever* be made to comply with Shari'a law under any set of circumstances.

This type of educational reform that focuses on emphasizing general agreement on fundamental principles could potentially influence the structure of the system's overall regulatory apparatus. A future generation of scholars with shared understandings of Islamic finance fundamentals, despite their differences, might be more capable of shaping an umbrella organization to whose guidelines all other Shari'a boards would willingly adhere. Such guidelines would inevitably be general and incomplete, but they could serve to insulate individual institutions and individuals from responsibility for engaging in transactions of which a small minority of scholars disapprove.

2. Avoiding Conflicts.—Limiting the number of boards on which individuals may sit should increase the number of opportunities for new

145. *See* Al-Shalhoob, *supra* note 54, at 9 (explaining the competing interpretations espoused by Islamic scholars concerning the "two conditions" requirement in the *tawarruq*).

scholars. It should also limit the impact of any one individual's opinion. If scholars were required to disclose publicly all their affiliations in the Islamic financial community and are limited in the number of board positions they may hold, then the pool of Islamic scholars would likely grow and become more equalized. In the case of *tawarruq* and other similar products, such changes may result in a more accurate gauge of the majority opinion. Because the current leaders and decision makers on prominent Shari'a boards are pulled from such a small group, any single group's prohibition has a disproportionate impact on the instrument's viability. Such an impact was especially pronounced in the case of Usmani's cautionary comments regarding *sukuk* issuances.[146] Thus, an emphasis on strengthening conflicts laws and increasing transparency with regard to Shari'a scholars' personal and professional dealings should lead to a more accurate measure of expert opinion and consensus. This may enable *tawarruq* arrangements to avoid the fate of *sukuk*.

3. New Ratings Methods.—If companies and institutions were required to rate the practices of other institutions in terms of Shari'a compliance, this may have a moderating effect on Islamic financial policy. This effect might be similar to that of increased conflicts-of-interest monitoring, which works to give institutions and individuals a better indication of where majority opinion lies. Additionally, with products such as *tawarruq* that have many variations, providing a type of peer-review rating might better enable parties to draw the line as to which conditions will push an acceptable product into ambiguous territory where transactions may be altogether impermissible. In other words, the availability of peer review might make it more obvious where public and expert opinion cluster with regard to Shari'a compliance and which factors most affect their findings. Peer review would allow institutions to judge various forms of *tawarruq*, forcing them to draw lines that they might not be in a position to draw with regard to their own institution's business.

VI. Conclusion

In the post-financial-crisis world economy, there will be opportunities for new financial products to develop in response to both the crisis and its regulatory fallout. The principles of Islamic finance may be just what are needed in these trying economic times. With an emphasis on morality, fairness, and aversion to excessive risk, these foundational premises can teach conventional finance a few helpful lessons. The key will be to find ways to translate Islamic finance's underlying goals and methods to a larger audience in the non-Muslim world.

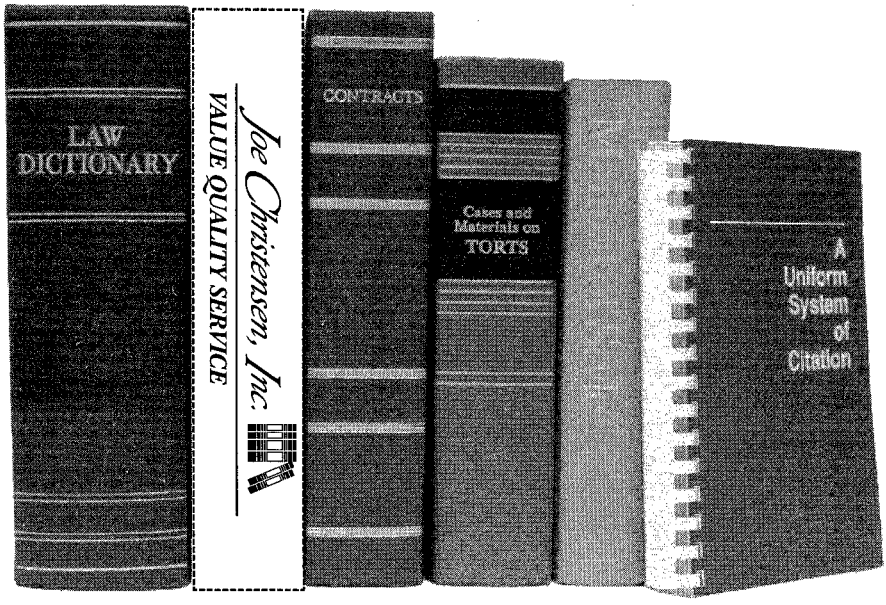
146. *See supra* notes 82–91 and accompanying text.

These proposals for streamlining educational standards, reducing conflicts among Shari'a law leadership, and making Shari'a-compliant investing more investor friendly might offer a way to move in that direction. Though these proposals face the same hurdles that other developments in Islamic finance have faced in the past, the Islamic financial system has proven resilient and responsive in the four decades since its inception.

Presumably, those involved in Islamic investing seek to grow the industry and provide Muslims the opportunity to invest with the knowledge that their religious commitments have not been compromised. If the restrictions on such investments are such that they discourage investment in Islamic financial institutions and products altogether, the worldwide Muslim community as a whole suffers. Therefore, any movement to change the Islamic financial system should acknowledge that the restrictions in place do not exist to prevent the growth of Islamic finance or to make it more difficult. The purpose of such restrictions is quite the opposite. The restrictions exist to achieve the religious goals that form the basis of Islamic investing. Thus, proposals that aim to grow the Islamic financial system must not do so at the risk of trampling these fundamental goals. For this reason, these above-mentioned proposals are potentially viable reforms, aiming to increase the overall investment in the Islamic financial system without contradicting the fundamental purposes of the religious restrictions that form the system's base.

—Holly E. Robbins

* * *



We Complete the Picture.

In 1932, Joe Christensen founded a company based on Value, Quality and Service. With 77 years of printing experience, 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


TEXAS
TLR
LAW REVIEW